Electronic Supplementary Information (ESI) for
Understanding the geometric diversity of inorganic and hybrid frameworks through structural coarse-graining

Thomas C. Nicholas,² Andrew L. Goodwin,² and Volker L. Deringer*

²Department of Chemistry, Inorganic Chemistry Laboratory, University of Oxford, Oxford OX1 3QR, UK

*volker.deringer@chem.ox.ac.uk

This ESI document contains additional information about the procedures used for:

- Selecting structural data (p. S2)
- Processing and “cleaning up” structural data (p. S3)
- Coarse-graining structural data for different material systems (p. S4)
- Re-scaling structural data as input for cg-SOAP, and why this is needed (p. S7).

In addition, this document contains additional information and references for all entries of the database; structures which are included (“y”) and structures not included for specified reasons (“n”) are listed in tables as follows:

- Table 1 (p. S8): ZIFs and related hybrid framework materials
- Table 2 (p. S11): Silica (SiO₂) polymorphs
- Table 3 (p. S12): Aluminophosphates (AlPOs)
- Table 4 (p. S18): Ice structures
- Table 5 (p. S19): Zn(CN)₂ polymorphs
- Table 6 (p. S20): Clathrate structures
- Table 7 (p. S21): Other inorganic structures
- References for all these are given at the end of the document, starting on p. S22.
Selection of structural data

Creating the database for this work consisted of two stages: compiling comprehensive lists of experimental structures representing chemical families typified by 4-fold connected networks, and curating the resulting database according to strict criteria. These stages are described in the following subsections. References for all original data are given starting on p. S8.

Data sources. A 2009 review of all synthesised ZIFs provided a list of hybrid inorganic-organic materials for inclusion (DOI: 10.1021/ar900116g). In the same year, a family of cadmium imidazolate frameworks (“CdIFs”) were also reported by Tian et al. (DOI: 10.1002/chem.200902729) and these structures are also included. All zeolites, AlPOs, and related structures in the Database of Zeolite Structures are included (http://www.iza-structure.org/databases/). The 9 distinct, naturally occurring crystalline silica polymorphs and the 17 experimentally realised ice polymorphs are also included. The cyanides were taken from a comprehensive report of various porous polymorphs of Zn(CN)₂ (DOI: 10.1021/ja4012707). The remaining inorganic and clathrate structures were chosen manually and taken from either the Cambridge Crystallographic Data Centre (CCDC) or the Inorganic Crystal Structure Database (ICSD).

Selection criteria. Each entry was manually checked, and only accepted into the curated database if:

(i) the framework has AB₂ stoichiometry;
(ii) each A site environment is tetrahedral;
(iii) the network is fully connected; and
(iv) the A–B–A bonding is well represented when the B site species is modelled as a single point.

It should be noted that this final point (iv) is a limitation of the present method. For example, the highly directional bonding associated with ligands in M[Au(CN)₂]₂ (M = Zn, Co) or large MOF ligands, in which the coordinating ligands atoms are greatly separated in space, are poorly represented by a single point. As such, these structure classes were not considered for the present work.
Cleaning up structures

Cleaning up the structures involves the removal of guest species (solvent molecules, (organic) SDAs, and counter-ions) and averaging partial-occupation and thermally disordered sites.

**Guest atoms.** Following the identification of the A and B site atoms of structures, all other atomic species are removed. This approach works provided that, for materials in which O is a constituent member of the framework, the pores do not contain O atoms.

Two scenarios are commonly found in which this is not the case. Firstly, if the guest species contains any O atoms (for example, if a solvent such as methanol resides in the pores). Secondly, in the crystallography of porous frameworks the electron density associated with guest species within the pores is often delocalised, and thus not resolved. This is often modelled as clusters of (artificial) O atoms.

Zeolite gismondine demonstrates both of these situations, and the guest removal procedure is illustrated in Figure 2.2. Guest O species are removed by two sequential “tidying” algorithms. The first removes clusters of O atoms by iterating over all O atom indices, and if a second O atom is found within a distance of 1.8 Å, both are removed (**Fig. S1**; step 1). Any remaining guest O atoms are then removed by iterating over all remaining O atom indices, and if no A site nodes lie within a sphere of radius 1.8 Å, about the O atom, it is removed (**Fig. S1**; step 2). The distance, 1.8 Å, was chosen based on typical Si–O and Al–O bond lengths (between 1.58 and 1.78 Å), and by manual inspection of the structures concerned. (The specific cut offs used may have been changed manually from these values for a given structure.)

**Fig. S1** Removal of interstitial oxygen atoms, exemplified for the zeolite gismondine.
**Disordered sites.** In many crystal structures atoms have partial occupancy, either as a result of thermal motion or occupational disorder; in contrast, the approach presented here requires well-defined A and B sites with unity occupation. To deal with disordered sites, we iterate over all indices of the framework atom species concerned and, if a second atom of the same species is found within a distance $r$ (where $r = 0.9$ Å for N atoms, else $r = 0.6$ Å), a new atom of the same atomic species is placed at the average position. The original two atoms are then removed.

These two cut-off distances were chosen to strike a balance between not missing any disordered sites while also avoiding mistakenly averaging closely situated atoms. A separate distance was used when considering N disorder because it was usually a result of rotational disorder in the imidazolate ligands, and therefore the separation of the two sites was larger, relative to thermal vibrational disorder.

**Coarse-graining**

The coarse-graining procedures identify and equalise all A sites to the same atomic species, and identify all B sites and place a dummy atom at the centre of the bonding interactions. All other atoms are then removed. Specific coarse-graining routines have been developed that identify and reduce the different types of frameworks to the fundamental AB$_2$ units.

**Inorganic networks.** The cyanides (as well as LiCo(CO)$_4$) are coarse-grained by identifying the midpoint of each CN (or CO) ligand, and placing the dummy B site atom there. Then all A site atoms (Li, Co, and Zn) are equalised to the dummy A atom. The coarse-graining procedure for the remaining inorganic networks is illustrated for gismondine in Fig. S2. First, the framework atoms are identified. We assumed that any Si, Al, P, and O atoms are a part of the underlying framework, and included an option to enter the atomic symbol manually for any node not included in this list. Following the identification of the underlying inorganic 4-c network atoms, all other atoms are removed. Any O atoms found within the pores are then removed by the tidying algorithms. Finally, all O atoms (B sites) are labelled with the dummy atom.
Fig. S2  Clean-up and coarse-graining of inorganic structures, exemplified for the zeolite gismondine: details of the clean-up step have been illustrated in Fig. S1, but we here show additionally the reduction to the fundamental \( \text{AB}_2 \) network by equalising the \{Al, Si\} A-sites.

Ices and clathrate hydrates. The ices incorporated in the curated database are classified as either H-ordered or H-disordered structures. The H-ordered ices have unity occupation for all protons and therefore have the correct \( \text{AB}_2 \) stoichiometry. All protons are replaced by the B-site dummy atom. The H-disordered ice and clathrate hydrate structures all have half-occupancy proton positions which are dealt with separately.

As illustrated by ice IV in Fig. S3, for a generic algorithm to be implemented it cannot be assumed that each pair of closest protons is necessarily from different O nodes. To overcome this, the constituent atoms of each tetrahedron are identified by taking the nearest four H neighbours to each O node. Then, iterating over all pairs of tetrahedra, if the distance between any two protons from different tetrahedra is less than 1.6 Å, a B-site dummy atom is placed at the mid-point and the protons are removed.

