Electronic Supplementary Material (ESI) for ChemComm.

Topology *versus* porosity: what can reticular chemistry tell about free space in metal-organic frameworks?

Alexander P. Shevchenko,[abc] Eugeny V. Alexandrov,[abc] Andrey A. Golov,[ab] Olga A. Blatova,[ab] Alexandra S. Duyunova,[a] Vladislav A. Blatov*[ab]

[a]*Samara Center for Theoretical Materials Science (SCTMS), Samara University, Ac. Pavlov St. 1, 443011 Samara, Russian Federation.*

[b]*Samara Center for Theoretical Materials Science (SCTMS), Samara State Technical University, Molodogvardeyskaya St. 244, 443100 Samara, Russian Federation.*

[c]*Samara Branch of P.N. Lebedev Physical Institute of the Russian Academy of Sciences, Novo-Sadovaya St. 221, 443011 Samara, Russian Federation.*

Supplementary Information

Computational methods

Simplification procedure is a method of splitting the coordination network to building units (complexing atoms, polynuclear complex groups, ligands) and representing the whole network as the *underlying net*, which if the net of the centres of the building units connected according to their connection in the initial network (Fig. S1).[1] The underlying net can be considered as encoding of the assembly of the coordination network from the building units. There are rigorous algorithms to separate the building units using chemical and topological criteria.[2] The connectivity of the building units and the underlying net as a whole is described by a number of topological descriptors; those used in this work can be arranged in three groups.

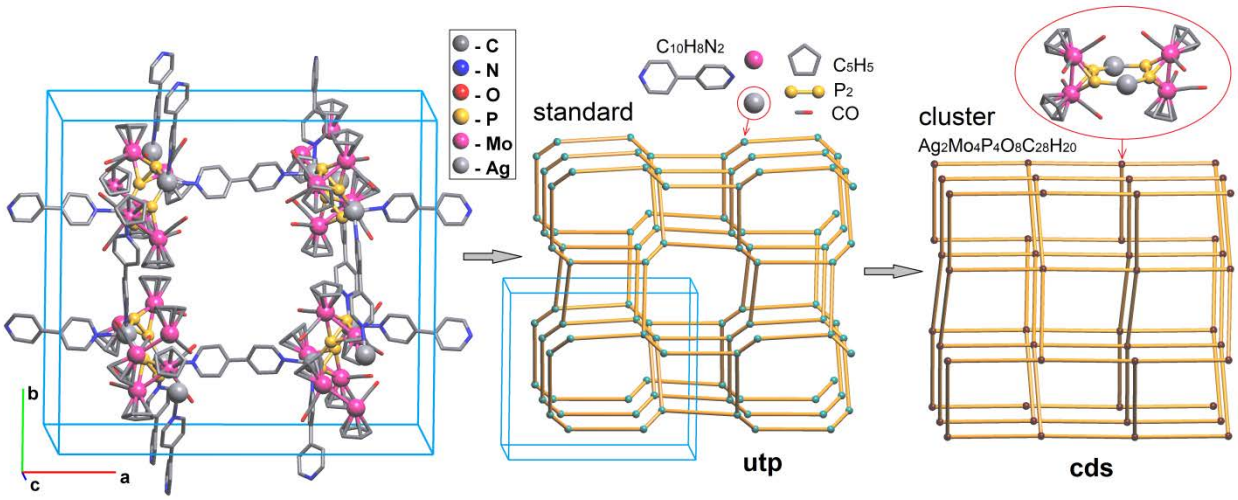

Figure S1. Simplified (underlying) nets of a coordination network (left) in the standard (centre) and cluster (right) representation for the crystal structure of [(4,4'-bipyridine)$_2$(diphosphide)$_2$(CO)$_8$(cyclopentadienyl)$_4$Mo$_4$Ag$_2$][(OC(CF$_3$)$_3$)$_4$Al)]$_2$·CH$_2$Cl$_2$ (YILMEO).[3] Hydrogen atoms and disordered anions [Al{OC(CF$_3$)$_3$}$_4$]$^-$ are not shown.

The result of the simplification procedure is a *representation* of the initial coordination network. The same network can be represented in different ways depending on which building units are chosen for the simplification. In this work, we use two kinds of representation of the coordination network:[4] (i) *standard representation*, where the building units are metal atoms and organic ligands, and *cluster representation*, where the building units are polynuclear complex groups or metal clusters (if any) and linkers between them; it also corresponds to the *single node* method. If the structure can be represented in the cluster representation, this representation is chosen as final in the tables of our knowledge database. Thus, although the same structure can be represented in different ways, all these representations are built according to rigorous algorithms, and if we consider the same type of representation for all structures under consideration there is no ambiguity in choosing the nodes of the underlying net.

An alternative topological description of the crystal structure is *tiling*. If the underlying net describes the topology of the coordination network, tiling deals with the topology of the free space in the network. Tiles are generalized polyhedra, which can contain curved faces and 2-coordinated vertices (Fig. S2). Usually, tiling is being built for the simplified (underlying) net following a rigorous algorithm.[5] In most cases, a unique *natural tiling* can be constructed with this algorithm; the natural tiles correspond to the smallest cages of the network, while the shared faces of the tiles mimic the channels between the cages. Not any net admits tiling; moreover, natural tilings can be constructed not for all nets that admit tilings. In fact, this means that the smallest cages cannot be selected in such nets. The tiles are designated by their face symbol, which contains a sequence of the face sizes and the numbers of the faces of a

particular size, for example the face symbol [$6^2.8^2$] means tile bounded by two 6-membered rings and two 8-membered rings. The topological types of tiles are rigorously determined by the graph of their edges; for the natural tiles of the zeolite frameworks the *t-xxx* symbols are applied, for example *t-kaa*, which has the face symbol [$6^2.8^2$] (Fig. S2).[6]

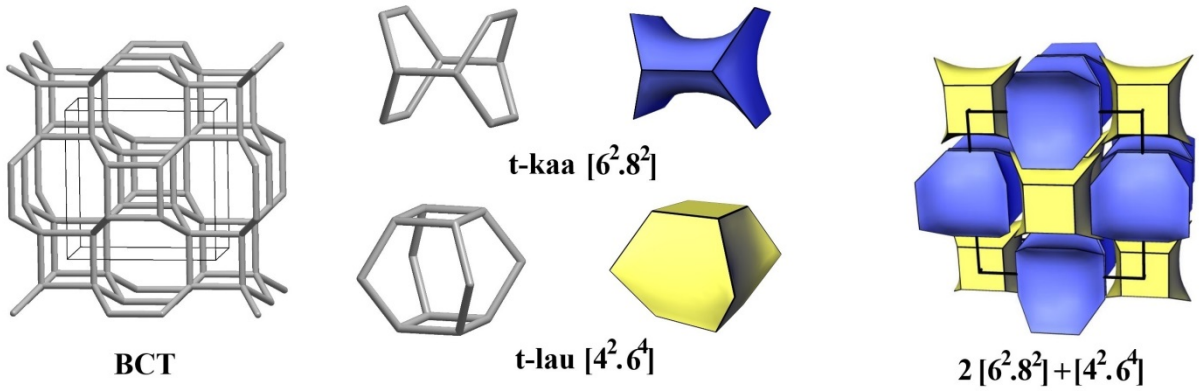

Figure S2. An example of natural tiles and tiling for zeolite BCT.

