

<https://doi.org/10.1038/s42004-024-01380-3>

The interplay between hydrogen bonds and stacking/T-type interactions in molecular cocrystals

Aurora J. Cruz-Cabeza ¹ ✉, Peter R. Spackman ² & Amy V. Hall ¹ ✉

Supramolecular synthon and hydrogen bond pairing approaches have influenced the understanding of cocrystal formation for decades, but are hydrogen bonds really the dominant interaction in cocrystals? To investigate this, an extensive analysis of 1:1 two-component cocrystals in the Cambridge Structural Database was undertaken, revealing that stacking and T-type interactions are just as, if not more important than hydrogen bonds in molecular cocrystals. A total of 84% of the most common coformers in the dataset are aromatic. When analysing cocrystal dimers, only 20% consist of solely strong hydrogen bonds, with over 50% of contacts involving stacking and T-type interactions. Combining interaction strength and frequency, both hydrogen bond and stacking/T-type interactions contribute equally to the stabilisation of cocrystal lattices. Therefore, we state that crystal engineering and cocrystal design concepts of the future should not solely revolve around supramolecular synthon pairing via hydrogen bonds, but instead consider optimising both hydrogen bonding and stacking/T-type interactions.

Cocrystals, generally defined as crystalline solids composed of two or more different neutral molecules in a stoichiometric ratio held together by non-covalent interactions, have applications spanning a wide range of industries, from pharmaceuticals^{1–3} to explosives^{4–6} to optoelectronics^{7–9} to dosimetry¹⁰. The practice of designing and synthesising molecular compounds is centuries old and so is their isolation in the crystalline state^{11,12}. The first ever cocrystal was reported by Wöhler in 1844^{13,14}. He produced green crystals containing quinone and hydroquinone from 1:1 solutions which he described as “one of the most beautiful substances known to organic chemistry” (Fig. 1a)^{13,14}. Wöhler’s discovery initiated the interest in molecular cocrystals, leading to the report of several cocrystals of phenazone in 1895¹⁵ and over 300 cocrystals of aromatic compounds in 1922¹⁶. A year after this, in 1923, the first X-ray structure of an organic molecule was reported which revolutionised what was possible for the characterisation of solid-state materials¹⁷. From the mid-1920s until the 1950s, Kofler, Kofler, and Kuhnert-Brandstatter synthesised hundreds of multicomponent crystals by contact method thermomicroscopy¹⁴, with nicotinamide used as a common coformer¹⁸. In 1965, an important repository to compile the structural data of organic small molecules was born in the form of The Cambridge Structural Database (CSD)¹⁹. For each entry in the CSD, all crystallographic information is reported, as well as the associated publication. The number of structures deposited in the CSD has grown exponentially since its creation and it now constitutes the single most important

source of small molecule crystal data in the world. This is indeed also the case for the number of reported cocrystals in the CSD, shown in Fig. 1b from 1965 to 2020. Whilst cocrystals have been known for centuries, it is perhaps only in the last 20 years or so that their potential to impact materials’ applications has been exploited, with drug delivery as an important example^{20,21}. The first cocrystal deposited in the CSD (the year of its creation, 1965) consists of isocytosine tautomers which resemble the base pair interactions in DNA, as shown in Fig. 1c^{22,23}.

Common to the quinone:hydroquinone and isocyanate tautomer cocrystals, is the fact that the two distinct molecules forming the cocrystal are held together via strong hydrogen bonds as well as by the strong stacking interactions of the planar molecules. The term ‘intermolecular interactions’ first appeared in the literature in 1954²⁴, however, ‘non-covalent interactions’ was only mentioned for the first time in 1974²⁵. The main types of non-covalent bonding found in molecular cocrystals are hydrogen bonds (HBs), halogen bonds (HaBs), weak hydrogen bonds (wHBs), and stacking/T-type interactions (STs) involving both aromatic and aliphatic compounds. Below we discuss the use of each of these types of non-covalent interactions in cocrystal design.