Fig. S3  The coarse-graining approach for the H-disordered ices and the clathrate hydrates, in which four H atoms surround a central O, each with an occupation of \( \frac{1}{2} \). First, these tetrahedra are identified (red). The “dummy” B atom (yellow) is then placed at the mid-point between nearest-neighbour protons from different tetrahedra. Finally, all protons are removed.
**Hybrid networks.** The hybrid networks in the database all have imidazolate-derivative organic linkers. The main framework atoms are defined as the metal nodes (A sites) and the N–C–N motifs in each imidazole linker (B sites).

To identify the B sites automatically, our algorithm first identifies the N–C–N motif (Fig. S4a). This is achieved by iterating through the C atoms and identifying if there are N atoms within a sphere of radius, 1.5 Å, centred on the C atom. If three N atoms are found it is assumed the imidazolate-derivative has a NO$_2$ side-chain. The nitrate N atom is then distinguished from the bonding N atoms by searching for the two O atoms anticipated within a sphere of radius, 1.4 Å. Once two bonding N atoms have been identified, the B site dummy atom is placed at the midpoint of the N···N contact. The structure is then coarse-grained by removing all but the A and B site dummy atoms, as illustrated for ZIF-74 in Fig. S4b–c.

There are two reasons for taking the midpoint of the N···N contact as the B site position instead of the centre of mass of the ligand. Firstly, large side-chains on the imidazolate ligand would skew the centre of mass away from the intuitively assumed midpoint of the link. (This effect would be exacerbated by thermal disorder in the ligand because the present approach does not assign fractional weights). This would result in a distortion to the bond angles and lengths, giving an unfaithful representation of local structure. Secondly, it ensures a uniform placement of the B site dummy atom in chemically modified series with different side-chains.

![Fig. S4](image)

**Fig. S4** The coarse-graining approach for hybrid structures. Panel (a) demonstrates a general scheme for identifying the (nitro-)Im ligands, as discussed above. Panels (b) and (c) illustrate the procedure for Zn(mbIm)(nlm) (ZIF-74) in which two different linkers are coarse-grained.
Rescaling

Finally, all structures are scaled down to a uniform minimum bond length ($r_0$). This scaling is of key importance, because it enables the comparison of structures with inherently different chemistries and bond lengths. The scaling is achieved by calculating the distance between all atoms, and using ASE’s `set_cell` function – which allows one to scale the dimensions of the unit cell and the atomic positions proportionately – to scale the structures by dividing by the minimum bond distance found (such that in all re-scaled structures, the new minimum distance is $r_0 = 1 \, \text{Å}$). **Fig. S5** illustrates the need for doing this by comparing (cg-) SOAP distances between pairs of relevant structures: geometric relationships are only properly recovered if the coarse-graining and re-scaling steps are applied.

**Fig. S5** Examples illustrating the requirement for both the coarse-graining and the re-scaling of structures in our approach. (a) Ice-VII and the Zn(CN)$_2$-I polymorph both crystallise in an anticuprite-like structure. The initial structures are fully dissimilar when using a chemically selective SOAP kernel (i.e., one that distinguishes between, say, O and Zn atoms on the A-site). The coarse-grained (AB$_2$) representations are still highly dissimilar ($d = 1.38$), because the A–B distances are very different (note that for the analysis of structures before re-scaling, the kernel cut-off and smoothness have been adjusted accordingly to $r_{\text{cut}} = 6.0 \, \text{Å and } \sigma = 0.4 \, \text{Å}$, respectively). Subsequent re-scaling to a uniform A–B distance, however, shows that the resultant structures are identical (with a cg-SOAP distance of zero). (b) Same for the isoreticular structures of ZIF-74 and gismondine, which have served as examples in the figures above.
The Zeolitic Imidazolate Frameworks review paper by Phan et al.\cite{1} provided a foundation dataset for the ZIF structures to be included. Twelve Cd analogues (CdIFs) have been reported by Tian et al.\cite{2} and have been included here.

### Table 1: ZIF structures in main, coarse-grained database

| Composition     | DepNum/CCDC Code     | Name   | Topology | Included? | Details                                      | Reference |
|-----------------|----------------------|--------|----------|-----------|----------------------------------------------|-----------|
| Zn(Im)₂         | 287181/MECWIB        | ZIF-14 | ANA      | y         | -                                            | [3]       |
| Cd(Im)₂         | 743555/GUPCOK        | CdIF-6 | ANA      | y         | -                                            | [2]       |
| Cd(nPrIm)₂      | 743560/GUPDIF        | CdIF-10| ANA      | y         | -                                            | [2]       |
| Cd(nBuIm)₂      | 743561/GUPDOL        | CdIF-11| ANA      | y         | -                                            | [2]       |
| Co(Im)₂         | 228486/EQOCES01      | -      | cag      | y         | -                                            | [4]       |
| Zn(nIm)₂        | 670170/GIZJOP        | ZIF-62 | cag      | y         | -                                            | [5]       |
| Co(Im)₂         | 223195/NAFGOR        | -      | cag      | y         | -                                            | [4]       |
| Zn(Im)₂         | 602536/VEJYUF        | ZIF-4  | cag      | y         | -                                            | [6]       |
| Zn(Im)₂         | 254158/VEJUF01       | -      | cag      | y         | -                                            | [7]       |
| Zn(Im)₂         | 701064/QOSYAZ        | TIF-4  | cag      | y         | -                                            | [8]       |
| Zn(Im)₂         | 212357/EQOCOC        | -      | coi      | y         | -                                            | [9]       |
| Co(Im)₂         | 118033/IMZYCO        | -      | coi      | y         | -                                            | [10]      |
| Zn(Im)₂         | 671071/GITTEJ        | ZIF-64 | crb(BCT) | y         | - Disordered ligand - take avg. pos.          | [5]       |
| Fe(mIm)₂        | 1208073/LODCUC       | -      | crb(BCT) | y         | - Disordered ligand - take avg. pos.          | [11]      |
| Co(Im)₂         | 223194/NAFGOR01      | -      | crb(BCT) | y         | -                                            | [4]       |
| Zn(Im)₂         | 602535/VEJYEP        | ZIF-1  | crb(BCT) | y         | -                                            | [6]       |
| Zn(Im)₂         | 254157/VEJYP01       | -      | crb(BCT) | y         | -                                            | [7]       |
| Zn₂(Im)₄        | 602536/VEJYIT        | ZIF-2  | crb(BCT) | y         | -                                            | [6]       |
| Zn(Im)₂         | 254159/VEJYIT01      | -      | crb(BCT) | y         | -                                            | [7]       |
| Pr(Im)₅         | 609489/LEMVOP        | -      | crs      | n         | Octahedral Pr environment                     | [12]      |
| Zn(Im)₂         | 254160/HIFVOI        | -      | dft      | y         | -                                            | [7]       |
| Zn₂(Im)₄        | 602537/VEJYPOZ       | ZIF-3  | dft      | y         | -                                            | [6]       |
| Zn₄(abIm)₂      | 647002/MIHOB         | ZIF-23 | dia      | y         | -                                            | [13]      |
| Cd₂(HIm)₃(Im)   | 655703/VIGHID        | -      | dia      | n         | Octahedral Cd environment                     | [14]      |
| Fe₄(abIm)₂      | 153520/XASGON        | -      | dia      | y         | -                                            | [15]      |
| Fe₂(bIm)₂       | 1312697/ZIMMMEN      | -      | dia      | n         | Octahedral Fe environment                     | [16]      |
| Cd(nim)₂        | 743556/GUPPCUQ       | CdIF-7α| dia      | y         | -                                            | [2]       |
| Cd(nim)₂        | 730757/GUPBOJ        | CdIF-7β| dia      | y         | -                                            | [2]       |
| Cd(PhIm)₂       | 743562/GUPDUR        | CdIF-12| dia      | y         | -                                            | [2]       |
| Hg(Im)₂         | 205573/BAYPUN        | -      | dia-c    | y         | -                                            | [17]      |
| Cd(Im)₂         | 205574/BAYQAU        | -      | dia-c    | y         | -                                            | [17]      |
| Cd(Im)₂         | 213239/BAYQAU01      | -      | dia-c    | y         | -                                            | [9]       |
| LiB(mIm)₄       | 699084/MOXXUG        | BIF-2Li| dia-c-b  | y         | -                                            | [18]      |
| CuB(mIm)₄       | 703703/MUCCLG        | BIF-2Cu| dia-c-b  | y         | -                                            | [18]      |
| CuBH(Im)₃       | 697962/MOXXEG        | BIF-6  | fes      | n         | Pyramidal Cu                                  | [18]      |
| Zn(nIm)₁.₇₄(mbIm)₀.₂₆ | 671083/GITEVOV      | ZIF-73 | fsl      | y         | Disordered ligand                            | [5]       |
| Zn(nIm)₂        | 671089/GITWIQ        | ZIF-77 | fsl      | y         | -                                            | [5]       |
| Zn₄₂(Im)₅       | 602539/VEJZAM        | ZIF-5  | gar      | n         | Octahedral In environment                     | [6]       |
| Zn(Im)₂         | 602540/EQOCOC01      | ZIF-6  | GIS      | y         | -                                            | [6]       |
| Zn₂(abIm)(nIm)  | 671085/GITVUB        | ZIF-74 | GIS      | y         | -                                            | [5]       |
| Zn(Im)₂         | 671086/GITWAI        | ZIF-75 | GIS      | y         | -                                            | [6]       |
| Zn(Im)₂         | 254162/HIPVUO        | -      | GIS      | y         | -                                            | [7]       |
Table 1: ZIF structures in main, coarse-grained database