Some of the most commonly used algorithms for calculation of topology and geometry of free space in crystal structures are based on the analysis of the graph (Voronoi net) obtained by the Voronoi decomposition of the crystal space. Previously we propose a number of such algorithms, and applied them to the zeolite frameworks.[7] Within this approach, the channels connectivity is fully characterized by the Voronoi net. The radii of cavities and channels are determined by the distances from vertices and edges of the Voronoi net to the closest points of the van der Waals surface of the structure. The shape, volume, and surface area of pores are calculated from geometries of the polyhedra constructed for the vertices of the Voronoi net. The total porosity is computed as $P = (V_{cell}-V_{vdw})/V_{cell}$, where V_{cell} is the volume of the unit cell and V_{vdw} is the part of the unit cell volume occupied by the atoms, which are represented by intersecting spheres of van der Waals radii. We use the Alvarez system of van der Waals radii.[8] The volume of an atom is calculated as the volume of Voronoi polyhedron, whose intermolecular faces are moved to the van der Waals surface of the atom. The free space descriptors used in this work are given below.

Topological descriptors

1. Local coordination descriptors include:

(i) *Coordination numbers* (CNs) of complexing metal atom and ligand are equal to the number of bonds or links, which are formed by the given structural group. Polynuclear complex group or polydentate molecular ligand can have two different CNs: the total number of bonds, which their donor atoms form with other building units (CN_{bond}), and the number of links with the building units in the underlying net (CN_{link}); the latter CN is equal to the number of building units connected to the group or ligand (Fig. S3). The types of We used the for the classification of the coordination figures have been determined with the library described in [9].

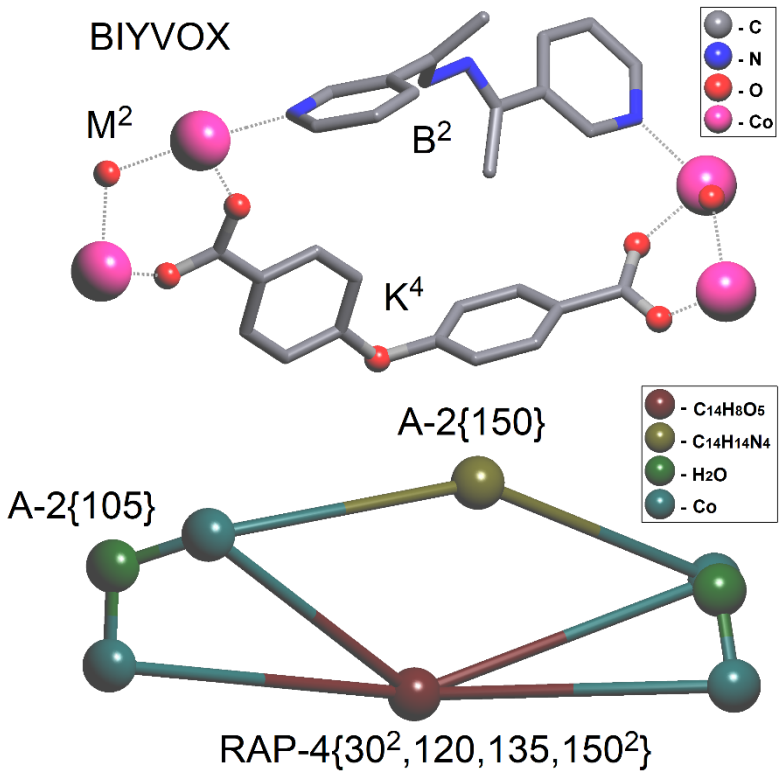

Figure S3. Coordination numbers, modes (top) and figures (bottom) of ligands in crystal structure [(3,3'-[hydrazine-1,2-diylidenebis(eth-1-yl-1-ylidene)]dipyridine)$_2$(4,4'-oxydibenzoato)$_2$Co$_2$]·DMAA (BIYVOX).[10] Hydrogen atoms are omitted for clarity. The types of coordination figures are designated according to Hartshorn,[11] and extended with the sequence of valence angle values in ascending order with a step of 15 degrees. Thus, symbol A-2{150} denotes a bent coordination figure with an angle of 150 degrees; symbol RAP-4{30^2,120,135,150^2} corresponds to a quasi-rectangular coordination figure.

(ii) *Coordination mode* of ligand describes how the ligand uses its donor atoms for the connection with other building units. We use the notation L^{mbtkpg} proposed in [12], where *L* = M, B, T, K, P or G depending on the number of the donor atoms of the ligand (1, 2, 3, 4, 5 or 6, respectively), and the integers *mbtkpg* are equal to the numbers of metal atoms connected with 1, 2, 3, 4, 5 or 6 donor atoms (Fig. S3). Obviously, coordination mode characterizes the ligand in the initial coordination network, not in the underlying net.

(iii) *Coordination figure* of structural unit shows the shape of its environment in the underlying net, i.e. the geometrical configuration of the centers of the adjacent structural units (Fig. S3). Coordination figure generalizes the well-known concept of coordination polyhedron; while the latter is always related to an atom, the former describes the coordination of any structural unit. This generalization leads to a much wider diversity of coordination figures compared to coordination polyhedra; special algorithms were developed to discriminate different forms of coordination figures.[13,14]

2. Overall topology descriptors include:

(i) *Dimensionality* of the coordination network or the corresponding underlying net; there can be molecular (0D), chain (1D), layered (2D) or framework (3D) structural groups. Mathematical rigor requires using the term *n-periodic* instead of *n-dimensional*, but the latter term is conventional in coordination chemistry.

(ii) *Underlying topology* characterizes the overall connectivity of the underlying net; there are several methods to determine the underlying topology; the nets, which have the same topology, can be matched without breaking or creating edges, adding or removing nodes. The coordination polymers,

which have the same underlying topology, belong to the same topological type. There are several nomenclatures to assign the names for topological types:

(a) *RCSR* three-letter symbols;[15] for example, **pcu** designates primitive cubic net;

(b) *ToposPro N*D*n* symbols,[2] where *N* is the set of coordination numbers of all independent nodes of the net; D is one of the letters C, L, or T, which designate the net dimensionality (C – chain, L – layered, T – framework); *n* enumerates topologically different nets with a given *N*D sequence. For instance, the symbol 3,4,5T2 denotes the 2nd (by the order) framework net with three 3-, 4- and 5-coordinated independent nodes. For the molecular structural groups the symbols *N*M*k-n* are applied, where k is the number of vertices in the group;

(c) *EPINET sqc*N symbols,[16] where N is an integer, for example, the primitive cubic net is designated *sqc*1;

(d) Fischer's symbols *k/m/fn* for three-periodic sphere packings;[17] the primitive cubic net has the Fischer's symbol 6/4/c1;

(e) Zeolite capital three-letter symbols.[18]

3. Tiling descriptors include:

(i) *Face symbol of tile* bears the information about the size and number of the tile faces. For example, hexagonal prism has face symbol $[4^6.6^2]$. Face symbols of all tiles together with their stoichiometric coefficients form the tiling signature (Fig. S2).

(ii) *Topological type of tile* is determined by the edge graph of the tile; the tiles belonging to the same topological type have isomorphic edge graphs. Topological types of all natural tiles in the 252 known zeolite frameworks are stored in the *ToposPro TTT* collection; their names are constructed from a three-letter symbol with the *t-* prefix, for example, *t-toc* corresponds to the tile in the sodalite framework (truncated octahedron with face symbol $[4^6.6^8]$; see also Fig. S2).