The term ‘crystal engineering’ was first coined by Pepinsky in 1955 in a meeting abstract²⁶, however, it was initially used in a full scientific paper by Schmidt in 1971²⁷. The practice of specifically designing cocrystals based on their potential to form HBs has been commonplace since the popularisation

¹Department of Chemistry, Durham University, Durham, DH1 3LE, UK. ²School of Molecular and Life Sciences, Curtin University, Perth, WA, 6845, Australia.

✉ e-mail: aurora.j.cruz-cabeza@durham.ac.uk; amy.v.hall@durham.ac.uk

Fig. 1 | Cocrystals and the CSD. **a** The structure of the first reported cocrystal: quinone and hydroquinone to form quinhydrone (QUIDON)^{13,66}. **b** The exponential growth of cocrystal entries in the CSD from 1965 to 2020, plotted from the 8408 dataset of two-component cocrystals (no water, solvents, or zwitterions)⁴⁹. **c** The first cocrystal structure deposited in the CSD (ICYTIN)²², consisting of two isocyanate tautomers displaying extensive hydrogen bonding and aromatic stacking.

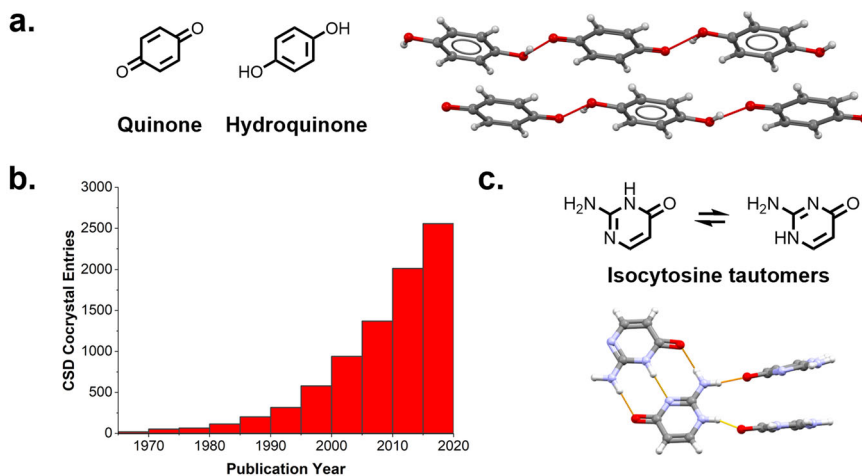
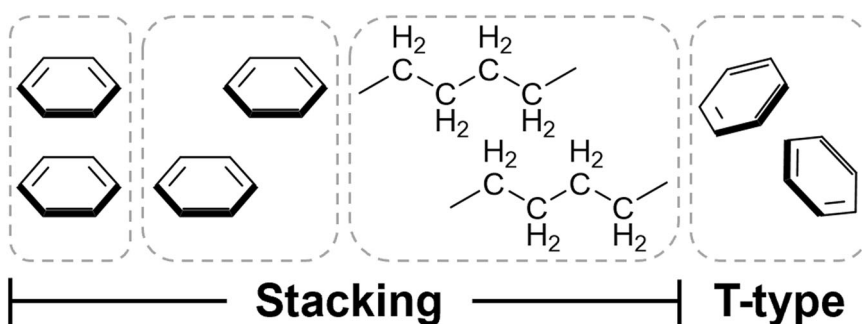


Fig. 2 | Examples of ST interactions. Left to right: eclipsed face-to-face aromatic stacking, offset face-to-face aromatic stacking, aliphatic alkane stacking, followed by T-type (also known as edge-to-face).



of crystal engineering in the late 1980s and mid-1990s^{28,29}. Over the last 30 years, the terms ‘cocrystal’ and ‘hydrogen bond’ appear simultaneously, seemingly due to the strength, directionality, and structural stability that HBs provide, alongside their predictability when considering HB supramolecular synthons and functional group pairings^{29,30}. Supramolecular synthons can be either homo (e.g. acid...acid) or hetero (e.g. acid...amide)³¹, with the acid...pyridine and acid...amide heterosynthons being examples of the most robust, illustrated by 1:1 and 2:1 benzoic acid-isonicotinamide cocrystals (CSD refcodes BUDWEC and MOVTOH)^{32,33}. Strong HBs and the accumulation of softer wHBs are exceedingly important in cocrystals³⁴, however, it is worth considering that they may not be the only significant type of intermolecular interaction that can influence cocrystal design, formation, and stability.