| Composition          | DepNum/CCDC Code | Name     | Topology | Included? | Details                                      | Reference |
|----------------------|------------------|----------|----------|-----------|----------------------------------------------|-----------|
| Zn(Im)(dmbIm)        | 701066/QOSYIH    | TIF-5Zn  | GIS      | y         |                                              | [8]       |
| Co(Im)(dmbIm)        | 701065/QOSYED    | TIF-5Co  | GIS      | y         |                                              | [8]       |
| Zn(bIm)(nIm)         | 671075/GITTUZ    | ZIF-68   | GME      | y         |                                              | [5]       |
| Zn(cblm)(nIm)        | 671076/GITVAH    | ZIF-69   | GME      | y         | Thermal disorder in ligand                   | [5]       |
| Zn(Im)_{1.13}(nIm)_{0.87} | 671078/GITVEL   | ZIF-70   | GME      | y         |                                              | [5]       |
| Zn(nblm)(nIm)        | 704995/YOZBIZ    | ZIF-78   | GME      | y         | Thermal disorder in ligand                   | [19]      |
| Zn(mblm)(nIm)        | 704997/YOZBOF    | ZIF-79   | GME      | y         |                                              | [19]      |
| Zn(dclm)(nIm)        | 704999/YOZBUL    | ZIF-80   | GME      | y         |                                              | [19]      |
| Zn(brbIm)(nIm)       | 705001/YOZCAS    | ZIF-81   | GME      | y         |                                              | [19]      |
| Zn(nIm)(nIm)         | 705002/YOZCEW    | ZIF-82   | GME      | y         |                                              | [19]      |
| Cd(mIm)_{2}          | 743554/GUPBOJ01  | CdIF-5   | lcs      | y         |                                              | [2]       |
| Zn(dcIm)_{2}         | 671082/GIZJUV    | ZIF-72   | lcs      | y         |                                              | [5]       |
| Zn(Im)(cbIm)         | 671087/GITWEM    | ZIF-76   | LTA      | y         |                                              | [5]       |
| Zn(pur)_{2}          | 646999/MIIHAN    | ZIF-20   | LTA      | y         |                                              | [13]      |
| Co(pur)_{2}          | 647000/MIHHER    | ZIF-21   | LTA      | y         |                                              | [13]      |
| Zn(5abIm)_{2}        | 647001/MIHHSV    | ZIF-22   | LTA      | y         |                                              | [13]      |
| Cd(Im)_{2}(bipy)     | 294000/DAYVIJ    |          |          |           | Octahedral Cd environment                    | [20]      |
| In_{5}(Imdc)_{10}    | 690432/EGEHOO    | usf-ZMOF | med      | y         |                                              | [21]      |
| Zn_{2}(Im)_{3}(Im)   | 671067/GITSUY    | ZIF-60   | MER      | y         |                                              | [5]       |
| Zn(Im)_{2}           | 602544/VEJZIU    | ZIF-10   | MER      | y         |                                              | [6]       |
| Cd(mIm)_{2}          | 743552/GUPCAW    | CdIF-2   | MER      | y         |                                              | [2]       |
| Cu(Im)_{2}           | 159163/CUIHDI03  |          |          |           |                                              | [22]      |
| Fe_{3}(Im)_{6}       | 1180191/MIDIME   | -        | mog      | n         | Octahedral Fe environment                    | [23]      |
| Fe_{3}(Im)_{6}       | 1180192/MIDIFE01 | -        | mog      | n         | Octahedral Fe environment                    | [24]      |
| Mn_{3}(Im)_{6}       | 1180193/MIDZIA   | -        | mog      | n         | Octahedral Mn environment                    | [25]      |
| Zn_{20}(cblm)_{35}(OH) | 668215/NOFQEF   | ZIF-100  | moz      | n         | Network modifier present                    | [26]      |
| Co(Im)_{2}           | 149555/EQOBUH    | -        | neb      | y         |                                              | [9]       |
| Co(Im)_{2}           | 212355/EQOCES    | -        | neb      | y         |                                              | [9]       |
| Co(Im)_{2}           | 212356/EQOCIW    | -        | neb      | y         |                                              | [9]       |
| Co_{5}(Im)_{10}      | 168798/AFIXAO    | -        | nog      | y         |                                              | [9]       |
| Co_{5}(Im)_{10}      | 168799/AFIXES    | -        | nog      | y         |                                              | [27]      |
| Zn(Im)_{2}           | 646539/HIFWAV    | -        | nog      | y         |                                              | [7]       |
| Zn(Im)(mbIm)         | 701063/QOSXUS    | TIF-3    | pcb(ACO) | y         | Different ligand - adapted cg procedure     | [8]       |
| Zn(cblm)_{2}         | 668214/NOFQAB    | ZIF-95   | poz      | y         |                                              | [26]      |
| Fe(mIm)_{2}          | 1119170/CAGLIEF  | -        | qtz      | y         |                                              | [28]      |
| Zn(dclm)_{2}         | 671080/GITVIP    | ZIF-71   | RHO      | y         |                                              | [5]       |
| Zn_{2}(elm)_{4}      | 287182/MECWOH    | -        | RHO      | y         |                                              | [3]       |
| In(Imdc)_{2}         | 294663/TEFWIL    | rho-ZMOF | RHO      | n         | In not tetrahedral                           | [29]      |
| Zn(bIm)_{2}          | 602545/VEJZOA    | ZIF-11   | RHO      | y         |                                              | [6]       |
### Table 1: ZIF structures in main, coarse-grained database

