Same or different – that is the question: identification of crystal forms from crystal structure data

Same or different – that is the question: identification of crystal forms from crystal structure data†

Pietro Sacchi,[a] Matteo Lusi,[b]* Aurora J. Cruz-Cabeza,[a]* Elisa Nauha[c] and Joel Bernstein[c]

The comparison of new structural data with crystal structures stored in the Cambridge Structural Database (CSD) is becoming ordinary when dealing with recognition and identification of original solid forms. Indeed, this is a consequence of the ever-increasing number of deposited crystal structures and of the development of new structural comparison tools. We performed a CSD search for single-component crystal structures of organic compounds, resulting in 8102 refcode families with more than one entry. The Crystal Structure Similarity (CSS) and the powder pattern similarity (PXS) tools implemented in the CSD Materials Software were used to quantitatively assess the degree of similarity for pairs of structures in each family, in order to identify potential polymorphs or structure redeterminations. A total of 47422 pairwise comparisons shows that in most cases (84%) both comparison methods agree in discriminating between the two categories. An analysis of selected examples for which the comparison methods disagree (16%) reveals that ambiguity is due to several reasons including experimental conditions employed for diffraction data collections (i.e. temperature and pressure) or crystallographic errors during structure solution. For those borderline cases, the decision to designate a pair of entries as redeterminations of the same structure or polymorphs still relies on the combination of several comparison techniques performed by an expert practitioner.

1. Introduction

A common understanding in chemistry is that molecular substances might crystallise into different forms broadly classified as polymorphs,[1, 2] solvates,[3] salts, co-crystals, tautomers[4-9] and solid solutions.[10, 11] Other forms might appear in response to external stimuli (thermal, electromagnetic, mechanical etc.)[12-14] or, with time, upon spontaneous conversion of a metastable one.[15] Each solid form constitutes a different physical phase, with specific chemical and physical properties – including solubility, thermal stability, reactivity, magnetism, conductivity *etc*. These properties depend on the chemical nature of the molecular components, their stoichiometry and their arrangement in space relatively to one another.[16-18]

The vastness of the solid form landscape represents both a challenge and an opportunity.[19] On one side, the potential for multiple products requires a careful optimisation of the crystallization conditions for the targeted product. On the other hand, crystalline phases might be engineered with improved properties for a given application.[20] The successful realization of both endeavours requires the correct identification of all the solid products. Furthermore, if the sought out properties depend of a particular arrangement of a molecule in space (structure), there might be an interest in identifying and isolating isostructural materials regardless of their chemical composition. Indeed, characterization and retrieval of crystal structures is as critical to a crystal engineer as chemical characterization is to a synthetic chemist.

Arguably, the method of choice for the characterization of crystalline materials insists on single crystal diffraction data, usually obtained with in-house X-ray diffractometers. Recently efficient strategies have been developed to solve crystal structures from powder samples as well.[21, 22] Single crystal and powder X-ray diffraction (SXRD and PXRD respectively) have contributed to the determination of over 1 million crystal structures that have been collected in over 50 years by the Cambridge Structural Database (CSD) and easily accessible through their suite of software.[23-25] The availability of such information implies that, in many cases, the identification of a crystallization product involves a simple search of the unit cell values of the new crystal with those recorded on the database. Alternatively, a comparison of the PXRD pattern measured for the new product with those calculated from the crystal structures in the database can help phase and purity identification.

In fact, common experience reveals that, even with accurate data in hand, structural differences could be so subtle that their correct identification might be difficult. In single-crystal diffraction experiments, phenomena such as twinning and disorder can further complicate the correct structure assessment.[26] In powder samples, particle size, shape and strain as well as impurities might strongly affect diffraction patterns, which, indeed, depend on the sample history and on the data collection conditions.[27]

The issue is well known to pharmaceutical and chemical companies that must guarantee the consistent production of the same form, as well as protect their intellectual property from infringement. The difficulty in assessing polymorphism can be exemplified by the case of form II aspirin,[28, 29] whose structural relationship was explained by visual inspection of the overlapping diffraction patterns generated for each phase.[30] In non-stoichiometric hydrates such as Paroxetine HCl, the complete characterization would require a chemical analysis as

[a.] Department of Chemical Engineering and Analytical Science, University of Manchester, Manchester, M13 9PL, United Kingdom. E-mail: aurora.cruzcabeza@manchester.ac.uk
[b.] SSPC, the SFI Research Centre for Pharmaceuticals, Bernal Institute, Department of Chemical Sciences, University of Limerick, Co. Limerick, Ireland. E-mail: matteo.lusi@ul.ie
[c.] NYU Abu Dhabi, Abu Dhabi, United Arab Emirates

well.[31] In many instances, though, the raw data are not available and the crystallographer or chemist has to make a decision based on processed data deposited in publications. In patent litigations, the notorious cases of Cefadroxil and Terazosin hydrochloride can serve as examples of the possible consequences.[18]

In this work, we study single component crystals with more than one crystal structure deposited in the CSD. These different crystal structures (those being either redeterminations or polymorphs) are then compared using two quantitative methods: the CSS and the PXS tools as implemented in the CSD systems. The purpose of this CSD analysis was to identify borderline cases where methods differed in their assessment. Examples of structures which are difficult to assess are then analysed and discussed with a number of other methods. Overall, we highlight the potential and limitations of various methods in assessing closely related crystal structures and how in those cases, the combination of various methods may be a necessity to arrive to informed conclusions.

2. Comparison Methods at a Glance

Besides unit cell and PXRD[32] comparison, other methods have been developed to help the correct phase identification. In particular, when reference crystal structures are available, a projection of the molecular Hirshfeld surfaces into 2D Fingerprint Plots (FP)[33] is often employed. More recently, the CCDC implemented two quantitative tools, namely the Crystal Structure Similarity (CSS) and the PXRD Similarity (PXS) calculators, as part of the Mercury suite of programs.[34] All these methods are available in widespread software making structure comparison fast and straightforward. Though each has some specificity as discussed below.

2.1 Unit cell comparison

When two structures are fully determined, a comparison of their unit cell metrics (and space groups) gives a first indication of how similar the structures could be. Correct scientific practice prescribes to report measurable quantities with an error - or a standard deviation when the value is calculated as an average of multiple measurements. Then, two dimensions are considered the same if their values fall within the accepted error or standard deviation. Unfortunately, XRD data can be collected under different temperature and pressure conditions causing significant variation of the unit cell parameters without causing any structural and phase transformation.

Moreover, unit cells can be reported in different space groups further complicating the comparison of structures. In some cases, different symmetries arise from poor modeling of disorder as in the case of IPEBIO03 ($Cmmm$ a 11.771(7) b 12.083(7) c 3.781(2)) and IPEBIO02 ($Pban$ a 11.753(2) b 12.107(3) c 3.7959(8)).[35] In other instances, the space group difference derives by the correct application of crystallographic conventions. For example, the salt bis(1,3-dibenzyl-1H-imidazol-3-ium) tetrachloro-palladium(ii) must be reported as $P2_1/c$ at 100 K and $P2_1/n$ at RT despite being the same phase. In

this case it might be useful to look at the reduced cell dimensions. Indeed $P2_1/c$ and $P2_1/n$ are alternative representations of the same symmetry and the reduced unit cell for the two structures are nearly identical (ref code ZAZHAL02 and ZAZHAL01 respectively).[36]

Finally, two genuine distinct phases could be described by isometric unit cells as exemplified by the above-mentioned forms of aspirin (ACSALA14 and ACSALA13). Both forms have a $P2_1/c$ symmetry but the cell dimensions of the latter, a 12.095(7) b 6.491(4) c 11.323(6) $β$ 111.509(9), convert into those of the first one a 11.2776(2) b 6.5517(1) c 11.2741(2) $β$ 95.837(1) through the transformation matrix [1 0 1 2 0 −1 0 0 0 −1].[28, 29]

2.2 PXRD comparison and the PXS

If single crystal data are not available, the characterization of a solid material often involves powder diffraction data. Experimental datasets may be compared to model patterns, simulated from single crystal data, or to the experimental pattern measured for a reference sample. The method is particularly effective in revealing crystalline impurities and phase mixtures in bulk products. Indeed, powder diffraction is regarded as the golden standard for the characterization of crystalline samples.

Qualitative analysis usually focuses on visual comparison of the diffraction profiles. A numerical value such as the sum of the differences (or squared differences) between data points can be calculated to measure the similarity between two PXRD profiles. These values are used for example in Rietveld refinement.[27] Gilmore *et al.* proposed an integrated full-pattern analysis instead.[37] De Gelder *et al.* suggested that the normalised integral of a weighted cross-correlation function could minimise the negative effect of the latter but such hypothesis was not confirmed on experimental data.[32] Schmidt et al. has also used cross-correlation functions to compare powder patterns which has been implemented in the FIDEL software.[38] Both methods, integrated full-pattern analysis and cross-correlation functions, are very sensitive to impurities and preferred orientation and therefore are useful as pre-screening steps in automated, high-throughput lines for quality control and polymorph screening when one type of pattern is expected. On the contrary, their application may be less straight forward when comparing diffraction data collected under different conditions (although this appears to work well when comparing patterns calculated from CIF files). A version of the de Gelder method is implemented in Mercury and applied for the CSD analysis and the various examples used in this contribution.