HaBs are specific and directional interactions that demonstrate considerable functionality in organic materials and crystal engineering^{35,36}. Halogen bonded homo and heterosynthons are common²⁹, however, HaBs have dwindling popularity when compared to HBs, partly as HaBs are restricted by the specific halogen atoms (F, Cl, Br, I) that must be present on the cofomer(s). In some cases, HaB cocrystals have shown to be more stable than HB cocrystals³⁴, with the example of a 1:1 cocrystal of tetraiodo-1,4-benzoquinone and tetrathiafulvalene, showing enhanced stability because of the I...S HaBs and charge transfer between the stacked components (HUJNUX)³⁷. Additionally, when halobenzene molecules stack, their influence is as great as hydrogen bonding^{38–40}.

The stacking of compounds can either be aromatic or aliphatic, depending on the nature of the molecules. Aromatic stacking allows aromatic rings to be closer in proximity to optimise intermolecular dispersion⁴¹, with the most energetically favourable configurations of rings being offset face-to-face and edge-to-face⁴², whereas eclipsed face-to-face stacking is unfavourable due to strong repulsion of the overlapping π electron clouds (Fig. 2)⁴³. Aromatic interactions provide the basis of organic optoelectronics in which new semiconductors are synthesised by manipulating the aromatic

stacking strengths through pairing of aromatic rings with complementarity, thus resulting in closer stacking distances and improved electrical properties⁴⁴. Cocrystals assembled by aromatic stacking interactions often exhibit intermolecular charge transfer because of the close packing of the π electron donors and acceptors, with perylene and tetracyanoquinodimethane 1:1 and 1:3 cocrystals as an example (PERTCQ01 and TCQPER01)⁴⁵. Furthermore, the role of ST interactions is traditionally ignored in crystal engineering when compared to HBs, which was recently exemplified by Friščić et al. where they concluded that π -systems should be considered as targets for directing cocrystal formation⁴⁶.

As the interest in cocrystal design has increased, crystal engineers have attempted to develop rules to predict and direct cocrystal formation. These rules have undoubtedly been dominated by HBs, with the design of heterosynthons based on them^{29,47}. Beyond HBs, HaBs and wHBs have been key players in cocrystal design, however, rules based on controlling dispersion-based interactions (including aromatic and aliphatic stacking and T-interactions) have been consistently underdeveloped, although there are exceptions⁴⁸. In this context, the aim of this study is simple. We seek to quantify the strength of intermolecular interactions in molecular cocrystals and compute their overall contributions to the cocrystal lattice energies. With this, we attempt to answer the question, ‘Are hydrogen bonds the most important interaction in molecular cocrystals?’ by quantitatively analysing the contributions of all interactions to the cocrystal lattice energies.

Results and discussion

The cocrystal dataset: overview

We have produced a subset of two-component cocrystals (AB) from the CSD starting from a cocrystal dataset with 9,464 structures – identified in a previous study⁴⁹. The previous study analysed all multicomponent crystals in the CSD and classified them according to the number of components, their protonation state, and their hydration/solvation. Hence, the original