| Composition | DepNum/CCDC Code | Name | Topology | Included? | Details | Reference |
|-------------|------------------|------|----------|-----------|---------|-----------|
| Co(bIm)$_2$ | 602546/VEJZUG    | ZIF-12 | RHO      | y         | -       | [6]       |
| Cd(eIm)$_2$ | 743553/GUPCEA    | CdIF-4 | RHO      | y         | -       | [2]       |
| Cd(nIm)$_2$ | 743559/GUPDEB    | CdIF-9 | RHO      | y         | -       | [2]       |
| Zn(ica)$_2$ | 693596/WOJGEI    | ZIF-90 | SOD      | y         | -       | [30]      |
| Zn(bIm)$_2$ | 211533/AKUGES    | -     | SOD      | y         | -       | [31]      |
| Cu(Im)$_2$  | 159161/CUIMDZ01  | -     | SOD      | y         | -       | [22]      |
| Co(nIm)$_2$ | 671072/GITTIN    | ZIF-65 | SOD      | y         | -       | [5]       |
| Co(nIm)$_2$ | 671073/GITTOT    | ZIF-67 | SOD      | y         | -       | [5]       |
| Zn(mIm)$_2$ | 287180/MECWEX    | -     | SOD      | y         | -       | [3]       |
| Zn(Im-d5)$_2$ | 652032/OFERUN | ZIF-8(D) | SOD | n | Same as ZIF-8(H) | [32] |
| In(Imdc)$_2$ | 294664/TEFWOR   | sod-ZMOF | SOD | n | In not tetrahedral | [29] |
| Co(bIm)$_2$ | 602543/VEJZEQ    | ZIF-9  | SOD      | y         | -       | [6]       |
| Zn(bIm)$_2$ | 602541/VELVIS    | ZIF-7  | SOD      | y         | -       | [6]       |
| Zn(mIm)$_2$ | 602542/VELVOY    | ZIF-8  | SOD      | y         | -       | [6]       |
| Cd(mIm)$_2$ | 743551/GUPBUP    | CdIF-1 | SOD      | y         | -       | [2]       |
| Cd(mIm)$_2$ | 743558/GUPCUQ01  | CdIF-8 | SOD      | y         | -       | [2]       |
| LiB(mIm)$_2$ | 703704/MUCLOM    | BIF-3Li | SOD-b    | y         | -       | [18]      |
| CuB(mIm)$_2$ | 697959/MOXJOZ   | BIF-3Cu | SOD-b    | y         | -       | [18]      |
| CuBH(eIm)$_3$ | 697964/MOKKOA   | BIF-8  | srs-c-b   | n | Pyramidal Cu | [18] |
| CuBH(mIm)$_3$ | 697963/MOKKIU   | BIF-7  | ths-c-b   | n | Not tetrahedral | [18] |
| Cd(mim)$_2$ | 736757/GUPBOJ    | CdIF-3 | yqt1      | n | Disordered ligand | [2] |
| Zn(dmbIm)$_2$ | 682400/QOGPIM   | TIF-1Zn | zea       | y         | -       | [33]      |
| Zn(Im)$_{1,16}$(mbIm)$_{0,9}$ | 701062/QOSXOM | TIF-2  | zeb       | y         | -       | [8]       |
| Zn(Im)$_2$  | 254161/HICGEG    | -     | zec       | y         | -       | [7]       |
| Zn(Im)(mIm) | 671069/GITTAF    | ZIF-61 | zni       | y         | -       | [5]       |
| Zn(Im)$_2$  | 1180194/IMIDZB   | -     | zni       | y         | -       | [25]      |
| Co(Im)$_2$  | 212358/IMZYCO01  | -     | zni       | y         | -       | [9]       |
| LiB(Im)$_2$ | 693499/MOXJEP    | BIF-1Li | zni-b    | y         | -       | [18]      |
| CuB(Im)$_2$ | 693500/MOXJIT    | BIF-1Cu | zni-b    | y         | -       | [18]      |
| CuCu(Im)$_3$ | 227576/BETHUE    | -     | -         | n | Linear Cu | [34] |
| CuCu[B(bIm)$_4$]$_2$ | 697960/MOXJUF | BIF-4  | -         | n | Ligand disorder modifier | [18] |
| Cu3I[B(bIm)$_4$]$_2$ | 697961/MOXKAM   | BIF-5  | -         | n | Network modifier | [18] |
All of the compact polymorphic forms of silica have been included in the dataset. Structures have been downloaded from the ICSD and the American Mineralogist Crystal Structure Database (AMCSD).[35]

| Mineral       | ICSD (AMCSD) Identifier | Reference |
|---------------|--------------------------|-----------|
| Coesite       | 156195                   | [36]      |
| Cristobalite-high | (0017665)             | [37]      |
| Cristobalite-low    | (0017659)             | [38]      |
| Tridymite-high    | (0020733)             | [39]      |
| Tridymite-low     | (0009625)             | [40]      |
| Keatite         | 34889                   | [41]      |
| Moganite        | 67669                   | [42]      |
| Quartz-high*    | -                       | [43]      |
| Quartz-low      | 173227                  | [44]      |
| Stishovite      | not included            | Octahedral Si environment |

*Crystallographic information used corresponds to the synthetic crystal (590°C), O general dataset, taken from Table 1 of this reference.
All structures for the experimentally realised Zeolites and related microporous materials have been taken from the IZA-database. For each framework type, the cif files provided have been included in the database.