(iii) *Geometrical distortion of tile* is estimated by the mean square deviation (Σ) of the vertices of the tile from the standard geometrical form. Both the tile and the form are reduced to the same volume by a similarity operation that is equivalent to reducing of the underlying nets to the same density (3.0 g/cm^3 with assigning a formal molar mass of 1.0 g/mol to all nodes). In this work, the standard geometrical forms of tiles were obtained from the most symmetrical embeddings of the corresponding underlying nets published in the RCSR database.[15] An increase in the variety of geometric shapes of tiles from left to right confirms a decrease in the rigidity of the net in the same sequence. The larger the part of tiles with a small Σ (**sod** or **nbo**, Fig. S4), the more rigid is the corresponding topological type of tile and the underlying net, and *vice versa*, a large number of tiles with large Σ indicates a deformable net (**cds**, **lvt** or **sra**, Figs. S4, S5).

(iv) *Connectivity Index* (CI) is calculated as the ration of two integers: the number of tiles edges (*e*) and the number of tile vertices (*v*). For example the tiles in the **dia**, **sra** and **lvt** nets have *e* = 12, 18, 20 and *v* = 10, 14, 16, thus giving rise to CI = 12/10 = 6/5, 18/14 = 9/7 and 20/16 = 5/4, respectively (Fig. S5). The integers *e* and *v* are interrelated with the number of tile faces (*f*) by the Euler equation: $v - e + f = 2$. This equation predetermines the minimal limit for CI: $v = e$, CI = 1, since on this condition $f = 2$, which is impossible for 3D polyhedra. CI = 3/2 is a minimal value for a tile, where each vertex is locally stable, *i.e.* its position cannot be changed without changing the positions of the neighbouring vertices and/or the edge lengths.

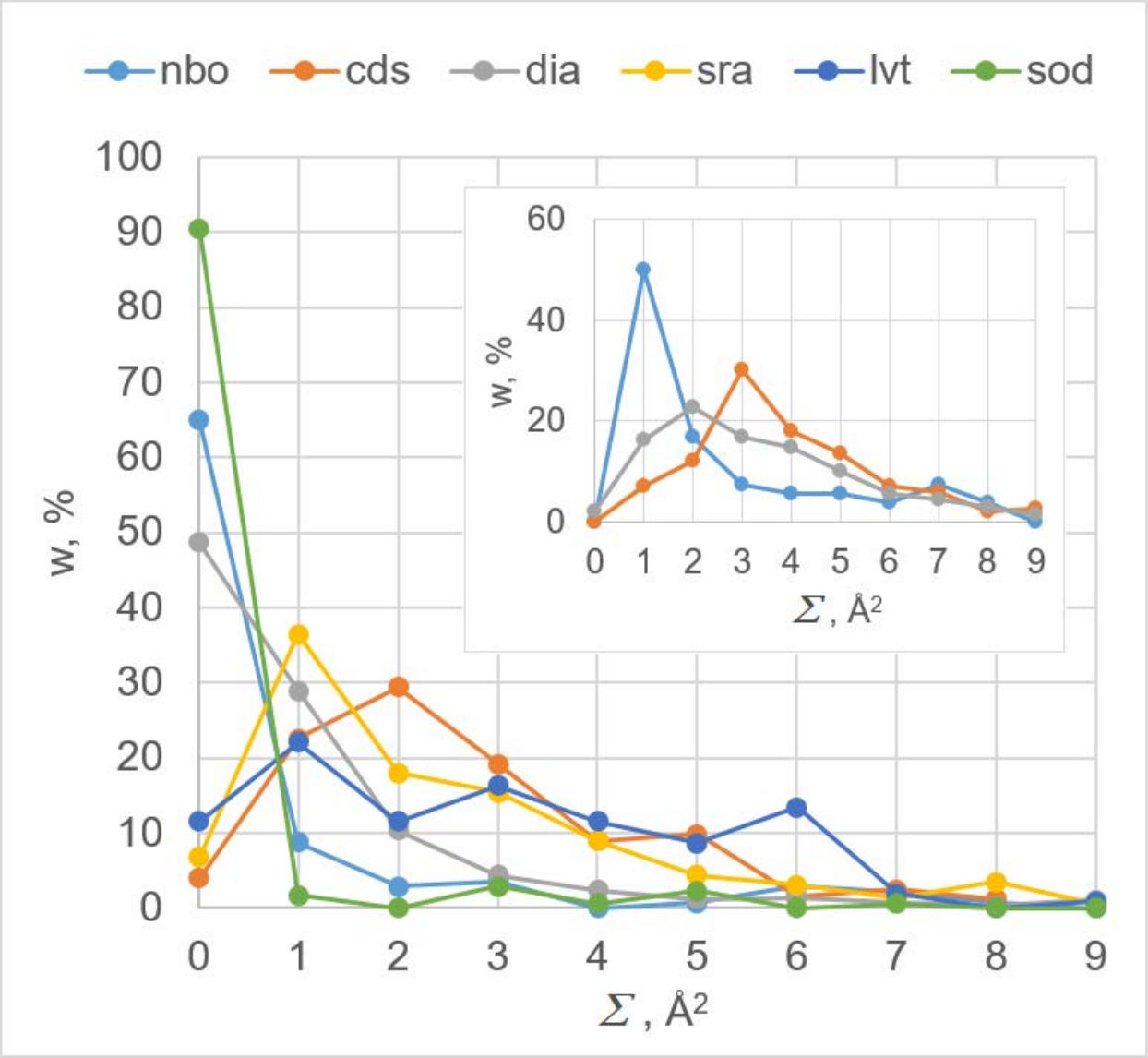

Figure S4. Dependence of the occurrence of tiles (w) on their degree of distortion Σ for various topologies in the single underlying nets of coordination polymers. The inset contains the same dependencies but for 50, 129, and 1034 structures containing interpenetrating arrays of nets with the **nbo**, **cds** or **dia** topology.