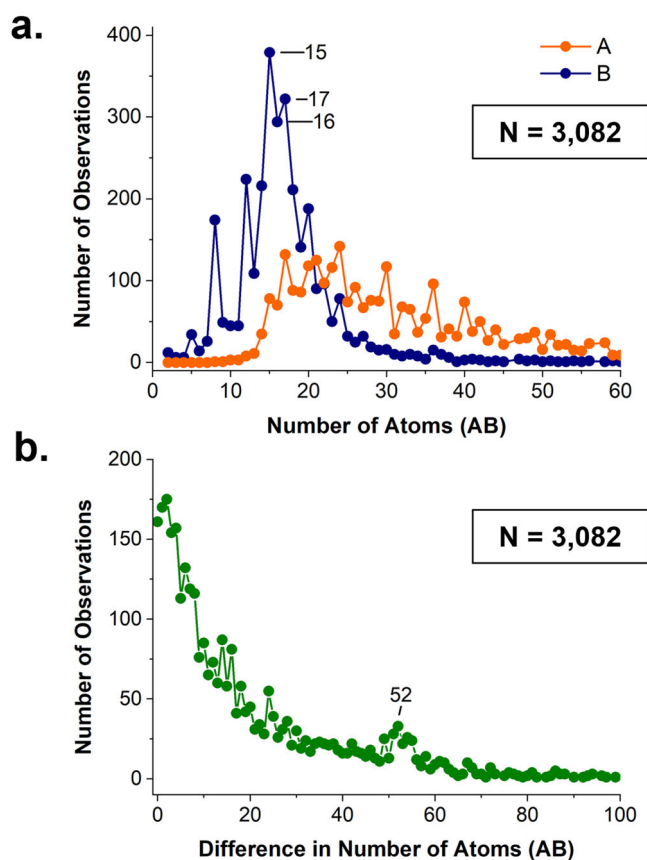


Fig. 3 | Cococrystal component size. A is the larger component and B is the smaller component in the AB cococrystal. **a** Distributions of sizes for the A and B components in the cococrystals and **b** the distribution of the difference between A and B components.

dataset only contains two component cococrystals from the CSD with two neutral (and non-zwitterionic) molecular components A and B, where A is larger than B. Structures containing the most common solvents were removed, so the dataset contains mostly cococrystals though less common solvent molecules may also be present. Starting from such a dataset, we took a subset of 3082 structures of cococrystals having 1:1 stoichiometry and $Z' = 1$ only, no compounds sitting in special positions, no halogen atoms larger than bromine, and no disorder. This filtering was done to facilitate the computations of interaction energies and their analysis.

The cococrystal dataset: size distribution

Here we analyse our dataset of 3082 cococrystals according to the size of its components, A and B. When referring to components within each cococrystal, A is always the larger molecule and B is the smaller molecule. The distribution of sizes of both AB components in our dataset is shown in Fig. 3a. We immediately note that the distribution of sizes for the smaller component B is significantly narrower than that of the larger component A, with peak maxima for B placed at 15, 16, and 17 atoms (Fig. 3a). This may be due to biases of crystal engineers who may often use a limited set of coformers based on structure, availability, and costs for designing cococrystals. The size of both components is smaller than most organic compounds in the CSD. The overall CSD database has a component size distribution peak maximum at 40 atoms and the subset of polymorphs at around 30 atoms⁵⁰; in the cococrystals, we find a peak at about 20 atoms for component B and around 20–30 for component A. This may reflect, again, trends in crystal engineering where cococrystal design has more often been attempted by academic groups working on smaller systems.

The distribution of the difference in atoms between components A and B clearly shows the decay in observations as such difference increases

(Fig. 3b). This suggests that cococrystals containing compounds of similar sizes are more common – consistent with the Fabian observation that compounds with complementarity, also in size, are more likely to cococrystallise⁵¹. There is a small band of observations at around an atom difference of 52 atoms which is related to inclusion structures containing a large host and a small guest.

The cococrystal dataset: composition

Analysis of the 3082 cococrystal dataset revealed 32 persistent coformers found in at least twenty or more different cococrystal structures (Fig. 4). 84% of the 32 coformers are aromatic. Isonicotinamide is the most common coformer within the dataset as it is found in 77 cococrystals, with its constitutional isomer, nicotinamide, found in 52. Hydroxybenzoic acid constitutional isomers are also found as top coformers with the para isomer having the greatest number of observations (43), followed by the ortho (41), and meta (24) isomers. Only 16% of the 32 coformers were aliphatic. Interestingly, the two most common aliphatic coformers in the dataset are two steroids, namely cholic acid and cholamid, which are well known to form inclusion compounds. Together, these two steroids are found in a total of 92 different cococrystal structures. Following the steroids, dicarboxylic acids are also in this set and include glutaric acid (37 observations), malonic acid (21), and 2-butenedioic acid (20). Analysis by acid/base nature of these top 32 coformers shows that there is an equal number of acids and bases (16 of each). The distribution of pK_a values (predicted by ChemAxon)⁵² for the acids lies between 1.4 and 18.3 and for the protonated bases is between –11.9 and 6.2, therefore, highlighting the coformers consisting of a range of mild to weak acids and bases.