Table 3: Aluminophosphate structures included in the database

| Mineral       | Topology | DepNum         | Description                                      | Reference |
|---------------|----------|----------------|--------------------------------------------------|-----------|
| Li-A(BW)      | ABW      | IZA-SC         | n - H$_2$O network-modifier                      | [47]      |
| ACP-1         | ACO      | 661416/BEFNIL   | y                                                 | [48]      |
| AIPO-18(calcined) | AEI        | IZA-SC         | y                                                 | [49]      |
| AIPO-11(calcined) | AEL        | IZA-SC         | y                                                 | [50]      |
| MAPO-11       | AEL      | IZA-SC         | y                                                 | [51]      |
| AIPO-53(A)    | AEN      | IZA-SC         | n - penta-coordinate Al                           | [52]      |
| AIPO-8        | AET      | 1635179        | y                                                 | [53]      |
| Afghanite     | AFG      | IZA-SC         | y                                                 | [54]      |
| AIPO-5        | AFI      | 1651199/(91673)| y                                                 | [55]      |
| AIPO-14(calcined) | AFN        | IZA-SC         | n - not fully connected                          | [56]      |
| AIPO-41(calcined) | AFO        | IZA-SC         | y                                                 | [57]      |
| SAPO-40       | AFR      | IZA-SC         | y                                                 | [58]      |
| MAPSO-46      | AFS      | IZA-SC         | y                                                 | [59]      |
| AIPO-52(calcined) | AFT        | IZA-SC         | y                                                 | [60]      |
| AIPO-57       | AVF      | IZA-SC         | y                                                 | [61]      |
| SAPO-56       |AFX       | IZA-SC         | y                                                 | [62]      |
| CoAPO-50      | AFY      | IZA-SC         | y                                                 | [59]      |
| AIPO-H2       | AHT      | IZA-SC         | n - Octahedral Al environment                     | [63]      |
| Analcime      | ANA      | IZA-SC         | y                                                 | [64]      |
| AIPO-C(hydrated) | APC        | IZA-SC         | n - Octahedral Al environment with H$_2$O network modifiers | [65]|
| AIPO-C(dehydrated) | APC        | IZA-SC         | y                                                 | [66]      |
| AIPO-D        | APD      | IZA-SC         | y                                                 | [66]      |
| Octadecasil   | AST      | IZA-SC         | y                                                 | [67]      |
| AIPO-16       | AST      | IZA-SC         | y                                                 | [68]      |
| ASU-7         | ASV      | IZA-SC         | y                                                 | [69]      |
| MAPO-39       | ATN      | IZA-SC         | n - Disordered                                   | [70]      |
| SAPO-31       | ATO      | IZA-SC         | y                                                 | [71]      |
| MAPO-36(calcined) | ATS        | IZA-SC         | y                                                 | [72]      |
| B-SSZ-55      | ATS      | IZA-SC         | y                                                 | [73]      |
| AIPO-12       | ATT      | IZA-SC         | y                                                 | [74]      |
| AIPO-25       | ATV      | IZA-SC         | y                                                 | [75]      |
| AIPO-78       | AVE      | IZA-SC         | y                                                 | [76]      |
| AIPO-59       | AVL      | IZA-SC         | n - disordered                                   | [61]      |
| AIPO-21       | AWO      | IZA-SC         | n - 5 coordinate Al                              | [75]      |
| AIPO-22       | AWW      | IZA-SC         | n - 5 coordinate Al                              | [77]      |
| Mg-BCTT       | BCT      | IZA-SC         | y                                                 | [78]      |
| FOS-5         | BEC      | IZA-SC         | y                                                 | [79]      |
| ITQ-17        | BEC      | IZA-SC         | y                                                 | [80]      |
| Bikitaite     | BIC      | IZA-SC         | n - H$_2$O network-modifier                      | [81]      |
| UCSB-15GaGe   | BOF      | IZA-SC         | y                                                 | [82]      |
| Boggsite      | BOG      | IZA-SC         | y                                                 | [83]      |
| Be-10         | BOZ      | IZA-SC         | y                                                 | [84]      |
| BePhosphate-H | BPH      | IZA-SC         | y                                                 | [85]      |
| Brewsterite   | BRE      | IZA-SC         | y                                                 | [86]      |
| UCSB-7        | BSV      | IZA-SC         | y                                                 | [82]      |
| Cancrinite    | CAN      | IZA-SC         | y                                                 | [87]      |
| Tiptopite     | CAN      | IZA-SC         | y                                                 | [88]      |
| Mineral                  | Topology | DepNum | Description | Reference |
|-------------------------|----------|--------|-------------|-----------|
| Cs-aluminosilicate      | CAS      | IZA-SC | y           | [89]      |
| UZM-25(calcined)        | CDO      | IZA-SC | y           | [90]      |
| CIT-5                   | CFI      | IZA-SC | y           | [91]      |
| CoGaPO-5                | CGF      | IZA-SC | y           | [92]      |
| CoGaPO-6                | CGS      | IZA-SC | y           | [93]      |
| Chabazite               | CHA      | IZA-SC | y           | [94]      |
| SAPO-47                 | CHA      | IZA-SC | y           | [95]      |
| CIT-1                   | CON      | IZA-SC | n - cif file missing some O positions | [96]      |
| CIT-7                   | CSV      | IZA-SC | y           | [97]      |
| Na-ZnPO                 | CZP      | IZA-SC | y - used P6\(_5\)22 as less disordered | [98]      |
| Dachiardite             | DAC      | IZA-SC | y           | [99]      |
| Deca-dodecasil-3R       | DDR      | IZA-SC | y           | [100]     |
| DAF-1                   | DFO      | IZA-SC | y           | [101]     |
| UiO-20                  | DFT      | IZA-SC | y           | [102]     |
| Dodecasil-1H            | DOH      | IZA-SC | y           | [103]     |
| UTD-1F                  | DON      | IZA-SC | y           | [104]     |
| Bellbergite             | EAB      | IZA-SC | y           | [105]     |
| TMA-E                   | EAB      | IZA-SC | y           | [106]     |
| Epistilbite             | EDI      | IZA-SC | y           | [107]     |
| K-F                     | EDI      | IZA-SC | y           | [108]     |
| SSZ-45                  | EEU      | IZA-SC | y           | [109]     |
| EMC-2(calcined)         | EMT      | IZA-SC | y           | [110]     |
| TNU-7                   | EON      | IZA-SC | y           | [111]     |
| Epistilbite             | EPI      | IZA-SC | y           | [112]     |
| Erionite                | ERI      | IZA-SC | y           | [113]     |
| AIPO-17                 | ERI      | IZA-SC | y           | [114]     |
| ERS-7                   | ESV      | IZA-SC | y           | [115]     |
| EU-12                   | ETL      | IZA-SC | y           | [116]     |
| ECR-34                  | ETR      | IZA-SC | y           | [117]     |
| EU-1                    | EOY      | IZA-SC | y           | [118]     |
| ZSM-50                  | EOY      | IZA-SC | y           | [119]     |
| EMM-26                  | EWS      | IZA-SC | y           | [120]     |
| EMM-3                   | EZT      | IZA-SC | y           | [121]     |
| Farnesite               | FAR      | IZA-SC | y           | [122]     |
| Faujasite               | FAU      | IZA-SC | y           | [123]     |
| Na-X                    | FAU      | IZA-SC | y           | [124]     |
| US-Y                    | FAU      | IZA-SC | y           | [125]     |
| Na-Y(siliceuous)        | FAU      | IZA-SC | y           | [126]     |
| Li-LSX                  | FAU      | IZA-SC | y           | [127]     |
| Ferrierite              | FER      | IZA-SC | y           | [127]     |
| Ferrierite(siliceous)   | FER      | IZA-SC | y           | [128]     |
| Franzinite              | FRA      | IZA-SC | y           | [129]     |
| Gismondine              | GIS      | IZA-SC | y           | [130]     |
| Amicite                 | GIS      | IZA-SC | y           | [131]     |
| Garronite               | GIS      | IZA-SC | y           | [131]     |
| Gobbsinite              | GIS      | IZA-SC | y           | [132]     |
| MAPO-43                 | GIS      | IZA-SC | y           | [133]     |
| Na-P1                   | GIS      | IZA-SC | y           | [134]     |
| Giuseppettite           | GIU      | IZA-SC | y           | [135]     |
| Gmelinite               | GME      | IZA-SC | y           | [136]     |
Table 3: Aluminophosphate structures included in the database

| Mineral              | Topology | DepNum   | Description                                      | Reference |
|----------------------|----------|----------|--------------------------------------------------|-----------|
| GUS-1                | GON      | IZA-SC   | y                                                | [137]     |
| Goosecreekite        | GOO      | IZA-SC   | y                                                | [138]     |
| Heulandite           | HEU      | IZA-SC   | y                                                | [139]     |
| Clinoptilolite       | HEU      | IZA-SC   | y                                                | [140]     |
| ITQ-51               | IFO      | IZA-SC   | y                                                | [141]     |
| ITQ-4(calcined)      | IFR      | IZA-SC   | y                                                | [142]     |
| ITQ-52(calcined)     | IFW      | IZA-SC   | y                                                | [143]     |
| SSZ-87(as-made)      | IFW      | IZA-SC   | n - use (simpler) calcined form                   | [144]     |
| SSZ-87(as-synthesised)| IFW    | IZA-SC   | n - use (simpler) calcined form                   | [144]     |
| SSZ-87(calcined)     | IFW      | IZA-SC   | y                                                | [144]     |
| ITQ-50(calcined)     | IFY      | IZA-SC   | y                                                | [145]     |
| ITQ-32               | IHW      | IZA-SC   | y                                                | [146]     |
| IM-5(calcined)       | IMF      | IZA-SC   | y                                                | [147]     |
| ITQ-49(calcined)     | IRN      | IZA-SC   | y                                                | [148]     |
| ITQ-44               | IRR      | IZA-SC   | y                                                | [149]     |
| ITQ-7(calcined)      | ISV      | IZA-SC   | y                                                | [150]     |
| ITQ-3(calcined)      | ITE      | IZA-SC   | y                                                | [151]     |
| ITQ-38(calcined)     | ITG      | IZA-SC   | y                                                | [152]     |
| ITQ-13               | ITH      | IZA-SC   | y                                                | [153]     |
| ITQ-34               | ITR      | IZA-SC   | y                                                | [154]     |
| ITQ-33               | IIT      | IZA-SC   | n - disordered                                   | [155]     |
| ITQ-12(calcined)     | ITW      | IZA-SC   | y                                                | [156]     |
| ITQ-24               | IWR      | IZA-SC   | y                                                | [157]     |
| ITQ-26               | IWS      | IZA-SC   | y                                                | [158]     |
| ITQ-27               | IWW      | IZA-SC   | y                                                | [159]     |
| ITQ-22               | IWW      | IZA-SC   | y                                                | [160]     |
| Nepheline Hydrate    | JBW      | IZA-SC   | n - not fully connected                          | [161]     |
| JU92-300             | JNT      | IZA-SC   | y                                                | [162]     |
| LSJ-10               | JOZ      | IZA-SC   | y                                                | [163]     |
| CoAPO-CJ40           | JRY      | IZA-SC   | n - disordered                                   | [164]     |
| CoAPO-CJ69           | JSN      | IZA-SC   | y                                                | [165]     |
| JU-64                | JSR      | IZA-SC   | y                                                | [166]     |
| GaGeO-CJ63           | JST      | IZA-SC   | y                                                | [167]     |
| CoAPO-CJ62           | JSW      | IZA-SC   | y                                                | [168]     |
| (Cs,K)-ZK-5          | KFI      | IZA-SC   | y                                                | [169]     |
| Laumontite           | LAU      | IZA-SC   | y                                                | [170]     |
| Leonhardite          | LAU      | IZA-SC   | y                                                | [171]     |
| CoGaPO               | LAU      | IZA-SC   | y                                                | No reference given |