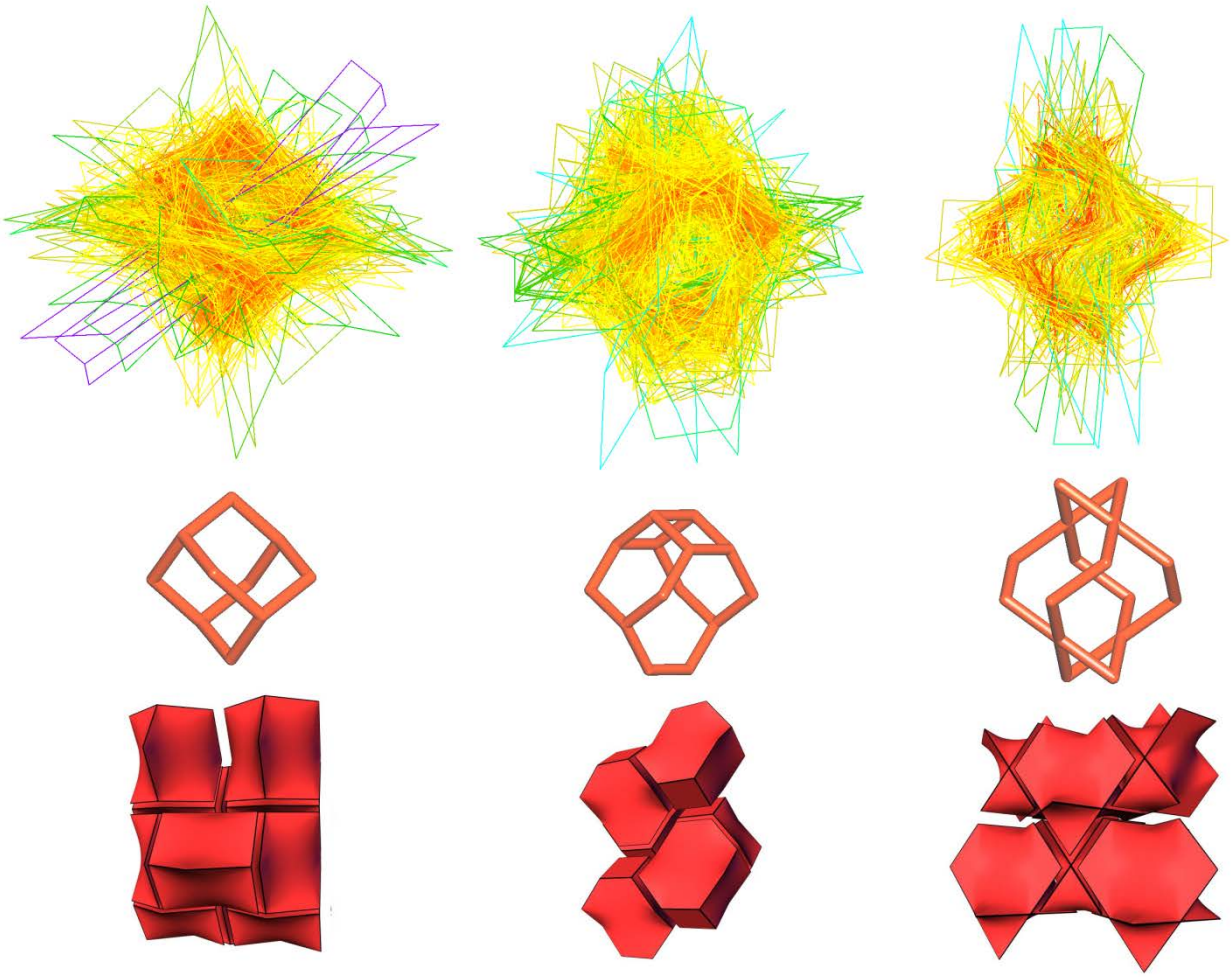

Figure S5. (Top) The overlapping diagrams for geometrically distorted tiles in 745 **dia** (927 tiles; CI = 6/5; Σ = 1.15 Å^2; left), 411 **sra** (455 tiles; CI = 9/7; Σ = 2.52 Å^2; centre) and 99 **lvt** (104 tiles; CI = 5/4; Σ = 3.09 Å^2; right) underlying nets of coordination polymers compared with the most symmetrical shapes of the tiles (middle) and the corresponding natural tilings (bottom). The colouring of the distorted tiles depends on their geometrical distortion (Σ), from red (small distortion) to blue (large distortion).

The simplification procedure, determination of the natural tiling and all descriptors described above are realized in the program package *ToposPro*,[19] which was used in this work. To determine the topological types of the underlying nets and tiles we used the *ToposPro TTD* and *TTT* collections.

Geometrical and topological descriptors of the free space in crystal

We used the following descriptors to characterize the free space:

(i) Total porosity (*P*) that is the part of the structure volume occupied by both accessible and isolated pores.

(ii) Radius of the largest included sphere (R_i) that can be embedded in the structure without intersection with the van der Waals surface of atoms. In other words, R_i corresponds to the largest cavity in the structure. For example, the structure of 2-fold interpenetrated zinc metal-organic framework $Zn_3(BTPCA)_2(H_2O)_3 \cdot 3py \cdot 3DMSO$[20] (H3BTPCA = 1,1' ,1''-(benzene-1,3,5-triyl)tripiperidine-4-carboxylic acid) has large tetrahedral cages (R = 4.5Å), which encapsulate pyridine and dimethylsulfoxide molecules. However, all these cages are isolated since the interpenetrated frameworks lock the windows of the cages.

(iii) Radius of free sphere (R_f) is the radius of the largest spherical probe that can migrate through the widest 1-, 2-, or 3-periodic channel systems. The descriptor can be used for the evaluation of the maximal size of molecules that are able to move through the structure.

(iv) Radius of free included sphere (R_{fi}) is the radius of the largest spherical probe that can be included in the widest 1-, 2-, or 3-periodic channel systems. Usually R_{fi} is equal to R_i. These parameters differ from each other only if the largest cavity is separated from the widest periodic system of channels.

(v) Periodicity of the channel system (*PC*).

(vi) Part of the structure volume occupied by the channel system. This descriptor accounts for only the pores that are accessible for the given spherical probe in a periodic channel system. Thus, this descriptor can be used for evaluation of the sorbent capacity with respect to specific molecules.

(vii) Number of independent channel systems (*Z*) accessible for a spherical probe of a given radius.

(viii) Direction of a one-periodic system of channels or direction of the normal to a two-periodic channels system. An example of a metal-organic coordination polymer containing multidirectional channel systems is bio-MOF-100,[21] where 1-periodic channels spread in four different directions. The independent intra-framework channel systems can have even higher periodicity. For example, the structure of 2-nitroimidazolate cadmium hydrate[20] contains two independent non-crossing 3-periodic channel systems.

All of these descriptors can be computed using *ToposPro* and special *Python* scripts according to the algorithms described in [2, 7, 19].

Data analysis

For the statistical analysis, we used the data from the Cambridge Structural Database, which contains more than one million records. The main stages of the data analysis and the extraction of knowledge from the data were described in detail in [9], where the prediction of the dimensionality and topology of the crystal structure was the main problem to be solved. The current paper is devoted to the analysis and prediction of the porosity of 3D frameworks, and this goal affected the conditions for the selection of the structural data, although the stages of the analysis remained the same.

An important step is the preparation of the structural data for the analysis. The structures with strongly disordered frameworks, in which ligands were found to be valence bonded to each other, were excluded from consideration. The extraframework species were removed for the correct estimation of the porosity. Duplicated structural studies were marked in the sample, and one of them was selected as the representative of the group. The structural studies were considered duplicated if their chemical composition, space-group symmetry, and framework topology, accounting for interpenetration, coincided. When constructing statistical distributions, 5735 duplicates were excluded from consideration (see Table S1, column *Doubles*).

We have used the *ToposPro* program package to filter data and to calculate the descriptors of atoms, structural building units, and the structures as a whole. The *ToposPro TTD*, *TTO* and *TTL* collections[2] were used as the sources of additional information on the descriptors. We examined the correlations of the porosity with the following topological and geometric descriptors:

- dimensionality of the coordination framework;
- topological type of the underlying net in the standard and cluster representations;
- number of interpenetrating frameworks;
- topology and geometric shape of tiles;
- dimensionality and radii of the free and included spheres for the channel systems;
- shape of coordination figures of structural building units.

The overlappings of the tiles from different structures were constructed with the *CFShape* program.[14] Before overlapping, the unit cell dimensions of the underlying nets were normalized to the same crystal density.

To search for correlations between overall topology and porosity of the framework (Table S6), we have split the porosity values into four groups (quartiles): Q1 with $P < 0.25$; Q2 with $P = 0.25–0.50$; Q3 with $P = 0.50–0.75$, and Q4 with $P > 0.75$.

Table S3. The occurrence of topological types of all 3D structures (All) and porous coordination polymers (MOFs) for single (Z=1) and interpenetrating (Z>1 is separated from the topology symbol by '#') networks in the standard representation.