The cococrystal dataset: lattice energies

Lattice energies for the 3082 cococrystal dataset were calculated using the Open Computational Chemistry (OCC) software which offers a command-line implementation of the interaction energy models available in the popular modelling software CrystalExplorer⁵³. The lattice energies are computed via summation of dimer interactions until convergence. The computed distribution of cococrystal lattice energies is shown in Fig. 5a with its fraction contributions by homo *versus* heterodimers split in Fig. 5b. There is a maximum in the distribution of cococrystal lattice energies at approximately –250 kJ/mol. Our distribution reflects of course the lattice energies of cococrystals of small molecules, with a relationship between size and lattice energy – similar to trends found in single-component systems⁵⁴. We note that the lattice energies of the cococrystals are given per pair of AB molecules, so this value can be compared to the sum of lattice energies of pure A and B. Lattice energies of pure compounds (similar to these) have been computed to lie around –100 to –110 kJ/mol, so our cococrystal lattice energies are slightly more negative than double that, as expected for cococrystal formation driven by an energy gain⁵⁴.

The partitioning of lattice energies into fractions of homo and heterodimer contributions shows a peak maximum for hetero contributions at around 60% and a peak maximum for homo contributions at around 40%. This suggests that, typically, the heterodimer interactions have an overall greater contribution to the lattice energy than the homodimer interactions (Fig. 5b).

To illustrate how different cococrystals can have markedly different hetero *vs.* homo contributions to the lattice energies, we have selected three illustrative examples from the dataset and shown them in Fig. 6. In the example AJUXUY, the hetero interactions between A and B are very strong while the homo (AA and BB) interactions are overall weak. A and B pack in a highly alternated manner, maximising the AB interactions and their fraction contribution to the lattice energy. In the example UPOQIA, there is generally a balance in the strength of AB, AA, and BB interactions. Here, the AB dimer has a strong contribution of HB energy whilst the AA and BB interactions are based on the stacking of AA and BB through translation symmetry. Finally, in GUNLAC, the homo interactions between the main and larger component A dominate significantly whilst B fills some spaces in the structure. This situation where the homo interactions are much more