2.3 Crystal Structure Similarity (CSS)

The CSS calculator is based on the COMPACK software[34, 39] and compares structures based on the determined atomic positions by comparing a cluster of molecules around a central one. This method is further explained in the methods section and was extensively used in our study. The software compares a cluster of molecules, 20 in our case, which usually extends beyond a single unit cell, thus compensating for missed symmetry or pseudo symmetry. The method, however, does not analyse

structures with disordered atoms since it is based on atomic positions. Indeed, when ordered structures are fully characterised and CIF files are available, CSS seems to be the ideal tool to compare crystal structures. Then, the question arises as to whether such calculation is always sufficient to identify polymorphism and, in general, how CSS compares to the other methods described above. Alternative programs are available to quantify cluster similarities but are not implemented in the database.[40]

2.4 Comparison of Hirshfeld Surfaces

Another easily accessible method to compare crystal structures is by making use of the Hirshfeld surfaces of molecules in a structure.[41] The intermolecular surface defined and the distance between atoms of adjacent molecules are calculated. The values are plotted on a two dimensional fingerprint plot (FP).[33] This method highlights very well the differences in the molecular environment (*i.e.* the intermolecular interactions in a compound) and newly developed graphic tools help visualizing such differences for a qualitative assesment.[42] Additional calculations can help identifying the different type of interactions in the structure. On the negative side, the realization and comparison of FP becomes difficult in the case of disordered structures and some information associated with unit cell size and symmetry might be overlooked.

3. Methods

3.1 CSD Analysis

Retrieval of the Dataset of (Single component) Refcode Families with Multiple Structures. The CSD (July 2018 version) was searched for single component organic crystal structures containing only common elements (H, D, C, O, N and halogens) through the Conquest interface. Polymeric structures, structures with errors, structures with no 3D coordinates and structures with disorder were excluded from the searches. Collection temperature and pressure were recorded for each entry. When no temperature was reported, a temperature of 298 K was assumed. Similarly, if no pressure was given, data were assumed to be collected at ambient pressure (0.0001 GPa). There were a few minor instances of "High Pressure" reported in the pressure entry information. Arbitrarily, we recorded those structures to have been collected at 1 GPa.

The resulting entries were then analysed by families: collections of entries that have the same chemical composition. In the CSD, each family is identified by a unique 6 letters refcode followed by a progressive 2-digit number for multiple entries. Multiple entries can either be polymorphs or redetermination of the same structure. Refcode families containing only one entry were then removed to return a dataset of 8102 families containing two entries or more.

Pairwise Comparison within Refcode Families. The CSD Python API was used to run the PXRD similarity algorithm (PXS) and the crystal structure similarity (CSS) algorithm as implemented in Materials Mercury on all the pairs of entries within each refcode family. *Vide infra* for details. A total of 47444 pairwise

comparisons were required for analysing our dataset of refcode families.

Of the 47444 structure comparisons performed, 22 (< 0.5%) gave errors due to, for example, missing coordinates for various atomic positions (*i.e.* TIETHE/TIETHE01 or DINICA11/DINICA14). These pairs were excluded reducing the dataset to 47422 comparisons.

3.2 Crystal structures comparison

PXRD similarity (PXS). Quantitative PXRD pattern analyses were performed with the PXRD comparison tool implemented in the CSD-Materials package. For each CIF file an ideal Cu Kα_1 (1.54056 Å) radiation powder pattern is calculated from 0.0 to 50.0° 2θ in steps of 0.02° and using a Pseudo-Voigt shape for diffraction peaks. The simulated PXRD patterns are then compared using a weighted cross-correlation function which is normalised[32] to return a numerical value between 0 and 1; with 1 indicating patterns identity whilst 0 being the limit for very dissimilar patterns. Experience suggests that values above 0.97 are to be expected for the same crystal form.[43] We have used a value of 0.965 as a reasonable cut-off above which two structures can be considered the same. The method was accessed via the CSD Python API as well as the CSD visualizer Mercury. In order to overcome some evident error with the deposited data (*vide infra*), atomic displacement parameters were eliminated from the CIF file before PXRD simulation.

CSD Crystal Structure Similarity (CSS) tool. The CSS tool was used as implemented in the CSD Python API and the CSD visualizer Mercury. For each CIF file a cluster of molecules is generated by picking a central molecule and its closest neighbours. The overlay of the two clusters being compared is then attempted. Once a good overlay is achieved, COMPACK returns the number of molecules which overlap within the tolerance parameters set. The software also returns the root mean square deviations in atomic positions (rmsd-r[N-match]) for the overlapping molecules in the cluster.

The default cluster size in the CSS tool is 15, which will use a central molecule with its 14 closest neighbours. Here the cluster size was extended to 20 molecules as experience suggests that such larger size is often necessary for structures with multiple molecules in the asymmetric unit. Indeed, in those cases different clusters can be generated. The default distance tolerances of 20% and angle tolerances of 20° were initially adopted. For comparisons affording a PXS of over 0.95, the distance and angle tolerances were increased to 50% and 50° respectively.

Hirshfeld surfaces. A visual comparison of Hirshfeld surfaces was carried out for a number of examples. Hirshfeld surfaces were generated using the software CrystalExplorer version 17.5.[44]

3.3 Calculation of rmsd for cell lengths (CL) and angles (CA)

The similarity between unit cells for selected pairs is expressed by calculating the root mean squared deviations of cell axes lengths *a*, *b* and *c* (rmsd-CL) and cell angles α, β and γ (rmsd-CA). The same applies for the parameters of the reduced cells.

$$rmsd - CL = \sqrt{\frac{(a_1 - a_2)^2 + (b_1 - b_2)^2 + (c_1 - c_2)^2}{3}} \qquad (1)$$

$$rmsd - CA = \sqrt{\frac{(\alpha_1 - \alpha_2)^2 + (\beta_1 - \beta_2)^2 + (\gamma_1 - \gamma_2)^2}{3}} \qquad (2)$$

3.4 DFT calculations

When needed, the molecular conformational energies were estimated using the Gaussian16 package.[45] All optimisations were performed on molecular structures extracted from the CSD2020 database at the B3LYP/6-31G** theory level and included the GD2 dispersion correction.[46]

4 Analysis of structure similarity in the CSD

The CSD search of single components, organic crystals produced 8102 refcode families containing multiple entries. A total of 47444 pairwise comparisons within each refcode family were performed and analysed to identify potential polymorphs or structure redeterminations. These were then reduced to 47422 after removing some defective entries for which only some atomic coordinates were provided. To this end the two quantitative tools available through the CCDC as implemented in Materials extension in Mercury (PXS and CSS), have been exploited, both accessed with the Python API.

The 47422 pairwise comparisons reveal that in almost 84% of the cases, the two quantitative methods agree - identifying the pair of entries either to be the same (*i.e.* redetermination of the same crystal structure - 47.4%, Group 1) or different (*i.e.* polymorphs - 36.4%, Group 2). In 14.1% of the comparisons (Group 3) the crystal structures compared are identified as being the same by the CSS algorithm but not by the PXRD similarity algorithm. In 2.2% of the comparisons (Group 4), the crystal structures compared are identified as being the same by the PXRD similarity algorithm but not by the CSS method. The

results are summarised in Table 1 and each case is discussed in detail below.

Table 1 Summary of the results obtained in comparing the crystal structures of single component systems belonging to 8102 refcode families with two or more refcodes.

	Group	Description	Comparisons	Percentage of Total
Methods Agree	1	Same by PXS & CSS	22 457	47.3%
	2	Different by PXS & CSS	17 264	36.4%
		Total Agree	**39 721**	**83.7%**
Methods Disagree	3	Same by CSS only	6 670	14.1%
	4	Same by PXS only	1 031	2.2%
		Total Disagree	**7 701**	**16.3%**
All		**Total**	**47 422**	**100%**

4.1 Group 1: Pairs of crystal structures identified as the same by both methods

Group one includes pairs of entries for which CSS comparison returns an overlap of 20 molecules and the PXRD similarity is higher than 0.965. 22457 pairs fall in this category. In 70% of the cases the PXRD similarity value is comprised between 1 and 0.99, as shown by the histogram in Fig. 1. A coarse analysis of the sources reveals that many of these data generally come from a duplication of the same crystal structure which was published in two different journals (XOFTET/XOFTET01 or JUSTEW/JUSTEW01). Usually the first publication is in a general chemistry journal whilst the second publication concerns to the crystal structure itself in a Crystal Structure report. In some instances, crystal structures are
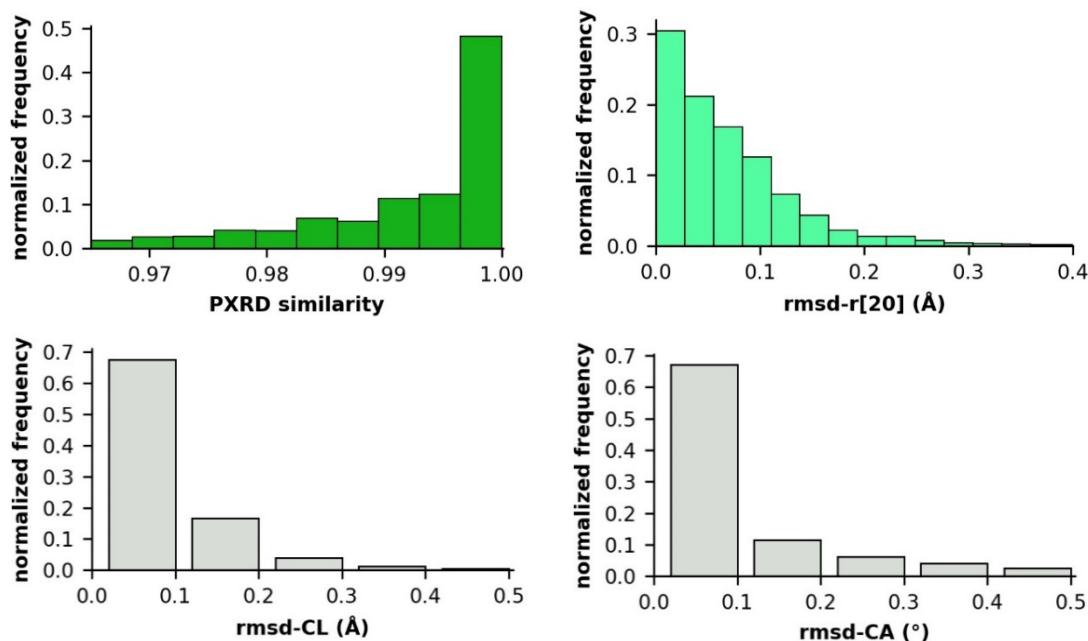