Topology	All, Z=1		MOFs, Z=1		Topology	All, Z>1		MOFs, Z>1	
	N	w, %	N	w, %		N	w, %	N	w, %
dia	640	2.71	401	3.19	4#**dia**	224	5.09	126	5.02
pcu	560	2.37	250	1.99	2#**dia**	220	5.00	145	5.78
pts	444	1.88	262	2.08	3#**dia**	196	4.45	117	4.67
sra	373	1.58	236	1.88	2#**xah**	176	4.00	81	3.23
cds	217	0.92	138	1.10	5#**dia**	123	2.79	73	2.91
xah	205	0.87	71	0.56	2#**pcu**	118	2.68	73	2.91
bpq	205	0.87	105	0.83	2#**ths**	103	2.34	33	1.32
srs	201	0.85	138	1.10	2#**fet**	78	1.77	41	1.63
bnn	195	0.82	123	0.98	2#**srs**	77	1.75	43	1.71
ins	172	0.73	107	0.85	2#*sqc*65	76	1.73	36	1.44
mog	160	0.68	107	0.85	2#**cds**	74	1.68	52	2.07
tcs	154	0.65	80	0.64	2#**pts**	64	1.45	36	1.44
nia	149	0.63	53	0.42	3#**xah**	60	1.36	38	1.52
dmc	135	0.57	90	0.72	2#**mog**	55	1.25	36	1.44
4,8T1	125	0.53	38	0.30	6#**dia**	52	1.18	36	1.44
Other	*19712*	*83.36*	*10378*	*82.52*	*Other*	*2708*	*61.49*	*1542*	*61.48*
Total	23647		12577		Total	4404		2508	

Table S4. The occurrence of topological types of all 3D structures (All) and porous coordination polymers (MOFs) for single (Z=1) and interpenetrating (Z>1) networks in the cluster representation.

Topology	All, Z=1		MOFs, Z=1		Topology	All, Z>1		MOFs, Z>1	
	N	w, %	N	w, %		N	w, %	N	w, %
pcu	1698	16.30	916	16.64	2#**pcu**	673	32.80	360	32.70
bcu	589	5.65	312	5.67	3#**pcu**	179	8.72	98	8.90
dia	335	3.22	202	3.67	2#**dia**	133	6.48	73	6.63
rtl	329	3.16	201	3.65	4#**pcu**	56	2.73	23	2.09
flu	205	1.97	112	2.03	3#**dia**	48	2.34	35	3.18
nbo	187	1.79	84	1.53	2#**fsc**	38	1.85	23	2.09
tfz-d	181	1.74	104	1.89	2#**tfz-d**	34	1.66	21	1.91
cds	175	1.68	109	1.98	2#**rtl**	32	1.56	16	1.45
fcu	171	1.64	48	0.87	2#**sit**	31	1.51	17	1.54
flu-3,6-*C*2/*c*	158	1.52	75	1.36	4#**dia**	22	1.07	10	0.91
fsc	138	1.32	68	1.24	2#**hex**	21	1.02	10	0.91
hex	135	1.30	82	1.49	2#**lvt**	21	1.02	10	0.91
rob	130	1.25	88	1.60	2#**the**	20	0.97	11	1.00
pts	127	1.22	88	1.60	2#**hms**	17	0.83	9	0.82
acs	126	1.21	74	1.34	2#**pto**	17	0.83	11	1.00
Other	*5734*	*55.04*	*2941*	*53.44*		*710*	*34.60*	*374*	*33.97*
Total	10418		5504			2052		1101	

Table S5. Occurrence of self-dual nets in coordination polymer structures

Symbol	Standard		Cluster		Symbol	Standard		Cluster	
	N(Z=1)	N(Z>1)	N(Z=1)	N(Z>1)		N(Z=1)	N(Z>1)	N(Z=1)	N(Z>1)
bbr	-	-	-	-	**pte**	-	-	-	-
cbs	-	-	-	-	**pyr**	16	1	39	10
cdq	3	-	3	7	**qtz-x**	-	-	7	1
cds	217	121	175	32	**rtw**	-	-	-	-
ctn	20	1	15	2	**sda**	-	-	3	-
dia	619	898	335	236	**smt**	-	-	-	-
est	1	-	-	-	**srs**	194	111	26	14
ete	-	1	-	-	**sto**	-	-	-	-
fsf	1	-	-	-	**svn**	-	-	-	-
ftw	1	-	32	-	**swl**	-	-	-	-
gsi	4	-	-	-	**sxd**	3	1	10	1
hms	28	37	3	17	**tfa**	5	11	3	-
hst	-	-	1	-	**tfc**	6	1	6	4
lcy	1	-	21	7	**ths**	122	182	1	5
mab	33	-	120	14	**tph**	-	-	1	-
mcf	-	-	-	-	**ttv**	-	-	-	-
mco	-	-	1	-	**unj**	1	-	5	-
mgc	-	-	-	-	**vck**	-	-	-	-
pcu	498	145	1698	915	**vtx**	-	-	-	-

Table S6. The distribution of frequently occurring topologies (standard and cluster) of the single coordination frameworks by quartiles of the porosity.

Topology*	Q1	Q2	Q3	Q4	Topology	Q1	Q2	Q3	Q4
dia	2.06	3.00	3.57	0.39	**alb**-4,8-*Imma*→**lvt**	-	-	0.04	1.97
pcu	2.79	1.56	3.22	0.79	**tbo**	-	-	0.04	1.97
pts	1.47	1.73	2.29	0.79	4,4,4,8T13→**cor**	-	-	-	1.97
srs	0.59	0.61	0.82	3.15	3,3,6,8,8T1→**csq**	-	-	-	1.97
3,4,8T15→**fcu**	0.15	-	0.79	3.15	**tcs**	0.88	0.59	0.36	-
fff→**pcu**	-	-	0.14	3.15	4,16T2→**scu**	-	-	0.64	1.18
4,12T1→**nts**	-	-	0.39	2.76	**sod**	-	0.17	1.11	0.39
3,6,8,8T1→**flu**	-	-	0.32	2.76	**mog**	0.88	0.53	0.25	-
alb-4,8-$P4_2/mmc$→**pts**	-	0.01	0.18	2.76	3,5T1	1.03	0.53	0.04	-
cds	0.74	1.23	0.43	0.39	**fsc**	0.15	0.42	0.89	-
4,4,4,4,4,6,6T1→**muo**	-	-	-	2.76	**flu**	1.03	0.27	0.04	-
bnn	1.32	0.92	0.50	-	**lvt**	0.29	0.52	0.43	-
sra	0.74	1.15	0.46	-	**ths**	0.15	0.45	0.64	-
4,8T24→**nbo**	-	-	0.75	1.57	4,4,6T19	0.88	0.19	0.14	-
bpq	0.15	0.71	1.47	-	4,8T1	-	0.08	1.11	-
nts→**tbo**	-	-	0.25	1.97	**fet**→**pcu**	-	0.47	0.64	-
rtl	1.62	0.46	0.11	-	**fsx**-4,5-*C*2/*c*	1.03	-	-	-
nia	1.32	0.33	0.54	-	**fsy**	-	0.50	0.43	-
5,8T16→**fsc**	-	0.01	0.11	1.97	**ins**	0.29	0.52	0.11	-
ith-3,9-*R*3*m*	2.06	-	-	-	**srs**→**dia**	0.29	0.50	0.11	-

*The topology of the cluster representation (if any) is indicated after the arrow.

Table S7. Abundant underlying nets of coordination polymers with isohedral tilings and their connectivity indices (CI).