Levyne               | LEV      | IZA-SC   | y                                                | [172]     |
Nu-3                 | LEV      | IZA-SC   | y                                                | [173]     |
Liottite             | LIO      | IZA-SC   | y                                                | [174]     |
Losod                | LOS      | IZA-SC   | y                                                | [175]     |
Lovdarite            | LOV      | IZA-SC   | y                                                | [176]     |
Linde-Type-A(hydrated)| LTA    | IZA-SC   | n - (used dehydrated form)                       | [177]     |
Linde-Type-A(dehydrated)| LTA    | IZA-SC   | y                                                | [178]     |
LZ-135               | LTF      | IZA-SC   | y                                                | [179]     |
Linde-Type-J         | LTJ      | IZA-SC   | y                                                | [180]     |
Table 3: Aluminophosphate structures included in the database

| Mineral          | Topology | DepNum | Description                  | Reference |
|------------------|----------|--------|------------------------------|-----------|
| Perlialite       | LTL      | IZA-SC | y                            | [181]     |
| Linde-Type-L     | LTL      | IZA-SC | y                            | [182]     |
| NaZ-21           | LTN      | IZA-SC | y                            | [183]     |
| Marinellite      | MAR      | IZA-SC | n - not fully connected      | [184]     |
| Mazzite          | MAZ      | IZA-SC | y                            | [185]     |
| ZSM-18           | MEI      | IZA-SC | y                            | [186]     |
| ZSM-11(calcined) | MEL      | IZA-SC | y                            | [187]     |
| ZSM-11(silicuous)| MEL      | IZA-SC | y                            | [188]     |
| Melanophlogite   | MEP      | IZA-SC | y                            | [189]     |
| Merlinoite       | MER      | IZA-SC | y                            | [190]     |
| BaClAPO          | MER      | IZA-SC | y                            | [191]     |
| ZSM-5            | MFI      | IZA-SC | n - use calcined instead     | [192]     |
| ZSM-5(calcined)  | MFI      | IZA-SC | y                            | [193]     |
| ZSM-57           | MFS      | IZA-SC | y                            | [194]     |
| Montesommaite    | MON      | IZA-SC | y                            | [195]     |
| Mordenite        | MOR      | IZA-SC | y                            | [177]     |
| Maricopaite      | MOR      | IZA-SC | n - not fully connected      | [196]     |
| ZSM-10           | MOZ      | IZA-SC | y                            | [197]     |
| ZSM-43           | MRT      | IZA-SC | y                            | [198]     |
| MCM-68           | MSE      | IZA-SC | y                            | [199]     |
| MCM-61           | MSO      | IZA-SC | n - unrecognised space-group | [200]     |
| MCM-35           | MTF      | IZA-SC | y                            | [201]     |
| Dodecasil-3C     | MTN      | IZA-SC | y                            | [202]     |
| ZSM-23           | MTT      | IZA-SC | y                            | [203]     |
| ZSM-12(calcined) | MTW      | IZA-SC | y                            | [204]     |
| MCM-70           | MVY      | IZA-SC | y                            | [205]     |
| ZSM-25           | MWF      | IZA-SC | y                            | [206]     |
| ITQ-1(calcined)  | MWW      | IZA-SC | y                            | [207]     |
| Nabesite         | NAB      | IZA-SC | y                            | [208]     |
| Natrolite        | NAT      | IZA-SC | y                            | [209]     |
| Mesolite         | NAT      | IZA-SC | y                            | [210]     |
| Scolecite        | NAT      | IZA-SC | y                            | [211]     |
| Nu-87            | NES      | IZA-SC | y                            | [212]     |
| Nonasil          | NON      | IZA-SC | y                            | [213]     |
| NPO-1            | NPO      | IZA-SC | y                            | [214]     |
| ONPO-2           | NPT      | IZA-SC | y                            | [215]     |
| Nu-6(2)          | NSI      | IZA-SC | y                            | [216]     |
| OSB-2            | OBW      | IZA-SC | y                            | [217]     |
| Offretite        | OFF      | IZA-SC | y                            | [218]     |
| COK-14           | OKO      | IZA-SC | y                            | [219]     |
| UiO-6(calcined)  | OSI      | IZA-SC | y                            | [220]     |
| OSB-1            | OSO      | IZA-SC | y                            | [217]     |
| UiO-28           | OWE      | IZA-SC | y                            | [221]     |
| Paulingite       | PAU      | IZA-SC | y                            | [222]     |
| IPC-4            | PCR      | IZA-SC | y                            | [223]     |
| Phillipsite      | PHI      | IZA-SC | y                            | [224]     |
| Harmotome        | PHI      | IZA-SC | y                            | [224]     |
| IST-1            | PON      | IZA-SC | y                            | [225]     |
| PST-13           | POR      | IZA-SC | y                            | [226]     |
| PST-14           | POR      | IZA-SC | y                            | [226]     |
| PUK-16           | POS      | IZA-SC | y                            | [227]     |
| Mineral          | Topology | DepNum | Description                        | Reference |
|------------------|----------|--------|------------------------------------|-----------|
| PST-6            | PSI      | IZA-SC | y                                  | [228]     |
| PUK-9            | PUN      | IZA-SC | y                                  | [229]     |
| PST-29           | PWN      | IZA-SC | y                                  | [230]     |
| PST-21           | PWO      | IZA-SC | y                                  | [231]     |
| PST-22           | PWW      | IZA-SC | y                                  | [231]     |
| Pahasapate       | RHO      | IZA-SC | n - not fully connected            | [232]     |
| Rho(hydrated)    | RHO      | IZA-SC | n - use deuterated form             | [233]     |
| Rho(deuterated)  | RHO      | IZA-SC | y                                  | [234]     |
| RUB-41           | RRO      | IZA-SC | y                                  | [235]     |
| RUB-17           | RSN      | IZA-SC | y                                  | [236]     |
| RUB-3            | RTE      | IZA-SC | y                                  | [237]     |
| RUB-13           | RTH      | IZA-SC | y                                  | [238]     |
| RUB-10           | RUT      | IZA-SC | y                                  | [239]     |
| RUB-24           | RWR      | IZA-SC | y                                  | [240]     |
| UCR-20           | RWY      | IZA-SC | y                                  | [241]     |
| STA-15           | SAF      | IZA-SC | y                                  | [242]     |
| STA-1            | SAO      | IZA-SC | y                                  | [243]     |
| STA-6            | SAS      | IZA-SC | y                                  | [244]     |
| STA-2            | SAT      | IZA-SC | y                                  | [245]     |
| STA-7            | SAV      | IZA-SC | y                                  | [246]     |
| UCSB-8Co         | SBE      | IZA-SC | y                                  | [247]     |
| UCSB-9           | SBN      | IZA-SC | y                                  | [248]     |
| UCSB-6GaCo       | SBS      | IZA-SC | y                                  | [247]     |
| UCSB-10GaZn      | SBT      | IZA-SC | y                                  | [247]     |
| SSZ-82           | SEW      | IZA-SC | y                                  | [249]     |
| SSZ-48(calcined) | SFE      | IZA-SC | y                                  | [250]     |
| SSZ-44(calcined) | SFF      | IZA-SC | y                                  | [251]     |
| SSZ-58           | SFG      | IZA-SC | y                                  | [252]     |
| SSZ-53           | SFH      | IZA-SC | y                                  | [253]     |
| SSZ-59           | SFN      | IZA-SC | y                                  | [253]     |
| SSZ-51           | SFO      | IZA-SC | y                                  | [254]     |
| SSZ-56           | SFS      | IZA-SC | y                                  | [255]     |
| SSZ-52           | SFW      | IZA-SC | y                                  | [256]     |
| Sigma-2          | SGT      | IZA-SC | y                                  | [257]     |
| SIZ-7            | SIV      | IZA-SC | y                                  | [258]     |
| Sodalite         | SOD      | IZA-SC | y                                  | [259]     |
| Bicchulite       | SOD      | IZA-SC | y                                  | [260]     |
| Tugtupite        | SOD      | IZA-SC | y                                  | [261]     |
| SU-15            | SOF      | IZA-SC | y                                  | [262]     |
| SCM-14           | SOR      | IZA-SC | y                                  | [263]     |
| SU-16            | SOS      | IZA-SC | y                                  | [264]     |
| SCM-15           | SOV      | IZA-SC | y                                  | [265]     |
| SSZ-65           | SSF      | IZA-SC | y                                  | [266]     |
| SSZ-60           | SSY      | IZA-SC | y                                  | [267]     |
| SSZ-35(calcined) | STF      | IZA-SC | y                                  | [251]     |
| Stilbite         | STI      | IZA-SC | y                                  | [268]     |
| Stellerite       | STI      | IZA-SC | y                                  | [269]     |
| Barrerite        | STI      | IZA-SC | y                                  | [270]     |
| SSZ-23           | STT      | IZA-SC | y                                  | [271]     |
| SU-32            | STW      | IZA-SC | y                                  | [262]     |
| SSZ-77           | SVV      | IZA-SC | y                                  | [272]     |
| STA-20           | SWY      | IZA-SC | y                                  | [273]     |
| Mineral                          | Topology | DepNum | Description | Reference |
|--------------------------------|----------|--------|-------------|-----------|
| SUZ-4                           | SZR      | IZA-SC | y           | [274]     |
| Terranovaite                    | TER      | IZA-SC | y           | [275]     |
| Thomsonite                      | THO      | IZA-SC | y           | [276]     |
| Tournikite-like                 | TOL      | IZA-SC | y           | [277]     |
| ZSM-22 (Theta-1)                | TON      | IZA-SC | y           | [278]     |
| Tschörtnerite                   | TSC      | IZA-SC | y           | [279]     |
| TNU-9                           | TUN      | IZA-SC | y           | [280]     |
| Mu-18                           | UEI      | IZA-SC | y           | [281]     |
| UZM-5                           | UFI      | IZA-SC | y           | [282]     |
| IM-16                           | UOS      | IZA-SC | y           | [283]     |
| IM-17                           | UOV      | IZA-SC | y           | [284]     |
| IM-10                           | UOZ      | IZA-SC | y           | [285]     |
| IM-6                            | USI      | IZA-SC | y           | [286]     |
| IM-12                           | UTL      | IZA-SC | y           | [287]     |
| IM-20                           | UWO      | IZA-SC | y           | [288]     |
| VPI-8                           | VET      | IZA-SC | y           | [289]     |
| VPI-5                           | VFI      | IZA-SC | y           | [290]     |
| VPI-9                           | VNI      | IZA-SC | n - not fully connected | [290]     |
| VPI-7                           | VSV      | IZA-SC | y           | [291]     |
| Weinebeneite                    | WEI      | IZA-SC | y           | [292]     |
| YNU-5                           | YFI      | IZA-SC | y           | [293]     |
| Yugawaralite                    | YUG      | IZA-SC | y           | [294]     |
| ZAPO-M1                         | ZON      | IZA-SC | y           | [295]     |
All seventeen experimentally realised ice forms have been considered. Entirely ordered and entirely disordered structures were included in the dataset.