Figure 1 Histograms of PXRD similarities (top left) and rmsd-r[20] values (top right), for the 22 457 comparisons of structures which were identified as the same by both PXRD similarity and CSS. Histograms of rmsd-CL (bottom left) and rmsd-CA (bottom right) for a reduced 19 405 dataset of structures determined with the same space group and settings
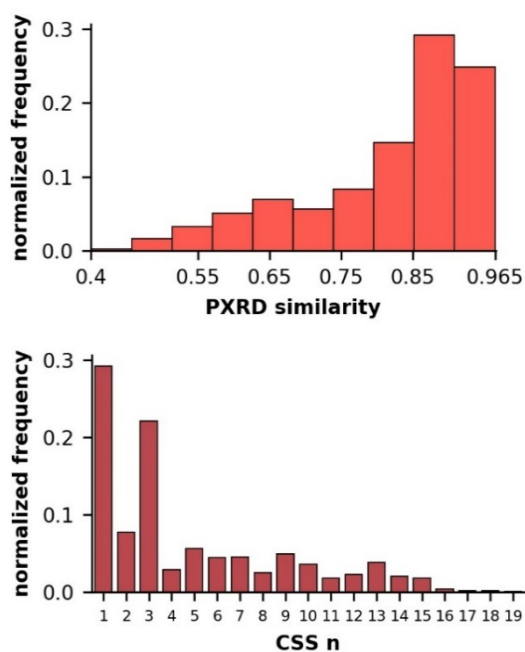
Figure 2 Histograms of PXRD similarities (top) and CSS (bottom) for the 17 264 pairs of structures identified as different by both PXRD and CSS similarity

Figure 3 Scatterplot of PXRD similarity versus difference in Pressure (top) and Temperature (bottom) conditions for the structure determination for 6670 pairs of structures in category 3

published first and the solution is revised by a different group. For example, Marsh reported that 10% of structures published in Cc with Z'=1 value should have been $C2/c$ with Z'=0.5.[47] Comparisons of structures with PXRD similarities less than 0.999 are usually redeterminations of the same crystal structures at different temperatures and pressures. The histograms of PXS, rmsd-r[20] and rmsd-CL values for these structure redeterminations reveal that when two structures are the same, the rmsd-r[20] value is usually less than 0.3 Å, the rmsd-CL less than 0.4 Å and rmsd-CA is less than 0.5°. The PXS values continuously decrease and the 0.965 value used here has been proposed as a reasonable cut-off.[40]

4.2 Group 2: Pairs of crystal structures identified as different by both methods

In Group 2, 17264 pairs of entries are recognised as different structures by both the PXRD and the CSS methods. Histograms of the PXS values and the numbers of overlapping molecules as presented in Fig. 2 reveal significant differences between most of the structures compared (CSS is mostly less than 15). These structures are in fact expected to be polymorphs.

In 57 cases (0.3%) the CSS algorithm matched more than 16 molecules. Of those, 23 correspond to data determined at different pressures (ΔP > 0.7 GPa), 18 correspond to pairs of structures determined at significantly different temperatures (ΔT > 99 K), and 16 to pairs of structures redetermined under very similar pressure and temperature conditions. These borderline cases can either be different phases, or different polymorphs or simply the same crystal structure determined under different conditions. Some of these will be discussed more in detail in Section 4.
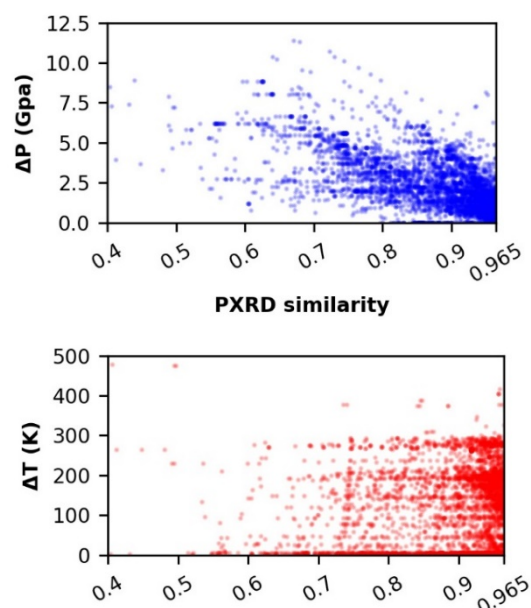
4.3 Group 3: Pairs of crystal structures identified as the same by CSS only

A set of 6670 pairs of entries are recognised as the same by the CSS method (n = 20) but as different by the PXS (< 0.965). The majority of these pairs were collected at significantly different pressures and/or temperatures.

In order to visualise the effect of pressure and temperature on the PXS values, the data were plotted in a scatter plot (Fig. 3). We can appreciate that, as the pressure difference increases, the PXRD similarity significantly decreases. The effect on temperature is not so clear and it strongly depends on the type of compound under study and its thermal expansion coefficients. Indeed it is known that pressure has a much profound effect of crystal compression.[48] Some examples in category 3 will be discussed in Section

4.4 Group 4: Pairs of crystal structures identified as the same by PXS only

The majority of the 1031 pairs in group 4 are reported as polymorphs in the literature. Indeed, the CSS correctly identifies these structures as different despite having a very similar PXRD pattern. Although the number in this group is relatively small (2.2% of the total), these examples highlight that polymorphism may be overlooked if its first identification is attempted by PXRD analysis alone. Hence, we may wonder how many new polymorphs with a PXRD very similar to known forms might have been overlooked.

The distribution of matching molecules by CSS values for the pairs in these categories shows a high degree of similarity (Fig.4). Some of these include *p*-aminobenzoic acid forms α and γ (AMBNAC06[49] and AMBNAC09[50]), mucochloric acid in the triclinic and orthorhombic forms (NISMUX01 and NISMUC02)[51]
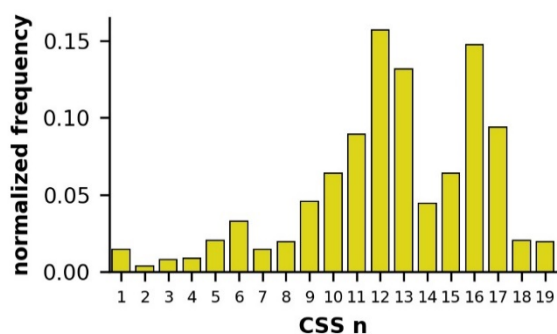
Figure. 4 Distribution of CSS values for pairs of structures affording a PXS greater than 0.965.

or chlorpropamide forms β2 and β3 (BEDMIG11 and BEDMIG12).[52]

5 Examples of selected pairs of structures compared with various methods

In order to explain, or perhaps exemplify, the issues with each method, a few cases were selected and individually screened. The results are described here.

5.1 Crystallographic errors

One of the most common causes of misidentification is due to crystallographic errors in structure solution. Three examples are summarised in Table 2. The differences between pairs of crystal data are tabulated together with the difference in temperature and pressure at which the XRD data collection was done, the results of the six similarity methods used in the comparisons and our conclusion statements. Our main conclusion for all these three pairs is that the compared structures are the same crystal form even though some of the methods fail to identify this. In each pair the differences appear to arise from crystallographic or data errors in one of the solved crystal structures. The first compound is a highly symmetrical molecule (Fig. 5) whose XRD data was initially collected and solved by a research group (ZEDCUG[53]) and resolved by a second group using the same data (ZEDCUG01).

Thus, ZEDCUG01[54] is a version of ZEDCUG[53] in which the space group has been updated from *Cc* to *C*2/*c* using the same dataset (*i.e.* the same unit cell metrics). PXRD similarity is unitary but no molecular match is found in the CSS similarity tool (1 out of 20 molecules). This is because the conformation
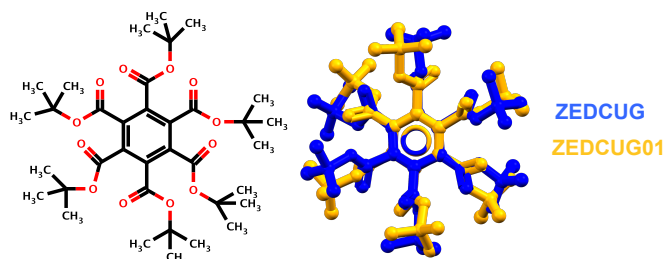

Figure 5 Molecular structure of compound in ZEDZUG/01 and overlay of the geometries resolved in the ZEDCUG (blue) and ZEDCUG01 structure solution

of the groups around the benzene ring change considerably (Figure 5) from the first to the second solution, thus the CSS similarity does not find a cluster match between these two structures. This example highlights the importance of making careful crystallographic choices when solving structures and how the same data can give rise to significantly different structure solutions. These are clearly two interpretations of the same XRD data. In cases with large and highly symmetrical molecules, crystallographers must be especially careful in the structure solution since many options may be possible from the same collected XRD data.

In the second example, TEVJIP[55] and TEVJIP01 differ in the way the nitrogen atoms on the central ring are located (Fig. 6a). As a consequence, although the PXRD similarity and cell/reduced cell comparisons are similar (Table 2), the comparison of specific atomic positions with the CSS tool and the interactions in the crystal structures analysed with the FP tool are significantly different (Fig. 6b). In this case, from a look at atomic displacement parameters and bond lengths and angles within the central ring it seems that the nitrogen atoms in TEVJIP01 are mistakenly placed. Moreover our knowledge as structural chemists suggests that this conformation would require the loss of four intramolecular CH-N bonds (which are present in TEVHIP) and a close proximity of the H-atoms from adjacent rings to about 1.9 Å (this is highlighted in the FP plot in TEVJIP01). To prove this, DFT-d was used to calculate the energy of the two conformations. This confirmed that the molecular conformation in TEVJIP is 51 kJ/mol more stable than in TEVJIP01. Beyond this, the R-factor of the TEVJIP01 structure is unreasonably large (15.5%), suggesting that there must be some errors in the structure.