Net	CI	Net	CI	Net	CI	Net	CI
dia	6/5	**srs**	15/14	**dmp**	11/10	**hex**	3/2
pcu	3/2	**bnn**	3/2	**ths**	10/9	**lvt**	5/4
sra	9/7	**bpq**	10/7	**bbf**	4/3	**qtz**	7/6
cds	7/6	**tcs**	14/9	**nbo**	4/3		

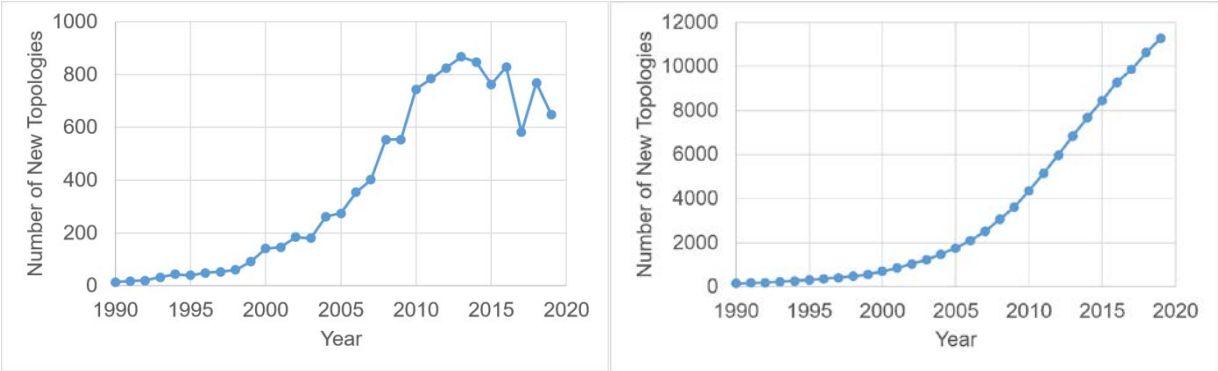

Figure S6. Dependence of number of new topological types of the underlying nets of 3D coordination polymers on the year of the first publication of a structure with this topology (left) and the corresponding integral dependence (right).

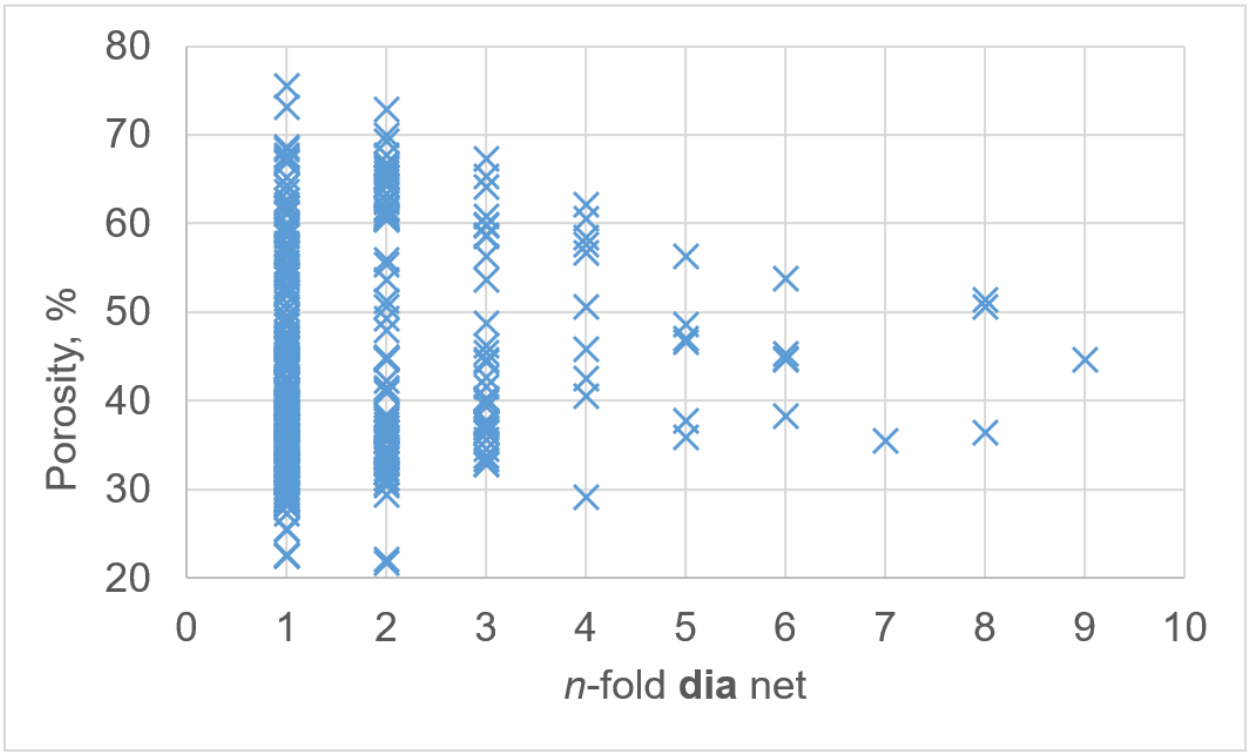

Figure S7. Dependence of porosity on the number of interpenetrating **dia** nets in the crystal structure of a coordination polymer.

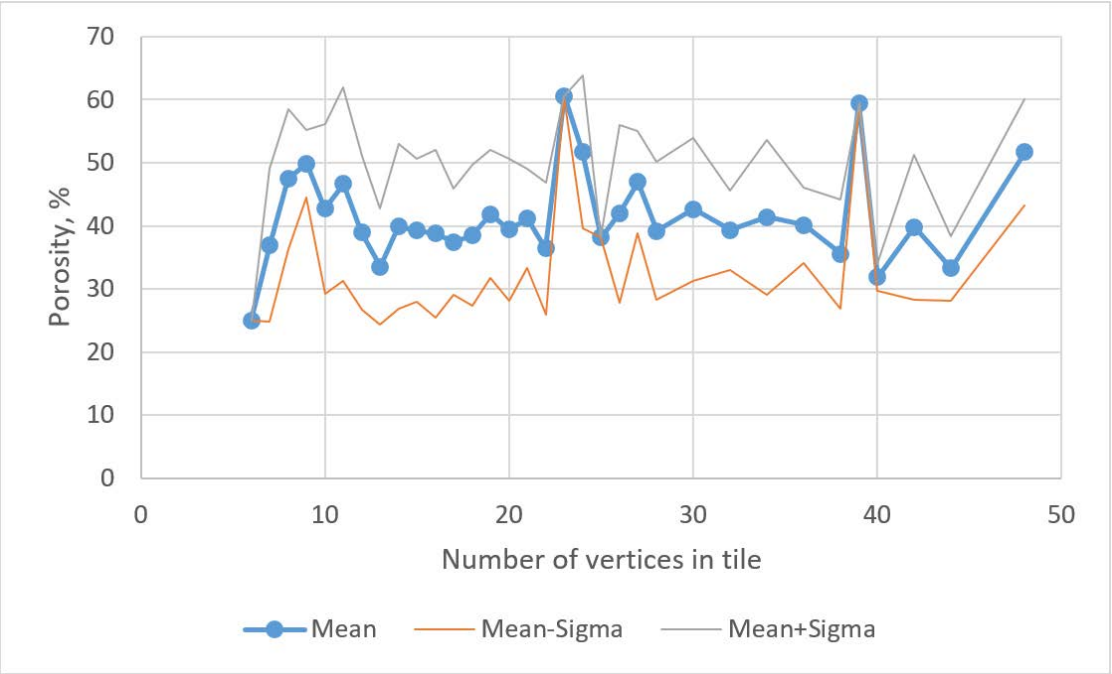

Figure S8. Dependence of the porosity of the coordination polymers on the number of vertices in the tile of the underlying net tiling. Only isohedral (consisting of the single type of tile) tilings are considered. The borders of the standard deviation gap are shown by red and grey thin lines.