Table 4: Ice structures included in the database

| Mineral | Topology | ICSD Identifier | Included? | H-Ordering       | Reference |
|---------|----------|-----------------|-----------|------------------|-----------|
| Ice-Ic  | dia      | 64775           | y         | Disordered       | [296]     |
| Ice-Ih  | lon      | 64776           | y         | Disordered       | [297]     |
| Ice-II  | ict      | 15907           | y         | Ordered          | [298]     |
| Ice-III | kea      | 64771           | y         | Disordered       | [299]     |
| Ice-IV  | icf      | 201179          | y         | Disordered       | [300]     |
| Ice-V   | icv      | 14318           | y         | Disordered       | [301]     |
| Ice-VI  | edi      | -               | y         | Disordered       | [302]     |
| Ice-VII | itv      | 31868           | y         | Disordered       | [303]     |
| Ice-VIII| dia      | 1613533         | y         | Ordered          | [304]     |
| Ice-IX  | kea      | 1614920         | y         | Ordered          | [305]     |
| Ice-X   | itv      | -               | n         | Symmetric        |           |
| Ice-XI  | lon      | -               | y         | Ordered          | [306]     |
| Ice-XII | itv      | -               | y         | Disordered       | [307]     |
| Ice-XIII| icv      | 1672152         | y         | Ordered          | [308]     |
| Ice-XIV | itv      | 1672153         | n         | Partial ordering | [308]     |
| Ice-XV  | edi      | 1682436         | y         | Ordered          | [309]     |
| Ice-XVI | mtn      | 1442023         | y         | Disordered       | [310]     |
| Ice-XVII| unj      | -               | y         | Disordered       | [311]     |
The cyanide structures included are taken from the paper “Exploiting High Pressures to Generate Porosity, Polymorphism, And Lattice Expansion in the Nonporous Molecular Framework Zn(CN)₂”, Saul H. Lapidus, Gregory J. Halder, Peter J. Chupas, Karena W. Chapman.