In the third example of this group, TCLBEN02[56] and 07[57] are two determinations reported for the β polymorph of 1,2,4,5-tetrachlorobenzene at 173 and 200 K respectively. The latter is the deuterated molecule whose structure is refined from
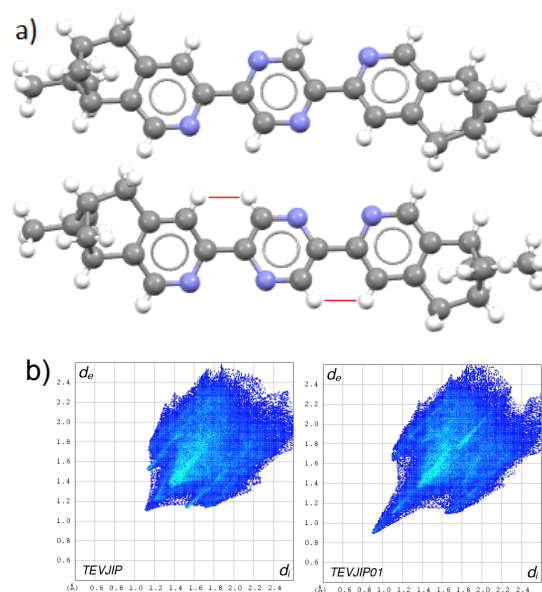

Figure 6 (a) Solved molecular structures and (b) FP of TEVJIP and TEVJIP01.

neutron powder diffraction. The unit cell and packing are essentially the same which is also reflected on the almost identical fingerprint plots (ESI) and an RMSD-r[20] of 0.021 Å. By contrast the PXRD similarity is calculated at 0.193 (Fig. 7) as one of the lowest in the series. The atomic displacement parameters for one record appear unusually large, which have a huge impact on the intensity of the diffraction peaks. Notably

the same data from an older version (CSD 2017) reports other APD values for the same entry. Another of such example of this is CILXAW[58] and CILXAW01[59]. As explained in the methods, this error can be avoided by removing the thermal ellipsoids before simulating and comparing the PXRD patterns for these structures. Must be noted though, that this method might overlook disordered phases in which ADPs are indeed large.

Table 2 Summary of pairs of crystal structures compared, the experimental conditions for XRD data collection and the results obtained with the similarity methods used for their comparison.

		Crystal Structure Comparisons		
Experimental		ZEDCUG-ZEDCUG01	TEVJIP-TEVJIP01	TCLBEN02- TCLBEN07
ΔT (K)		0 (RT-RT)	25 (175 -200)	25 (175 – 200)
ΔP (GPa)		0 (RP-RP)	0 (RP-RP)	0 (RP-RP)
Similarity Method				
Crystallography	Z'	1 – 0.5	1 - 1	0.5 – 0.5
	SG	*Cc – C2/c*	*P2₁ - P2₁*	*P2₁/n - P2₁/n*
	R-factor (%)	8.1 - NA	5.3 - 15.5	3.7 – 5.3
Cell	rmsd-CL (Å)	0.000	0.016	0.016
	rmsd-CA (°)	0.000	0.171	4.866
Red-Cell	rmsd-rCL (Å)	0.000	0.016	0.016
	rmsd-rCA (°)	0.000	0.171	0.008
PXRD		1.000	0.999	0.193 (improved after correction)
CSS (20 mol)	N-match	1	2	20
	rmsd-r (Å)	1.857	0.826	0.021
FP plots		Small Differences	Significant Differences	Identical
Issue		Wrong SG in ZEDCUG	Wrong position of a N atom in TEVJIP01	Wrong thermal ellipsoids stored in the CSD.
Conclusion		Same form.	Same form.	Same form.

5.2 Temperature and Pressure effects

The second set of examples explores the impact of temperature and pressure variation on the different indicators of isomorphic forms and polymorphism. The three examples selected are summarised in Table 3. According to literature evidences, these pairs of entries represent redeterminations of the same crystal form under significantly different experimental conditions. In all these cases, the PXS value is below 0.965 suggesting a possible polymorphic relationship between the two forms. These low values are justified by the significant changes in the PXRD

patterns due to large unit cell expansions or contractions as a function of temperature or pressure.

The first comparison looks at the crystal structure of acetic acid (orthorhombic polymorph) determined at two different temperatures: 175 K (ACETAC)[60] and 40 K (ACETAC07).[61] Changes in the unit cell parameters, FP plots and simulated PXRD are, however, significant. In this pair, thus, a comparison of the PXS values identifies the pair as polymorphs (Table 3 and Fig. 8). Comparison of the molecular packing with the CSS, however, returns 20 molecules in common with a low rmsd-
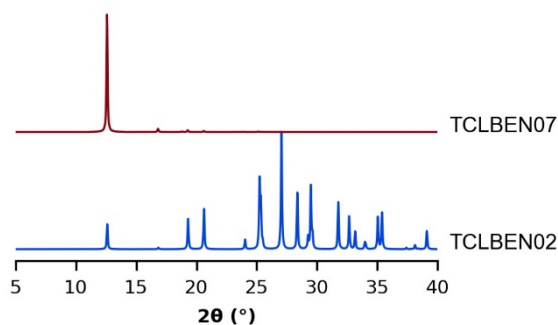

Figure 7 Simulated PXRD patterns for TCLBEN07 (red upper) and TCLBEN02 (blue lower) from the structures in the CSD2020 geometries resolved in the ZEDCUG (blue) and ZEDCUG01 structure solution
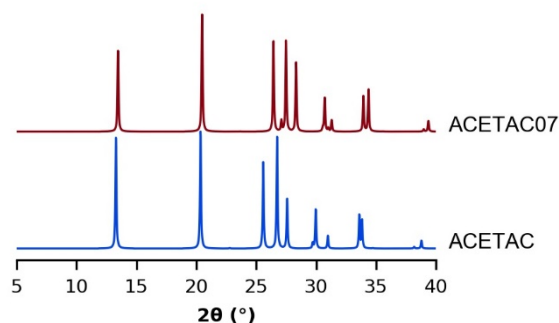

Figure 8 Simulated PXRD patterns for ACETAC07 (red upper) and ACETAC (blue lower). geometries resolved in the ZEDCUG (blue) and ZEDCUG01 structure solution.

r[20] of 0.162 Å confirming thus a very high structural similarity between these two structures.

Table 3 Summary of pairs of crystal structures compared, the experimental conditions for XRD data collection and the results obtained with the similarity methods used for their comparison.

		Crystal Structure Comparisons		
Experimental		ACETAC-ACETAC07	DIJHUA-DIJHUA01	JAYDUI-JAYDUI06
ΔT (K)		135 (175 − 40)	170 (123-293)	270 (30 − 300)
ΔP (GPa)		0 (RP − RP)	0 (RP − RP)	6 (0 - 6)
Similarity Method				
Crystallography	Z′	1 - 1	1 - 1	1 - 1
	SG	$Pna2_1$ - $Pna2_1$	$P2_1/n$ - $P2_1/n$	$P2_1/n$ - $P2_1/n$
	R-factor (%)	16.1 − 4.0	3.6 − 7.2	4.7 − 7.6
Cell	rmsd-CL (Å)	0.133	1.130	0.778
	rmsd-CA (°)	0.000	0.117	0.462
Red-Cell	rmsd-rCL (Å)	0.133	1.130	0.778
	rmsd-rCA (°)	0.000	0.117	0.462
PXRD		0.928	0.944	0.311
CSS (20 mol) si	N-match	20	20	20
	rmsd-r (Å)	0.162	0.932	0.587
FP plots		Significant Differences	Some Differences	Significant Differences
Issue		Large thermal expansion	Large thermal expansion	Large temperature difference and pressure contraction
Conclusions		Same form with significant expansion.	Same form with significant expansion.	Same form with significant compression.

The second example corresponds to the crystallographic data for 2-Cyano-3-(dimethylamino)-3-hydroxy-N-phenylacrylamide collected at 123 and 293 K (DIJHUA and DIJHUA01 respectively)[62]. In the two datasets, the cell symmetry is maintained although the variation in cell metrics (Table 3) indicates a large negative thermal expansion (α_b = -5x10^{-4} K^{-1}). The calculated rmsd-CL > 1 Å. The variation in cell metrics translates into significant differences in the PXRD patterns (Fig. 9) with a PXS value of 0.944. Some differences are also appreciated in the FP plots due to the rotation of the phenyl and methyl groups (ESI). Here it is worth highlighting that Hirschfield surfaces are very sensitive to the position of H atoms, which, in turn are often misplaced in a crystal solved by X-ray radiation.[63-65] Nevertheless, even when the most ambiguous H atoms are removed from the structure, the FP plots show some evident differences (ESI). The CSS similarity returns 20 molecules in common but with a high rmsd-r[20] of 0.932 Å, which, despite being well above the typical rmsd-r[20] for redeterminations (< 0.3 Å), suggests a correspondence for the molecules in the two clusters and, essentially, isostructurality (Figure 9). In the original article the two structures are considered as isomorphous.[62]

The final example corresponds to the crystal structure of propane as determined at 30 K and room pressure (JAYDUI)[66] and at room temperature and 6 GPa (JAYDUI06).[67] PXRD profiles look similar although considerably shifted to wider angles in the
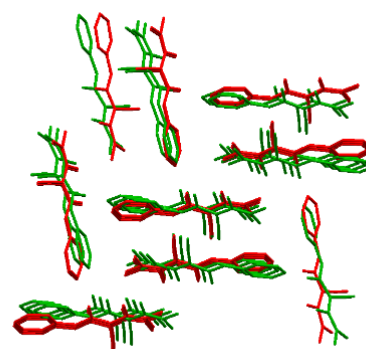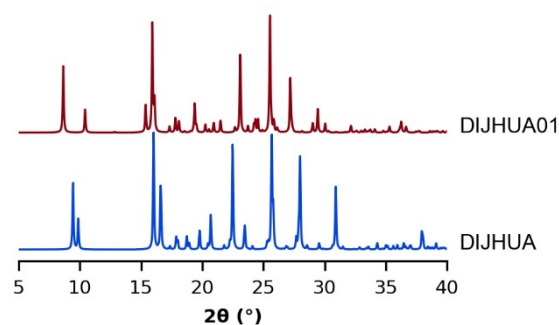