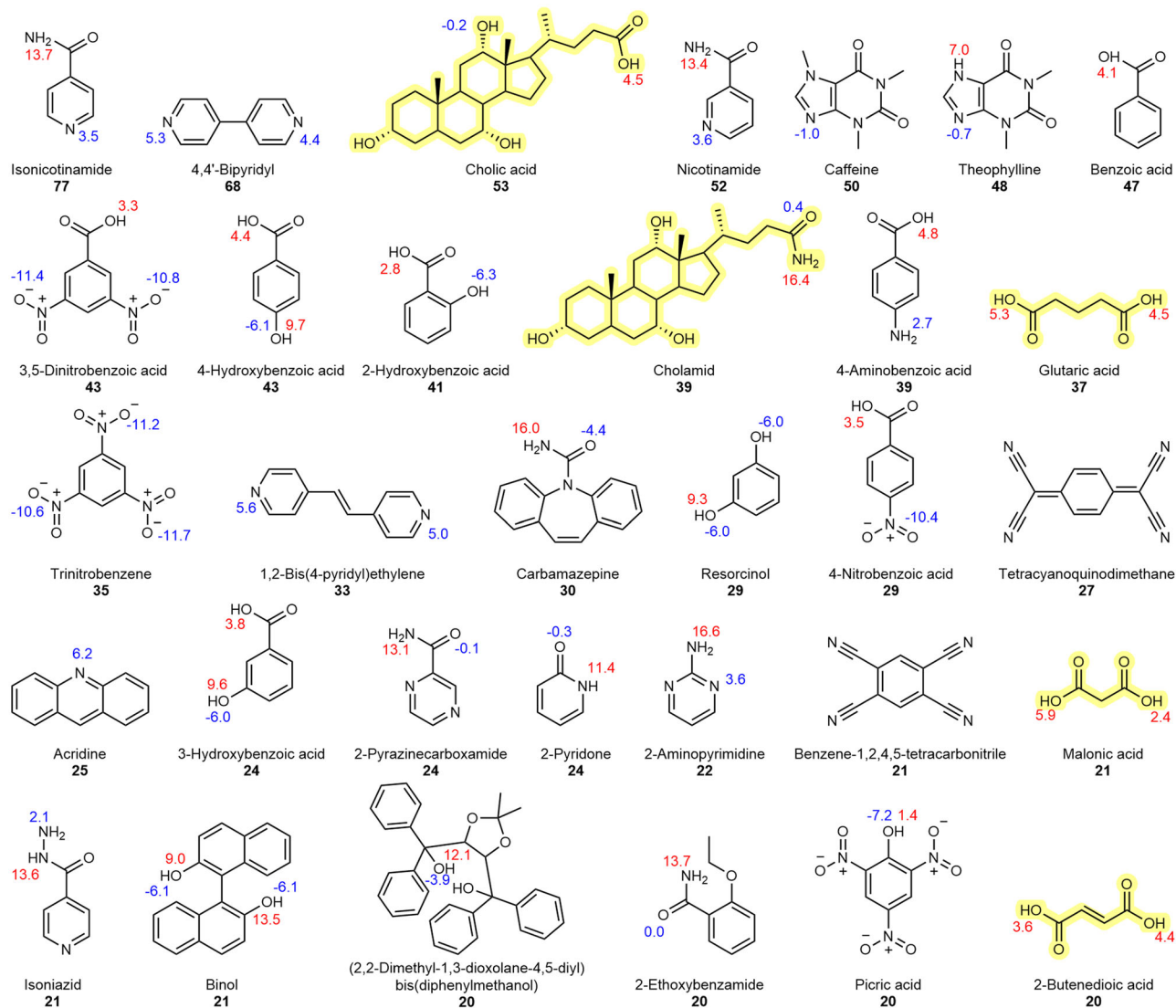


Fig. 4 | Cocrystal composition. Molecular structures of the top 32 cocrformers found in the cocrystal dataset in decreasing order of observations (value in bold) along with their common name and pK_a values (acid values in red, protonated base values in blue). The five aliphatic cocrformers are highlighted in yellow.

dominant than the hetero interactions match with inclusion-like structures. Structures of such type tend to have a markedly difference in size between the A and B components with the larger component (and thus the homo interaction) dominating the stabilisation of the lattice and the smaller component playing a minor role.

Dimers: hierarchy, classification, and examples

Almost 1.5 million unique dimer interactions were computed within the 3082 cocrystal dataset with OCC. Of those, 69,089 dimers were found to have interaction energies of less than -1 kJ/mol. In this section, we analyse the structure and interaction energies of the 69,089 dimers. On average, ~22 relevant dimers are analysed per cocrystal structure. Molecule-molecule dimers contain a combination of contacts and thus a 'pure' classification of such dimers is not possible – a hierarchy must be established for a meaningful classification. The following hierarchy of interactions was adopted to classify each of the dimers: HB > HaB > ST > wHB. The hierarchy simply assigns priorities for the classification. For example, a dimer containing wHB and HB would be classified as HB since HB is higher in the hierarchy. For a dimer to be classified as ST, at least two ST contacts must exist in the dimer whereas only one contact of the type was required for the other classifications. With this hierarchy, we were able to classify ~40% of the unique 69,089 dimers into these four types (Table 1). The adopted hierarchy