Table 5: Cyanide structures included in the database

| Formula   | Phase | Description                  | Included? | Reference |
|-----------|-------|------------------------------|-----------|-----------|
| Zn(CN)₂   | I (dia) | interpenetrated diamondoid   | y         | [312]     |
| Zn(CN)₂   | II (dia) | distorted interpenetrated diamondoid | y         | [312]     |
| Zn(CN)₂   | dia    | diamondoid                  | y         | [312]     |
| Zn(CN)₂   | lon    | lonsdaleite               | y         | [312]     |
| Zn(CN)₂   | pyr    | pyrite                      | n - Octahedral Zn environment | [312]     |
| Name                        | Database Identifier | Clathrate Type | Included? | Reference |
|-----------------------------|---------------------|----------------|-----------|-----------|
| Acetylene clathrate         | 247246/NAHClI       | Type-I         | y         | [313]     |
| HPF₆·7.67H₂O                | 1706893*           | Type-I         | y         | [314]     |
| Methane clathrate           | 247245/NAHCEF       | Type-I         | y         | [313]     |
| Ethylene Oxide Hydrate      | 1150521/ETHYLO      | Type-I         | y         | [315]     |
| Propane clathrate           | 247247/NAHCP        | Type-II        | y         | [313]     |
| Methane-propane clathrate   | 247248/NAHCUV       | Type-II        | y         | [313]     |
| XeCCl₄·136D₂O               | 1194852/KELKUH      | Type-II        | y         | [316]     |
| Methane-adamantane clathrate| 247249/NAHDA        | Type-H         | y         | [313]     |

*In the absence of a CCDC Identifier, an ICSD Identifier is provided.*
Table 7: Inorganic structures included in the database

| Name                | ICSD Identifier | Included? | Reference |
|---------------------|-----------------|-----------|-----------|
| alpha-BeCl₂         | 92583           | y         | [317]     |
| beta-BeCl₂          | 92586           | y         | [317]     |
| alpha-ZnCl₂         | 26154           | y         | [318]     |
| beta-ZnCl₂          | 26153           | y         | [318]     |
| gamma-ZnCl₂         | 26152           | y         | [318]     |
| SiS₂                | 291208          | y         | [319]     |
| LiCo(CO)₂           | 1712939*        | y         | [320]     |

* Downloaded from the “Pauling File Multinaries Edition - 2012” database.[321]
Bibliography

[1] Phan, A.; Doonan, C. J.; Uribe-Romo, F. J.; Knobler, C. B.; O’Keeffe, M.; Yaghi, O. M. Acc. Chem. Res. 2010, 43, 58–67.
[2] Tian, Y. Q.; Yao, S. Y.; Gu, D.; Cui, K. H.; Guo, D. W.; Zhang, G.; Chen, Z. X.; Zhao, D. Y. Chem. - A Eur. J. 2010, 16, 1137–1141.
[3] Huang, X.-C.; Lin, Y.-Y.; Zhang, J.-P.; Chen, X.-M. Angew. Chemie Int. Ed. 2006, 45, 1557–1559.
[4] Tian, Y. Q.; Chen, Z. X.; Weng, L. H.; Guo, H. B.; Gao, S.; Zhao, D. Y. Inorg. Chem. 2004, 43, 4631–4635.
[5] Banerjee, R.; Phan, A.; Wang, B.; Knobler, C.; Furukawa, H.; O’Keeffe, M.; Yaghi, O. M. Science (80-. ). 2008, 319, 939–943.
[6] Park, K. S.; Ni, Z.; Côte, A. P.; Choi, J. Y.; Huang, R.; Uribe-Romo, F. J.; Chae, H. K.; O'Keeffe, M.; Yaghi, O. M. Proc. Natl. Acad. Sci. U. S. A. 2006, 103, 10186–10191.
[7] Tian, Y. Q.; Zhao, Y. M.; Chen, Z. X.; Zhang, G. N.; Weng, L. H.; Zhao, D. Y. Chem. - A Eur. J. 2007, 13, 4146–4154.
[8] Wu, T.; Bu, X.; Zhang, J.; Feng, P. Chem. Mater. 2008, 20, 7377–7382.
[9] Tian, Y. Q.; Cai, C. X.; Ren, X. M.; Duan, C. Y.; Xu, Y.; Gao, S.; You, X. Z. Chem. - A Eur. J. 2003, 9, 5673–5685.
[10] Sturm, M.; Brandl, F.; Engel, D.; Hoppe, W. Acta Crystallogr. Sect. B Struct. Crystallogr. Cryst. Chem. 1975, 31, 2369–2378.
[11] Rettig, S. J.; Storr, A.; Summers, D. A.; Thompson, R. C.; Trotter, J. Can. J. Chem. 1999, 77, 425–433.
[12] Müller-Buschbaum, K. Zeitschrift für Naturforsch. B 2015, 61, 792–798.
[13] Hayashi, H.; Côté, A. P.; Furukawa, H.; O’Keeffe, M.; Yaghi, O. M. Nat. Mater. 2007, 6, 501–506.
[14] Fu, Y. M.; Zhao, Y. H.; Lan, Y. Q.; Wang, Y.; Qiu, Y. Q.; Shao, K. Z.; Su, Z. M. Inorg. Chem. Commun. 2007, 10, 720–723.
[15] Rettig, S. J.; Sánchez, V.; Storr, A.; Thompson, R. C.; Trotter, J. J. Chem. Soc. Dalt. Trans. 2000, 21, 3931–3937.
[16] Lorente, M. A. M.; Dahan, F.; Sanakis, Y.; Petrouleas, V.; Bousseksou, A.; Tuchagues, J. P. Inorg. Chem. 1995, 34, 5346–5357.
[17] Masciocchi, N.; Ardizzoia, G. A.; Brenna, S.; Castelli, F.; Galli, S.; Maspero, A.; Sironi, A. Chem. Commun. 2003, 3, 2018–2019.
[18] Zhang, J.; Wu, T.; Zhou, C.; Chen, S.; Feng, P.; Bu, X. Angew. Chemie - Int. Ed. 2009, 48, 2542–2545.
[19] Banerjee, R.; Furukawa, H.; Britt, D.; Knobler, C.; O’Keeffe, M.; Yaghi, O. M. J. Am. Chem. Soc. 2009, 131, 3875–3877.
[20] Han, J. Y.; Fang, J.; Chang, H. Y.; Dong, Y.; Liang, S. Acta Crystallogr. Sect. E Struct. Crystallogr. Sect. E Struct. Reports Online 2005, 61.,
[21] Liu, Y.; Kravtsov, V. C.; Eddaoudi, M. Angew. Chemie - Int. Ed. 2008, 47, 8446–8449.
[22] Masciocchi, N.; Bruni, S.; Cariati, E.; Cariati, F.; Galli, S.; Sironi, A. Inorg. Chem. 2001, 40, 5897–5905.
[23] Lehner, R.; Seel, F. Z. anorg. alg. Chem. 1978, 444, 91–96.
[24] Rettig, S. J.; Storr, A.; Summers, D. A.; Thompson, R. C.; Trotter, J. J. Am. Chem. Soc. 1997, 119, 8675–8680.
[25] Lehner, R.; Seel, F. ZAAC à J. Inorg. Gen. Chem. 1980, 464, 187–194.
[26] Wang, B.; Côté, A. P.; Furukawa, H.; O’Keeffe, M.; Yaghi, O. M. Nature 2008, 453, 207–211.
[27] Tian, Y.-q.; Cai, C.-x.; Ji, Y.; You, X.-z.; Peng, S.-m.; Lee, G.-h. Angew. Chem. 2002, 114, 1384–1386.
[28] Spek, A. L.; Duisenberg, A. J. M.; Feiters, M. C. Acta Crystallogr. Sect. C Cryst. Struct. Commun. 1983, 39, 1212–1214.
[29] Liu, Y.; Kravtsov, V. C.; Larsen, R.; Eddaoudi, M. Chem. Commun. 2006, 14, 1488–1490.
[30] Morris, W.; Doonan, C. J.; Furukawa, H.; Banerjee, R.; Yaghi, O. M. J. Am. Chem. Soc. 2008, 130, 12626–12627.
LiCo(CO)₄ (LiCo[CO]₄) Crystal Structure: Datasheet from "PAULING FILE Multinaries Edition à 2012" in SpringerMaterials (https://materials.springer.com/isp/crystallographic/docs/sd1712939)"