Figure 9 Top: Calculated powder patterns of DIJHUA (blue) and DIJHUA01 (red). Right: Overlay of the 20 molecule clusters in the CSS tool for DIJHUA and DIJHUA01.
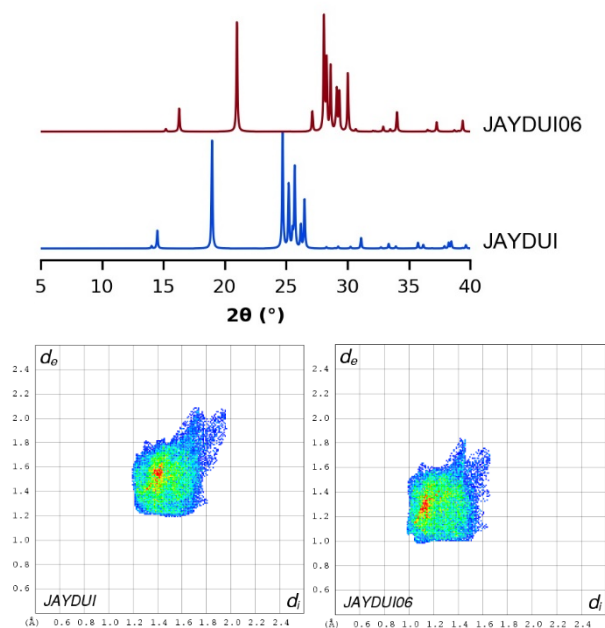
Figure 10 Top: Calculated powder patterns of JAYDUI (blue) and JAYDUI06 (red). Bottom: FP plots for JAYDUI and JAYDUI06.

Figure 11 Top: calculated powder patterns of MOSTIX (blue) and MOSTIX01 (red). Bottom: FP of MOSTIX and MOSTIX01.

high-pressure structure (Fig. 10), the unit cell lengths changing significantly (Table 3). The PXS comparison tool returns a similarity value of just 0.311. The CSS tool returns 20 molecules in common with a relatively high rmsd-r[20] of 0.587 Å. The FP plots are also similar but squashed towards shorter contacts at high pressure (Fig. 10) Also in this case several complementary methods reveals that the structures are probably the same form but with significant differences due different experimental conditions. The structures of benzene BENZEN17,[68] RT and 0.97(5) GPa and BENZEN24,[69] 773 K and 8.3 GPa show the same behaviour.

5.3 High versus Low Temperature Phases

Here three examples describe cases of different crystal phases at high and low temperatures (Table 4).

The radical 8-Chloro-4-methyl-4H-bis((1,2,3)dithiazolo)(4,5-b:5',4'-e)yridine-3-yl is reported at 25 and 293 K (MOSTIX[70] and MOSTIX01[71] respectively). The unit cells are almost isometric with the exception of minor anisotropic thermal expansion (ESI). Thus rmsd-CL and rmsd-rCL are small. The calculated PXRDs show only minor differences in peak position and intensity (Figure 11), but the similarity value is relatively low at 0.92. FP plots appear different and close structural analysis shows a rotation of 60° for the hydrogen atoms in the methyl group. In fact, when those atoms are constrained in the same positions some of the differences observed in the fingerprint plots fade (Fig. 11). The CSS calculation reports a match of 20 molecules with a rmsd-r[20] of 0.527 Å but, the two structures are reported as polymorphs and such conclusion is justified by a change in the magnetic susceptibility upon cooling.[71]

The structures of 4-Cyano-1,3,2-benzodithiazole (ECOBUU and ECOBUU01) were collected at 293 and 180 K respectively.[72] The unit cells have the same symmetry, but the *a* axis in
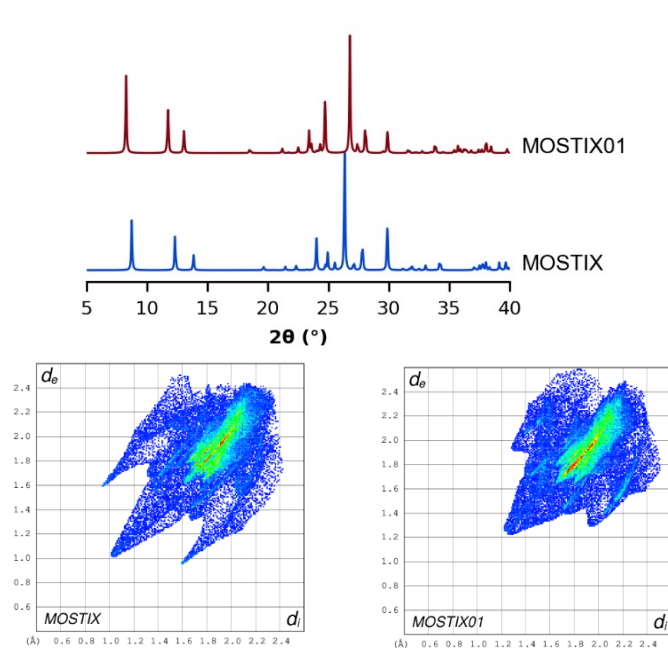
ECOBUU is half than in ECOBUU01 whilst the other axes are roughly the same. Consequently, the low temperature cell has a Z′ of 2 and the high temperature cell has a Z′ of 1. Despite such difference the PXRD patterns are extremely similar (R = 0.97) except for the position of the peaks at around 25° in 2θ (ESI). In this case differences are more evident in the fingerprint plots due to the number of independent molecules (Fig. 12). CSS indicates a 20-molecule cluster match with a rmsd-r[20] of 0.2 Å. Once again, experimental data confirm the existence of a phase transition upon cooling.[72]

In the final case, we analyse forms I and II of the drug compound Simvastatin (EJEQAL08 and EJEQAL02). The PXS in

Figure 12 Fingerprint plots of ECOBUU and ECOBUU01 for its two symmetry independent molecules.

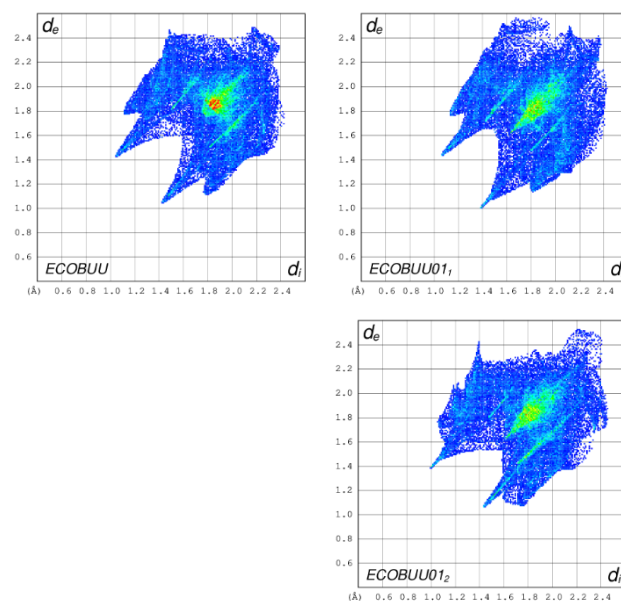these two forms is high (0.988) indicating that these might be the same crystal structure. However, the CSS algorithm only returns one molecule in common. Form II (EJEQAL02) is known to be the room temperature phase, which transforms to form I upon cooling at 275 K. The phase change is confirmed by a variation in the heat capacity.[73] A further analysis reveals that the major structural change involves the conformation around the peripheral group of the compound (see Fig. 13 left). Notably, because of this conformational change, the CSS fails to identify a cluster overlay despite the fact that the majority of atoms in both forms remain in the same position (Figure 13 right). Crystal forms that are related by small conformational changes without a major structure rearrangement have been referred to as conformational polymorphs.[74]They are usually enantiotropically related and interconvert with temperature.
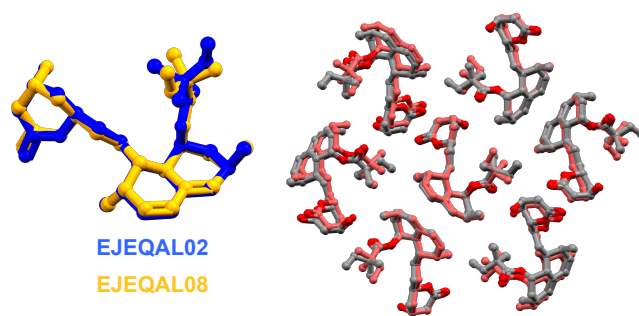

EJEQAL02
EJEQAL08

Figure 13 One molecule (left) and 20 molecules (right) overlay for forms I and II Simvastatin (EJEQAL08 and EJEQAL02 respectively)

Table 3 Summary of pairs of crystal structures compared, the experimental conditions for XRD data collection and the results obtained with the similarity methods used for their comparison.