correlates well with the average interaction energies for each of the types (Table 1) except for HaB. We note that our dataset excludes the iodine atom from the calculations (no 6-31 G** basis set was available for iodine) causing the average energy of HaB classified dimers to be lower than that expected for HaBs containing iodine. Furthermore, there is only a small number of structures and dimers of molecules containing halogens with close contacts. The contacts with no classification are on average soft interactions of approximately -5 kJ/mol and in most of these interactions, the molecules in the dimers are significantly separated and the repulsion contributions are 0. As clearly seen in Table 1, this classification shows that ST interactions are the most abundant, followed by wHBs and HBs, with HaB being a very small fraction. Further, when symmetry and all dimers (not just the symmetry unique) are taken into consideration, the fractions remain similar. This is simply because the symmetry relations affect all types of dimers equally. To illustrate the effectiveness of our classification rules, we show four classified dimers within the crystal structures of four cocrystals in Fig. 7 together with their interaction energies. Hetero-HB interactions are identified in three cocrystals (ACOYUM, ADETOT, ISIJAW) and are, in all cases, the strongest interactions with stabilisation energies between -48 and -62 kJ/mol. The two R₂²(8) motifs are more stabilising than the two D₁¹(2) motifs. ST interactions are identified in six dimers (both hetero and homo) with stabilisation energies between -14 and -43 kJ/mol. We note that ST

interactions are defined broadly and may involve the stacking or T interaction between molecules of any type. The homo-ST in ACOYUM is a classic aromatic stack example with inversion symmetry. The stacking of azelaic acid in the ISIJAW cocrystal is an example of a homo-ST interaction of an aliphatic compound. wHBs are found in three cocrystals (ACOYUM,

ADETOT, and HOZNIV) with stabilisation energies ranging between -9 and -28 kJ/mol. Finally, HaB of Br \cdots O and Br \cdots Br are identified in the HOZNIV cocrystal with stabilisation energies ranging from -9 to -14 kJ/mol.

Dimers: analysis

In this section, the data on dimers are analysed and meaningful distributions are presented. The distribution of interaction energies by dimer types is shown in Fig. 8a with examples of HB and ST interactions of different strengths in Fig. 8b. Since the HaB interactions are so scarce, they will not be discussed.

We first note the shape of the energy distribution by interaction types (Fig. 8a). The distribution for HBs is broad, for STs narrower and for wHBs the narrowest. A reason for this could be our classification rules. Since the hierarchy of these interactions is HBs > STs > wHBs, dimers classified with interaction types higher in the hierarchy may consist of a blend of interactions. As the classification goes down the hierarchy, the types certainly become more specific – the dimers are less composed of a blend of interactions and are ‘purer’ in nature. This is seen in the example UKUSOJ in Fig. 8b. Whilst the highlighted UKUSOJ dimer is classified as HB, it contains very important ST contributions.

We secondly note the strength of dimer interactions by types (Fig. 8a). The peak maxima for the distributions of the different interaction types (the most common interactions in the dataset) are approximately -50 kJ/mol for HBs, -20 kJ/mol for STs, and -10 kJ for wHBs (Fig. 8a). To shed some light on the interaction types and strengths, examples of HB and ST interactions of high (-100 kJ/mol), medium (-50 kJ/mol) and low (-20 kJ/mol) strengths are shown in Fig. 8b. For the HBs, low-strength HBs are weaker with a suboptimal geometry (UHAFEP), medium-strength HBs are strong with an optimal geometry (YIPQIA) and high-strength HBs are interactions containing strong HBs but also contributions from ST interactions (UKUSOJ). For the STs, we observe a very clear correlation between the ST energy and the size of the molecule. Therefore, the low, medium, and high-strength STs contain, for example, stacking dimers of one, two, and three ring-containing compounds (Fig. 8b). Further to this, aromatic rings with electron-withdrawing groups are found to have strong ST interactions with aromatic rings with electron-donating groups⁵⁵.

We thirdly note the volume of the interaction types (Figs. 8a and 9a). Whilst STs and wHBs are on average weaker than HB interactions, they are significantly more frequent. This can be seen in Fig. 8a but also is shown more clearly in a pie chart in Fig. 9a. ST interactions account for 53% of all classified unique interactions found in these cocrystals, thus significantly dominating the cocrystals by volume. This is followed by wHBs (26%) and finally HBs (at only 20%). Similar statistics are derived if symmetry is accounted for in the counting of the interactions, which is not shown here for simplicity – our analysis, as described before, only counts symmetry-independent dimers. These hetero-ST motifs are commonly found in cocrystals designed for electronic, optoelectronic, and magnetic applications^{56,57}.