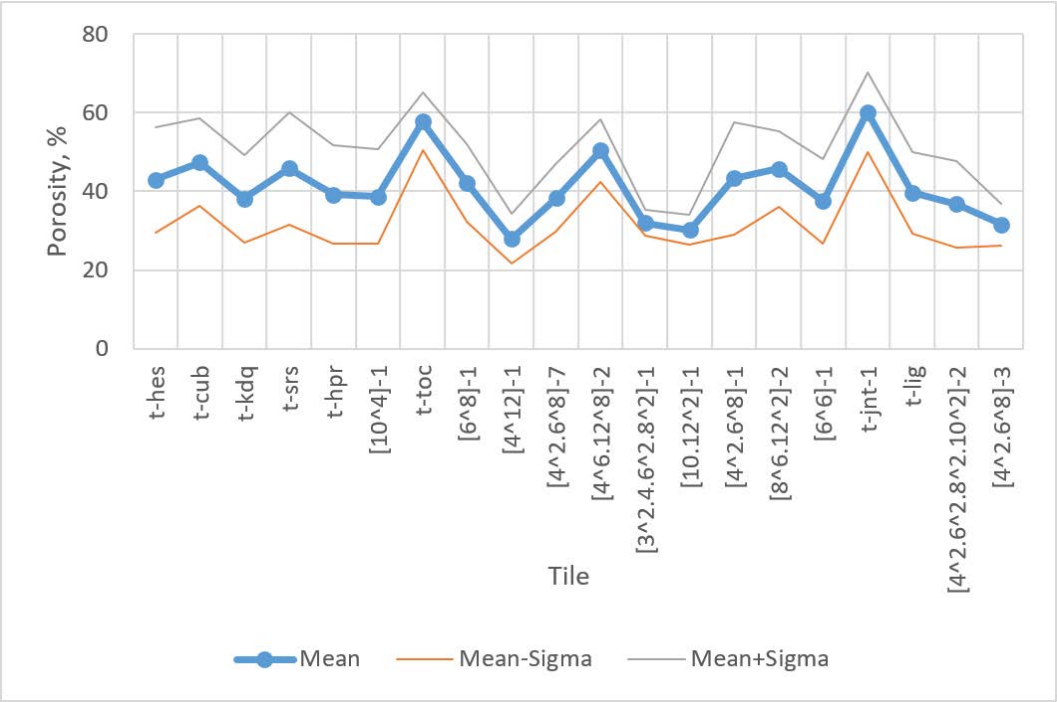

Figure S9. Dependence of the porosity of the coordination polymers on the topological type of the tile of the underlying net tiling. The first 20 abundant tiles of the isohedral tilings and considered. The borders of the standard deviation gap are shown by red and grey thin lines.

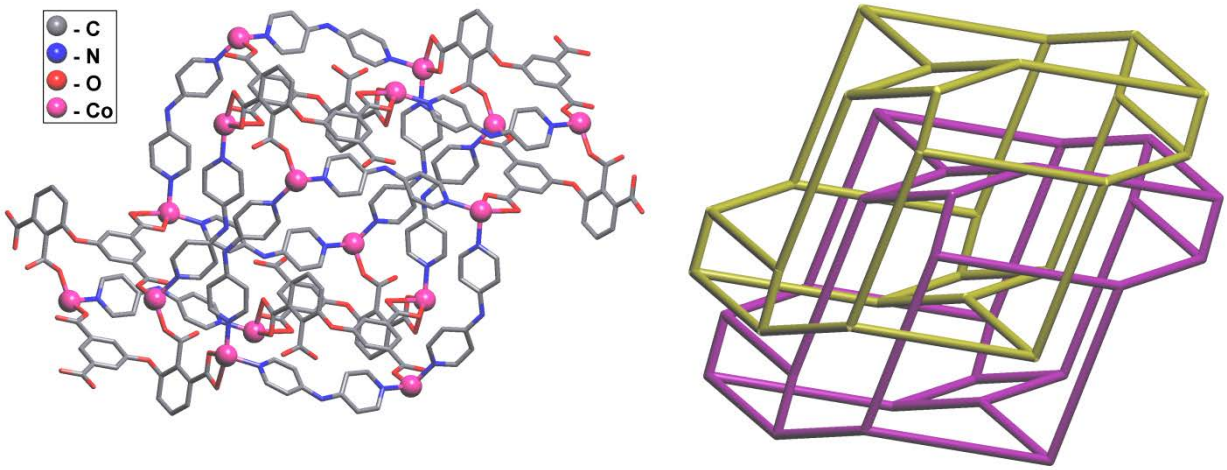

Figure S10. A fragment of the coordination network (left) and the corresponding natural tiles (right) in the interpenetrating array of two **sod** underlying nets in the crystal structure of [((3,5-dicarboxylatophenoxy)phthalato)(N-(pyridin-4-yl)pyridin-4-amine)Co_2] (IJOGAR).[23]