		Crystal Structure Comparisons		
Experimental		MOSTIX-MOSTIX01	ECOBUU-ECOBUU01	EJEQAL02-EJEQAL08
Polymorph name (literature)		LT Polymorph- HT Polymorph	HT Polymorph – LT Polymorphs	Polymorph II – Polymorph I
ΔT (K)		268 (25 – 293)	113 (293-180)	30 (258-298)
ΔP (Gpa)		0 (RP – RP)	0 (RP – RP)	0 (RP – RP)
Similarity Method				
Crystallography	Z'	1 – 1	1-2	1-1
	SG	$P2_12_12_1 – P2_12_12_1$	$P2_1/c – P2_1/c$	$P2_12_12_1 – P2_12_12_1$
	R-factor (%)	4.9 – 3.1	6.2-3.9	14.7 – 7.9
Cell	rmsd-CL (Å)	0.689	**2.025**	0.525
	rmsd-CA (°)	0.000	0.396	0.000
Red-Cell	rmsd-rCL (Å)	0.689	**2.025**	0.525
	rmsd-rCA (°)	0.000	0.396	0.000
PXRD		**0.917**	0.970	0.987
CSS (20 mol) si	N-match	20	20	**1**
	rmsd-r (Å)	0.527	0.222	0.842
FP plots		**Small Differences**	**Small Differences**	Small Differences
Further Characterisation		Different magnetic susceptibilities	Phase Transition (cell doubling) on Cooling	Phase transition with conformational change on cooling
Conclusions		**High-Low temperature phases**	**High-Low temperature phases**	**High-Low temperature conformational phases**

5.4 Polymorphs with high Similarity

In the last set of examples, we present true polymorphic structures that are determined at the same temperature and appear isostructural by comparison of PXRD patterns. Four of such pairs of polymorphs are given in Table 5, including *p*-aminobenzoic acid forms α and γ, a system which has been discussed in length elsewhere.[75]

Comparison of simulated PXRD patterns alone would have resulted in the erroneous assumption that the compared structures are all the same. In all cases, the PXS is above 0.965. For the third pair, DAWFUE-DAWFUE01,[76] the PXS is close to the limit usually chosen for differentiating structures (0.972). Simulated PXRDs for the pairs are plotted in Fig. 14 for visual inspection. The similarity of the patterns is very high with only small differences apparent to the naked eye. Perhaps this evidence, together with the awareness that the data are simulated from model that are determined under the same condition, may have revealed the authenticity of the polymorphism. The differences in the PXRD patterns, however, are much smaller than those arising due to temperature or pressure variations that are described in the previous section. Remarkably the PXS value for the two forms collected at the same temperature (DAWFUE01 and DAWFUE02) is lower than the one for the same forms measured at 20 K difference (DAWFUE-DAWFUE02) 0.972 and 0.998 respectively.

Given that experimentally measured PXRDs suffer from preferred orientation, impurity effects or particle

size/amorphous contents effects, it would not be surprising that polymorphs with such high PXRD similarity may be passed unrecognised.

When these pairs of structures are compared with the CSS tool (Table 5), it becomes clear that these crystals are indeed true polymorphs with a high degree of similarity (they all have common layers of molecules in common). Fabian and Kalman have referred to these pairs as polymorphs with 1D or 2D isostructurality.[77]

6. Discussion

Most of the time the recognition of different crystal structures and phases from CIF files is straightforward. Unit cell metrics, FP plots, PXS and CSS directing towards the same answer. In a considerable amount of instances (over 16% according to our CSD analysis), however, different methods give opposite answers and each comparison method results sensitive to certain factors.

The most obvious causes of ambiguity are crystallographic errors. Such errors may be difficult to identify since they originate in crystallographic choices adopted during structure solution. Modern diffractometers are almost completely automated perhaps inducing overlooking causes of errors, especially when crystal systems show disorder. Other causes of errors are poor chemical analysis of minor components in non-stoichiometric crystals. In those cases, Andrew Bond reminds us of the importance of a detailed solution by an expert crystallographer.[78] High R-factors, and unusual symmetries, short contacts or conformations may flag possible errors in the structure solution.

Beside the evident cases of crystallographic errors, the examples selected in this contribution highlight that FP plots and CSS are particularly sensitive to misplaced atoms while PXS and unit cell metrics are strongly affected by pressure and temperature conditions during data collection. In particular, relatively to the quantitative PXS and CSS methods, the effect of different factors is summarised in Table 6.

PXS seems particularly useful to highlight real changes in symmetry between structures as well as for comparing structures with atomic disorder or partially resolved atomic positions (for which other methods simply cannot be used).

Moreover, the comparison can be performed for experimental PXRD patterns, although impurities and preferred

Table 4 Summary of pairs of crystal structures compared, the experimental conditions for XRD data collection and the results obtained with the similarity methods used for their comparison.

|		Crystal Structure Comparisons			
Experimental		AMBNAC07-AMBNAC09	DIWKOK-DIWKOK01	DAWFUE-DAWFUE01	XOCVAN-XOCVAN02
		Form α – Form γ	Monoclinic - Orthorhombic	Form β – Form α	
ΔT (K)		1 (101 – 100)	0 (173-173)	0 (300-300)	22 (122-100)
ΔP (GPa)		0 (RP – RP)	0 (RP – RP)	0 (RP – RP)	0 (RP – RP)
Similarity Method					
Crystallography	Z'	2 - 2	1 - 1	1 - 1	1-2
	SG	$P2_1/n$ - $Pna2_1$	$P2_1$ - $P2_12_12_1$	$P2_1/c$ - $Pca2_1$	$P2_1$ - $P2_1$
	R-factor (%)	5.1 - 5.1	4.0 – 4.6	6.3 - 6.0	2.2 – 9.6
PXRD		0.991	0.982	0.972	0.984
CSS (20 mol)	N-match	18	14	9	9
	rmsd-r (Å)	0.097	0.023	0.251	0.082
FP plots		Very Similar	Very Similar	Very Similar	Very Similar
Conclusions		Different polymorphs	Different polymorphs	Different polymorphs	Different polymorphs
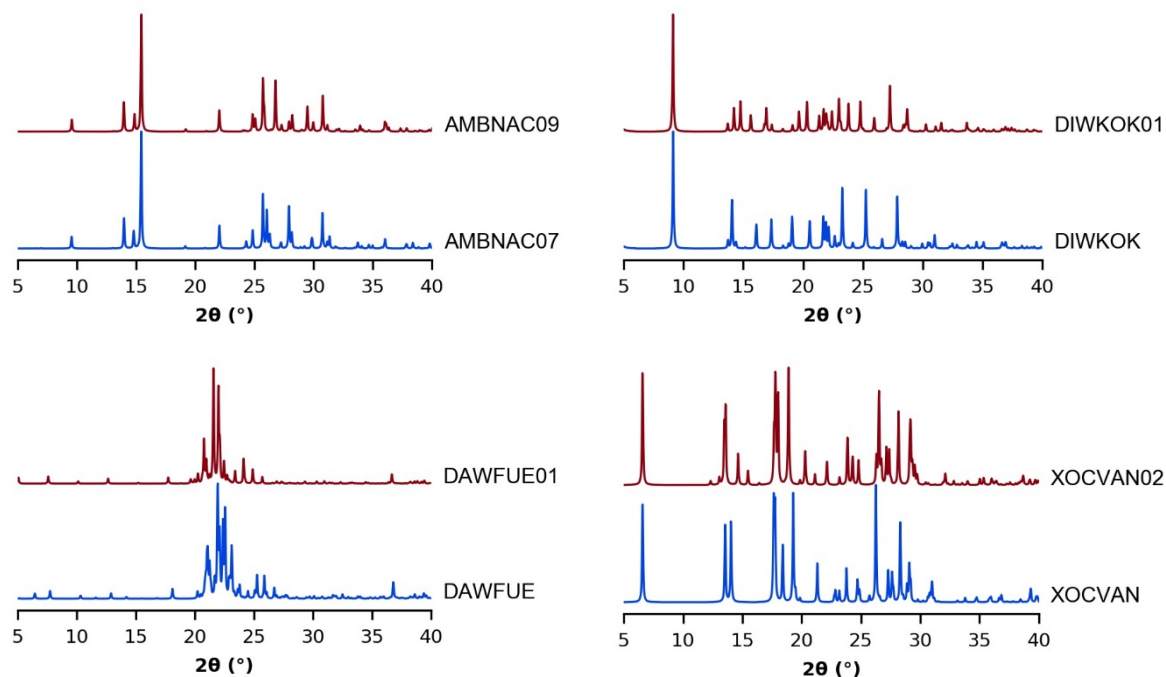
Figure 14 Simulated PXRD for the crystal structures with CSD refcodes (upper left) AMBNAC07 (blue, α-form) and AMBNAC09 (red, γ-form) (upper right) DIWKOK and DIWKOK01 (lower left) DAWFUE and DAWFUE01 and (lower right) XOCVAN and XOCVAN02

orientation need to be considered in those cases which may distort the outcome.

PXS, however, is very sensitive to large anisotropic expansion and thus the method may fall short when comparing structures collected under different temperature and pressure conditions. In this view, the cross correlation function embedded in the CSD algorithm does not always correct for strong peak shifts and isostructural entries are recognised as polymorphs.[32,43] As discussed previously, the impact of temperature and pressure on the PXS is very much case dependent. This is illustrated by the changes in PXS values as a function of ΔT and ΔP for a number of selected systems (Fig. 15). For example, the effect of temperature on the PXS of acetic acid (ACETAC) is considerably higher than on the PXS of dl-alanine (DLALNI05). The effect of pressure is even more pronounced and, again, impacts some crystal structures more than others. For example, it is far more marked in dl-alanine (DLALNIN05) than in dl-serine (DLSERN16). Comparison of PXRD

patterns calculated for structures collected at the same temperature and pressure will eliminate that risk.

In this work we noted that PXS is also sensitive to differences in thermal ellipsoids, some of which are particularly large (an example of this is given in the previous section). In our analysis the issue was avoided by removing thermal parameters from all the entries and all PXRD simulations and comparisons were performed using atomic positions alone. Whilst this approach eliminates the problem for a good number of comparisons, real differences due to order disorder transitions might be overlooked.

Ultimately to eliminate the effect of both, different temperature and pressure or thermal ellipsoids, it has been suggested that structures can be relaxed with an appropriate energy model before being compared.[33, 41] In those conditions, multiple data for the same form should yield a PXS value close to 0.99. The PXS method, however, can fail the identification of real cases of polymorphism when the usual cut-off of 0.965 is

Table 5 Summary of borderline examples in the various comparison groups and which methods fail or work.

| Comparison Group	Structures	Fails	Works
Crystallographic Errors	Errors	No particular method works or fails. It depends on the error.	
High versus Low Temperature or Pressure Structures	Redeterminations under different conditions	PXS	CSS
High versus Low Temperature or Pressure Phases	Different Phases	PXS or CSS alone	PXS & CSS with Thermal Analysis
High versus Low Temperature Conformational Phases	Different Conformational Phases	CSS	PXS
Different Polymorphs with high similarity	Different Polymorphs	PXS can be corrected visually	CSS
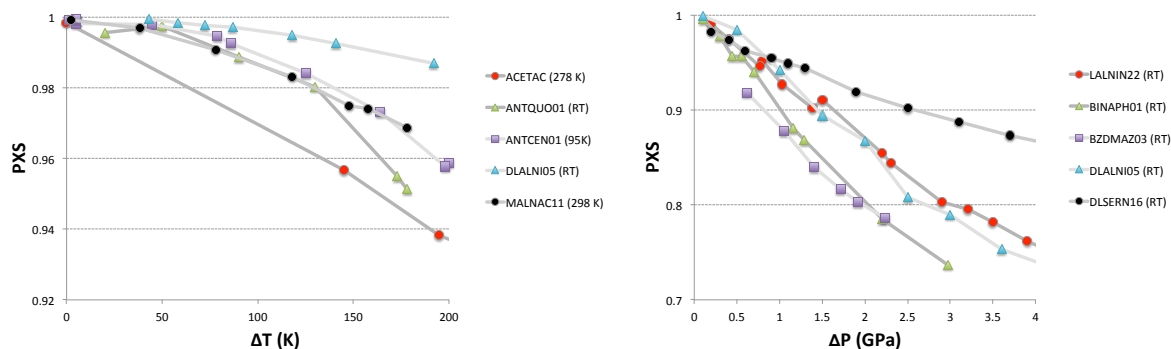
Figure 15 Impact of pressure and temperature differences to the PXS value when comparing structures determined at different temperatures (left, pressure is room pressure for all) and pressures (right, temperature is around RT for all and reference structure is always at room pressure).

used. In our CSD analysis this occurs in only 2.2% of cases. Those pairs of polymorphs usually have a high degree of 2D isostructurality and extremely similar PXRD patterns but do actually crystallise in different crystal structures often also with different space groups. We wonder whether these true polymorphs are routinely overlooked due to their very similar PXRD patterns.

The CSS is the most robust method for identifying whether two entries describe different structures or the redetermination of the same one. Indeed, the choice of CSS cut-off parameters is important. For example, the recommended cluster size for CSS comparisons is 20 since smaller sizes can fail when comparing structures with high Z' values. For large and flexible molecules, the usual tolerance thresholds used in the CSS algorithm (20% distance tolerance and 20° angle tolerance) does not always afford good matches for structures described in the literature as being redeterminations of the same polymorph. To avoid this issue in our CSD comparison, if a structure comparison had a PXS over 0.95 the CSS threshold parameters were increased to 50% distance tolerance and 50° angle tolerance. Our choice of cut-offs was done on the basis that our work required an analysis of a large number of structures. When looking at individual examples, however, one must carefully analyse the problem under consideration and the appropriate cut-offs adopted in a case-by-case basis. Our histograms of PXS, rmsd-r[20], rmsd-CL and rmsd-CA values for redetermined structures (Figure 1) should be useful references for these cut-off choices. The obvious limitation of the CSS method is that the method cannot be applied to disordered structures. Disorder needs to be removed before the method can be used. Such practice may hide real order-disorder phase transitions.

Cases of closely related polymorphs have attracted considerable attention in recent years since the discovery of form II aspirin.[28, 29] The PXS of forms I and II is 0.956 for the structures solved at 100 K (ACSALA12 and 13) and 0.962 for the structures solved at room temperature (ACSALA01 and 17), having 13 out of 20 molecules in common. These polymorphs have layers of identical structure but pack differently (through shifts in the packing or different symmetry). Various types of disorder have been observed in some of these polytypic pairs,[78] examples of these being aspirin,[30] or tazofelone.[79] Whether we should refer to these cases as crystals with static or dynamic disorder or polymorphs is not always clear. This question perhaps becomes even more important when comparing predictive landscapes with experimental structures.[80]

Related to this topic, crystal structures determined under significantly different temperature or pressure conditions may indeed be modulations of the same structure or they may actually correspond to different phases.[81] Some phase transitions occur with very minimal changes in crystal structure and these structure have been referred to as *isostructural polymorphs* (see FIZPEL FIZPEL01).[82, 83] Some other phase transitions only involve small conformational changes of flexible groups in the molecular periphery without significant crystal structure changes and these structures have been referred to as conformational phases.[74] Once again, whether these high/low temperature/pressure phases should be referred to as polymorphs is a matter of debate which extends beyond the scope of this paper.[81, 84, 85] What must be noted is that, whilst a structural analysis can help measuring the similarity and differences between the structures, it will fail in recognising whether those differences are linked to a phase change or not. In these instances, the use of other methods such as thermal analysis is a necessity.

Ultimately this work reveals that, a simple comparison (unit cell, PXRD, or crystal packing alone) alone may afford the wrong conclusions in the identification of crystal forms. It is thus our recommendation that at least a number of different methods are used for structures with high similarity.

7. Conclusions

The correct identification and recognition of different solid forms is critical in chemical crystallography. Traditionally this task required careful crystallographic analyses of the experimental crystals' structure, morphology and physical properties. In the information era, however, and with the growing availability of structures in the CSD, we rely more and more in the comparison of experimental data to data available in databases stored often as structural CIF files. To do such comparisons, we rely heavily in structural data and comparative methods based on structural data. Available methods to decide whether or not two structures are the same or are different are very often in agreement but in a considerable number of cases (~16%) they are not. In those borderline difficult cases, the CSS

method based on COMPACK seems to be the most robust, but it certainly also comes with limitations. In these difficult cases, the use of various structural comparison methods and the close look of an expert in the field remains necessary (and maybe non-sufficient). Ultimately, in those difficult cases, the identification of whether crystal forms are redeterminations, different phases or polymorphs is a complex problem of solid-state chemistry that requires a collection of techniques, including thermal analytic and computational ones, and their correct utilization by expert practitioners.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

Notes and references

1. K. S. Gushurst, J. Nyman and S. X. M. Boerrigter, *CrystEngComm*, 2019, **21**, 1363-1368.
2. K. Edkins, G. J. McIntyre, C. Wilkinson, V. Kahlenberg, D. Többens, U. J. Griesser, J. Brüning, M. U. Schmidt and J. W. Steed, *Cryst. Growth Des.*, 2019.
3. D. E. Braun, S. Lingireddy, M. Beidelschies, R. Guo, P. Müller, S. L. Price and S. M. Reutzel-Edens, *Cryst. Growth Des.*, 2017.
4. S. L. Childs, G. P. Stahly and A. Park, *Mol. Pharm.*, 2007, **4**, 323-338.
5. C. C. P. da Silva, R. d. O. Pepino, C. C. de Melo, J. C. Tenorio and J. Ellena, *Cryst. Growth Des.*, 2014, **14**, 4383-4393.
6. K. L. Cavanagh, C. Maheshwari and N. Rodríguez-Hornedo, *J. Pharm. Sci.*, 2018, **107**, 113-120.
7. O. N. Kavanagh, G. Walker and M. Lusi, *Cryst. Growth Des.*, 2019, **19**, 5308-5313.
8. P. M. Bhatt and G. R. Desiraju, *Chem. Commun.*, 2007, 2057-2059.
9. M. R. Chierotti, L. Ferrero, N. Garino, R. Gobetto, L. Pellegrino, D. Braga, F. Grepioni and L. Maini, *Chem. Eur. J.*, 2010, **16**, 4347-4358.
10. M. Lusi, *CrystEngComm*, 2018, **20**, 7042-7052.
11. M. Lusi, *Cryst. Growth Des.*, 2018, **18**, 3704-3712.
12. Q. Jiang, C. Hu and M. D. Ward, *J. Am. Chem. Soc.*, 2013, **135**, 2144-2147.
13. M. A. Malwitz, S. H. Lim, R. L. White-Morris, D. M. Pham, M. M. Olmstead and A. L. Balch, *J. Am. Chem. Soc.*, 2012, **134**, 10885-10893.
14. F. P. A. Fabbiani and C. R. Pulham, *Chem. Soc. Rev.*, 2006, **35**, 932-942.
15. E. C. van Tonder, M. R. Caira, S. A. Botha and A. P. Lötter, *Int. J. Pharm.*, 1990, **63**, 35-42.
16. B. Rodríguez-Spong, C. P. Price, A. Jayasankar, A. J. Matzger and N. r. Rodríguez-Hornedo, *Adv. Drug Deliver. Rev.*, 2004, **56**, 241-274.
17. P. Vishweshwar, J. A. McMahon, J. A. Bis and M. J. Zaworotko, *J. Pharm. Sci.*, 2006, **95**, 499-516.
18. J. Bernstein, *Polymorphism in Molecular Crystals*, First edn., Oxford, 2007.
19. R. M. Bhardwaj, J. A. McMahon, J. Nyman, L. S. Price, S. Konar, I. D. H. Oswald, C. R. Pulham, S. L. Price and S. M. Reutzel-Edens, *J. Am. Chem. Soc.*, 2019.
20. M. K. Corpinot and D.-K. Bučar, *Cryst. Growth Des.*, 2019, **19**, 1426-1453.
21. W. I. F. David, K. Shankland, J. van de Streek, E. Pidcock, W. D. S. Motherwell and J. C. Cole, *J. Appl. Crystallogr.*, 2006, **39**, 910-915.
22. A. Altomare, C. Cuocci, C. Giacovazzo, A. Moliterni, R. Rizzi, N. Corriero and A. Falcicchio, *J. Appl. Crystallogr.*, 2013, **46**, 1231-1235.
23. F. Allen, *Acta Crystallogr. Sect. B*, 2002, **58**, 380-388.
24. C. R. Groom, I. J. Bruno, M. P. Lightfoot and S. C. Ward, *Acta Crystallogr. Sect. B*, 2016, **72**, 171-179.
25. R. Taylor and P. A. Wood, *Chem. Rev.*, 2019.
26. P. Müller, *Crystallography Reviews*, 2009, **15**, 57-83.
27. H. Rietveld, *J. Appl. Crystallogr.*, 1969, **2**, 65-71.
28. P. Vishweshwar, J. A. McMahon, M. Oliveira, M. L. Peterson and M. J. Zaworotko, *J. Am. Chem. Soc.*, 2005, **127**, 16802-16803.
29. A. D. Bond, R. Boese and G. R. Desiraju, *Angew. Chem. Int. Ed. Engl.*, 2007, **46**, 615-617.
30. A. D. Bond, R. Boese and G. R. Desiraju, *Angew. Chem. Int. Ed. Engl.*, 2007, **46**, 618-622.
31. M. F. Pina, J. F. Pinto, J. J. Sousa, L. Fábián, M. Zhao and D. Q. M. Craig, *Mol. Pharm.*, 2012, **9**, 3515-3525.
32. R. de Gelder, R. Wehrens and J. A. Hageman, *J. Comput. Chem.*, 2001, **22**, 273-289.
33. M. A. Spackman and D. Jayatilaka, *CrystEngComm*, 2009, **11**, 19-32.
34. C. F. Macrae, I. J. Bruno, J. A. Chisholm, P. R. Edgington, P. McCabe, E. Pidcock, L. Rodriguez-Monge, R. Taylor, J. van de Streek and P. A. Wood, *J. Appl. Crystallogr.*, 2008, **41**, 466-470.
35. C. Hu, Q. Li and U. Englert, *CrystEngComm*, 2003, **5**, 519-529.
36. C. J. Adams, M. Lusi, E. M. Mutambi and A. G. Orpen, *Cryst. Growth Des.*, 2017, **17**, 3151-3155.
37. C. J. Gilmore, G. Barr and J. Paisley, *J. Appl. Crystallogr.*, 2004, **37**, 231-242.
38. S. Habermehl, P. Mörschel, P. Eisenbrandt, S. M. Hammer and M. U. Schmidt, *Acta Crystallogr. Sect. B*, 2014, **70**, 347-359.
39. J. A. Chisholm and S. Motherwell, *J. Appl. Crystallogr.*, 2005, **38**, 228-231.
40. T. Gelbrich and M. B. Hursthouse, *CrystEngComm*, 2006, **8**, 448-460.
41. J. J. McKinnon, F. P. A. Fabbiani and M. A. Spackman, *Cryst. Growth Des.*, 2007, **7**, 755-769.
42. D. J. Carter, P. Raiteri, K. R. Barnard, R. Gielink, M. Mocerino, B. W. Skelton, J. G. Vaughan, M. I. Ogden and A. L. Rohl, *CrystEngComm*, 2017, **19**, 2207-2215.
43. J. van de Streek and S. Motherwell, *Acta Crystallogr. Sect. B*, 2005, **61**, 504-510.