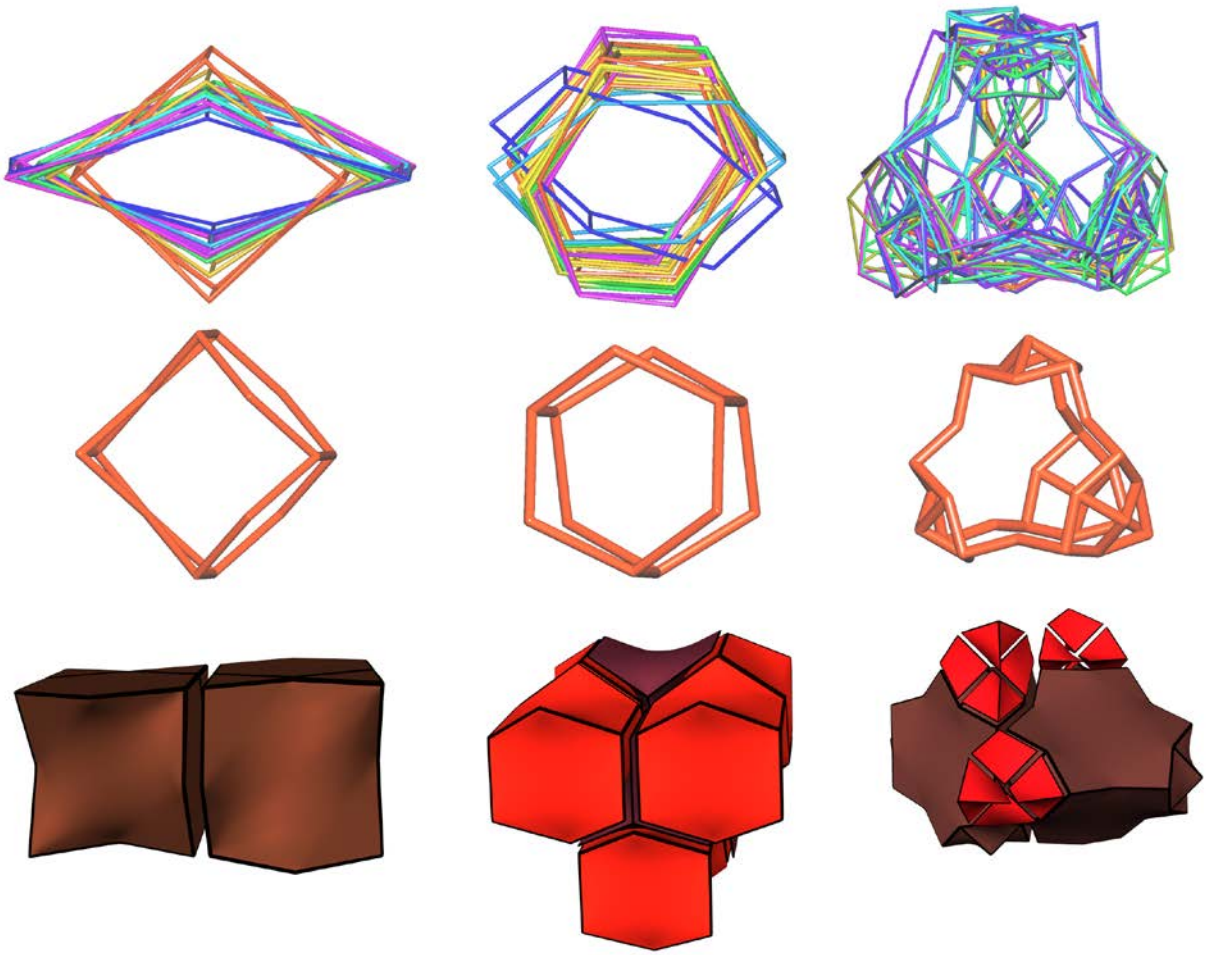

Figure S11. (Top) The overlapping diagrams for geometrically distorted tiles [$3^4.6^2.8^2$] (CI = 10/7), [6^5] (CI = 5/4), and [$4^{12}.12^4$] (CI = 24/17) in 14 **bpq** (left), 11 **lon** (centre) and 25 **met** (right) underlying nets of coordination polymers compared with the most symmetrical shapes of the tiles (middle) and the corresponding natural tilings (bottom). The colouring of the distorted tiles depends on their geometrical distortion (Σ), from red (small distortion) to blue (large distortion).

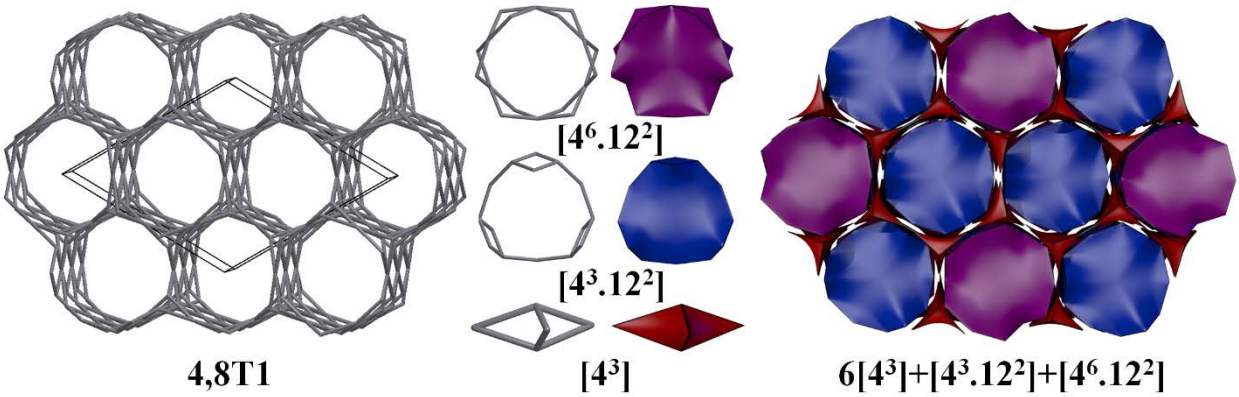

Figure S12. The natural tiles and tiling of the 4,8T1 net.

References

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

2. V. A. Blatov, E. V. Alexandrov and A. P. Shevchenko, Topology: ToposPro, in *Reference Module in Chemistry, Molecular Sciences and Chemical Engineering*, Elsevier, 2019, DOI: 10.1016/b978-0-12-409547-2.14576-7.

3. M.E. Moussa, E.V. Peresypkina, A.V. Virovets, D. Venus, G. Balazs and M. Scheer, *CrystEngComm*, 2018, **20**, 7417.

4. C. Bonneau, M. O'Keeffe, D. M. Proserpio, V. A. Blatov, S. R. Batten, S. A. Bourne, M. Soo Lah, J.-G. Eon, S. T. Hyde, S. B. Wiggin and L. Öhrström, *Cryst. Growth Des*. 2018, **18**, 3411.

5. V. A. Blatov, O. Delgado-Friedrichs, M. O'Keeffe and D. M. Proserpio, *Acta Cryst*. 2007, **A63**, 418.

6. N. A. Anurova, V. A. Blatov, G. D. Ilyushin and D. M. Proserpio, *J. Phys. Chem. C*, 2010, **114**, 10160.

7. O.A. Blatova, A.A. Golov and V. A. Blatov, *Z. Kristallogr*. 2019, **234**, 421.

8. S. A. Alvarez, *Dalton Trans.* 2013, **42**, 8617.

9. A. P. Shevchenko, R. A. Eremin, V. A. Blatov, *CrystEngComm* 2020, DOI: 10.1039/D0CE00265H.

10. P. Arun, C. Santanu, J.C. Boquera, F. Lloret, Jian-Bin Lin, C.P. Shyam and M.C. Das, *Inorg. Chem*. 2019, **58**, 6246.

11. R. M. Hartshorn, E. Hey-Hawkins, R. Kalio and G. J. Leigh, *Pure Appl. Chem*. 2007, **79**, 1779.

12. V. N. Serezhkin, A. V. Vologzhanina, L. B. Serezhkina, E. S. Smirnova, E. V. Grachova, P.V. Ostrovac and M. Yu. Antipin, *Acta Cryst.*, 2009, **B65**, 45.

13. A. P. Shevchenko, I. A. Blatov, E. V. Kitaeva and V. A. Blatov, *Cryst. Growth Des*. 2017, **17**, 774.

14. I. A. Blatov, E. V. Kitaeva, A. P. Shevchenko and V. A. Blatov, *Acta Cryst*. 2019, **A75**, 827.

15. M. O'Keeffe, M. A. Peskov, S. J. Ramsden and O. M. Yaghi, *Acc. Chem. Res*. 2008, **41**, 1782.

16. S. J. Ramsden, V. Robins and S. T. Hyde, *Acta Cryst.* 2009, **A65**, 81.

17. E. Koch, W. Fischer and H. Sowa, *Acta Cryst*. 2006, **A62**, 152.

18. http://www.iza-structure.org/databases/.

19. V. A. Blatov, A. P. Shevchenko and D. M. Proserpio, *Cryst. Growth Des*. 2014, **14**, 3576.

20. Xiaoliang Zhao, Haiyan He, Tuoping Hu, Fengna Dai and Daofeng Sun, *Inorg. Chem*. 2009, **48**, 8057.

21. Jihyun An, O. K. Farha, J. T. Hupp, E. Pohl, J. I. Yeh and N. L. Rosi, *Nature Commun*. 2012, **3**, 604.

22. Yun-Qi Tian, Shi-Yan Yao, Dong Gu, Ke-Hui Cui, Dong-Wei Guo, Gao Zheng, Zhen-Xia Chen and Dong-Yuan Zhao, *Chem Eur. J*. 2010, **16**, 1137.

23. Xiu-Yan Dong, Chang-Dai Si, Yan Fan, Dong-Cheng Hu, Xiao-Qiang Yao, Yun-Xia Yang and Jia-Cheng Liu, *Cryst. Growth Des*. 2016, **16**, 2062.