44. J. J. M. M. J. Turner, S. K. Wolff, D. J. Grimwood, P. R. Spackman, D. Jayatilaka and M. A. Spackman, , *CrystalExplorer17*, (2017) University of Western Australia.

45. M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, Williams, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman and D. J. Fox, *Gaussian 16 Rev. C.01*, (2016), Wallingford, CT.

46. S. Grimme, *J. Comput. Chem.*, 2006, **27**, 1787-1799.

47. R. Marsh, *Acta Crystallogr. Sect. B*, 2004, **60**, 252-253.

48. R. M. Hazen, *Am. Mineral.*, 1976, **61**, 1280-1293.

49. S. Athimoolam and S. Natarajan, *Acta Crystallogr. Sect. C*, 2007, **63**, o514-o517.

50. R. Benali-Cherif, R. Takouachet, E.-E. Bendeif and N. Benali-Cherif, *Acta Crystallogr. Sect. C*, 2014, **70**, 323-325.

51. O. A. Lodochnikova, R. M. Khakimov, L. Z. Latypova, A. R. Kurbangalieva and I. A. Litvinov, *Russ. Chem. Bull.*, 2015, **64**, 2444-2453.

52. T. N. Drebushchak, V. A. Drebushchak and E. V. Boldyreva, *Acta Crystallogr. Sect. B*, 2011, **67**, 163-176.

53. M. Yasuda, G. Kuwamura, T. Nakazono, K. Shima, Y. Inoue, N. Yamasaki and A. Tai, *Bull. Chem. Soc. Jpn.*, 1994, **67**, 5.

54. R. E. Marsh, M. Kapon, S. Hu and F. H. Herbstein, *Acta Crystallogr. Sect. B*, 2002, **58**, 62-77.

55. X.-L. Li, Y.-L. Gao, X.-L. Feng, Y.-X. Zheng, C.-L. Chen, J.-L. Zuo and S.-M. Fang, *Dalton Trans.*, 2012, **41**, 11829-11835.

56. D. G. Anderson, A. J. Blake, R. Blom, S. Cradock and D. W. H. Rankin, *Acta Chem. Scand.*, 1991, **45**, 6.

57. S. A. Barnett, C. K. Broder, K. Shankland, W. I. F. David, R. M. Ibberson and D. A. Tocher, *Acta Crystallogr. Sect. B*, 2006, **62**, 287-295.

58. C. M. Weeks, W. L. Duax, R. A. Finnegan, D. J. Delecki and B. Kojic-Prodic, *Acta Crystallogr. Sect. C*, 1984, **40**, 1376-1378.

59. H. Irngartinger, A. Altreuther, T. Sommerfeld and T. Stojanik, *Eur. J. Org. Chem.*, 2000, **2000**, 4059-4070.

60. R. E. Jones and D. H. Templeton, *Acta Crystallogr.*, 1958, **11**, 484-487.

61. R. Boese, D. Blaser, R. Latz and A. Baumen, *Acta Crystallogr. Sect. C*, 1999, **55**.

62. A. Basheer, H. Yamataka, S. C. Ammal and Z. Rappoport, *The Journal of Organic Chemistry*, 2007, **72**, 5297-5312.

63. T. Steiner, *J. Chem. Soc., Chem. Commun.*, 1995, 1331-1332.

64. M. Lusi and L. J. Barbour, *Cryst. Growth Des.*, 2011, **11**, 5515-5521.

65. M. Woińska, S. Grabowsky, P. M. Dominiak, K. Woźniak and D. Jayatilaka, *Science Advances*, 2016, **2**.

66. R. Boese, H.-C. Weiss and D. Bläser, *Angew. Chem. Int. Ed. Engl.*, 1999, **38**, 988-992.

67. M. Podsiadło, A. Olejniczak and A. Katrusiak, *The Journal of Physical Chemistry C*, 2013, **117**, 4759-4763.

68. A. Katrusiak, M. Podsiadło and A. Budzianowski, *Cryst. Growth Des.*, 2010, **10**, 3461-3465.

69. A. D. Chanyshev, K. D. Litasov, S. V. Rashchenko, A. Sano-Furukawa, H. Kagi, T. Hattori, A. F. Shatskiy, A. M. Dymshits, I. S. Sharygin and Y. Higo, *Cryst. Growth Des.*, 2018, **18**, 3016-3026.

70. L. Beer, J. L. Brusso, A. W. Cordes, R. C. Haddon, M. E. Itkis, K. Kirschbaum, D. S. MacGregor, R. T. Oakley, A. A. Pinkerton and R. W. Reed, *J. Am. Chem. Soc.*, 2002, **124**, 9498-9509.

71. C. M. Robertson, A. A. Leitch, K. Cvrkalj, R. W. Reed, D. J. T. Myles, P. A. Dube and R. T. Oakley, *J. Am. Chem. Soc.*, 2008, **130**, 8414-8425.

72. A. Alberola, R. J. Collis, S. M. Humphrey, R. J. Less and J. M. Rawson, *Inorg. Chem.*, 2006, **45**, 1903-1905.

73. R. G. Simões, C. E. S. Bernardes, A. Joseph, M. F. M. Piedade, W. Kraus, F. Emmerling, H. P. Diogo and M. E. Minas da Piedade, *Mol. Pharm.*, 2018, **15**, 5349-5360.

74. F. J. Zuniga, A. J. Cruz-Cabeza, X. M. Aretxabaleta, N. de la Pinta, T. Breczewski, M. M. Quesada-Moreno, J. R. Aviles-Moreno, J. J. Lopez-Gonzalez, R. M. Claramunt and J. Elguero, *IUCrJ*, 2018, **5**, 706-715.

75. A. J. Cruz-Cabeza, R. J. Davey, I. D. H. Oswald, M. R. Ward and I. J. Sugden, *CrystEngComm*, 2019, **21**, 2034-2042.

76. P. K. Tarafdar, S. T. Reddy and M. J. Swamy, *Cryst. Growth Des.*, 2012, **12**, 1132-1140.

77. L. Fabian and A. Kalman, *Acta Crystallogr. Sect. B*, 2004, **60**, 547-558.

78. A. D. Bond, *CrystEngComm*, 2012, **14**, 2363-2366.

79. L. S. Price, J. A. McMahon, S. R. Lingireddy, S.-F. Lau, B. A. Diseroad, S. L. Price and S. M. Reutzel-Edens, *J. Mol. Struct.*, 2014, **1078**, 26-42.

80. A. J. Cruz-Cabeza, G. M. Day and W. Jones, *Phys. Chem. Chem. Phys.*, 2011, **13**, 12808-12816.

81. A. Gavezzotti, *J. Pharm. Sci.*, 2007, **96**, 2232-2241.

82. S. J. Coles, T. L. Threlfall and G. J. Tizzard, *Cryst. Growth Des.*, 2014, **14**, 1623-1628.

83. K. K. Jha, S. Dutta, V. Kumar and P. Munshi, *CrystEngComm*, 2016, **18**, 8497-8505.

84. G. R. Desiraju, *Cryst. Growth Des.*, 2008, **8**, 3-5.

85. J. Bernstein, J. D. Dunitz and A. Gavezzotti, *Cryst. Growth Des.*, 2008, **8**, 2011-2018.