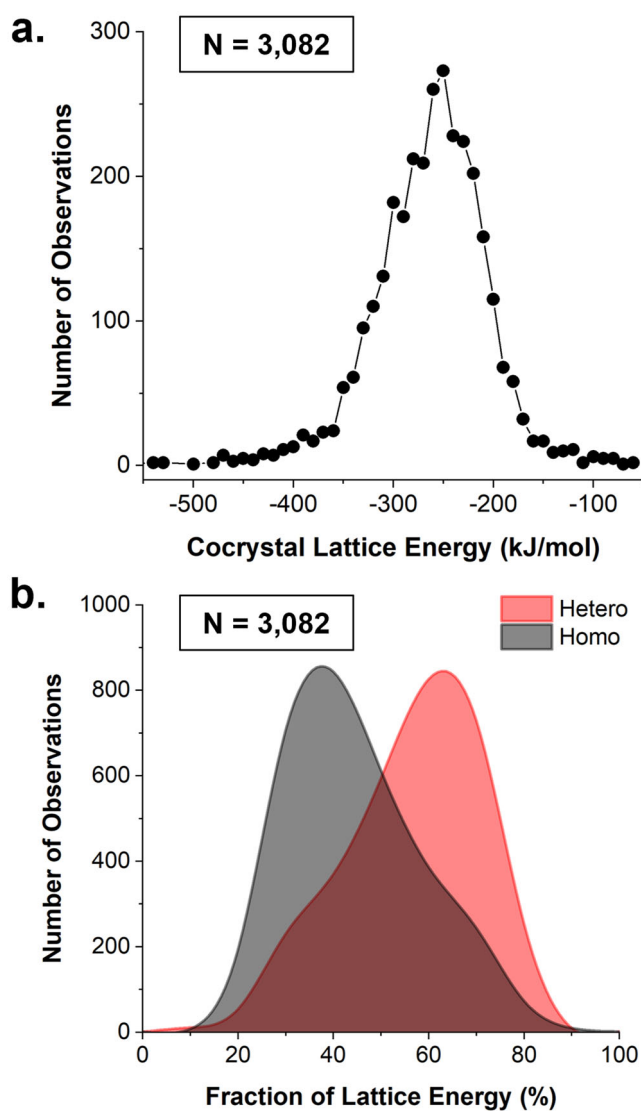


Fig. 5 | Cocystal lattice energies. a Distribution of cocystal lattice energies and b distribution of hetero- vs. homo- fraction of the cocystal lattice energies.

Fig. 6 | Cocystal lattice energy examples. Examples of cocystals with different homo (AA and BB) and hetero (AB) contributions to the lattice energies. Hydrogen atoms are omitted for clarity.

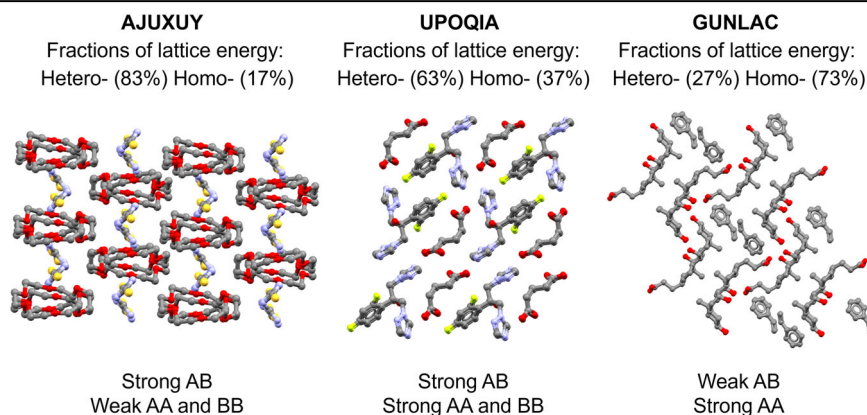
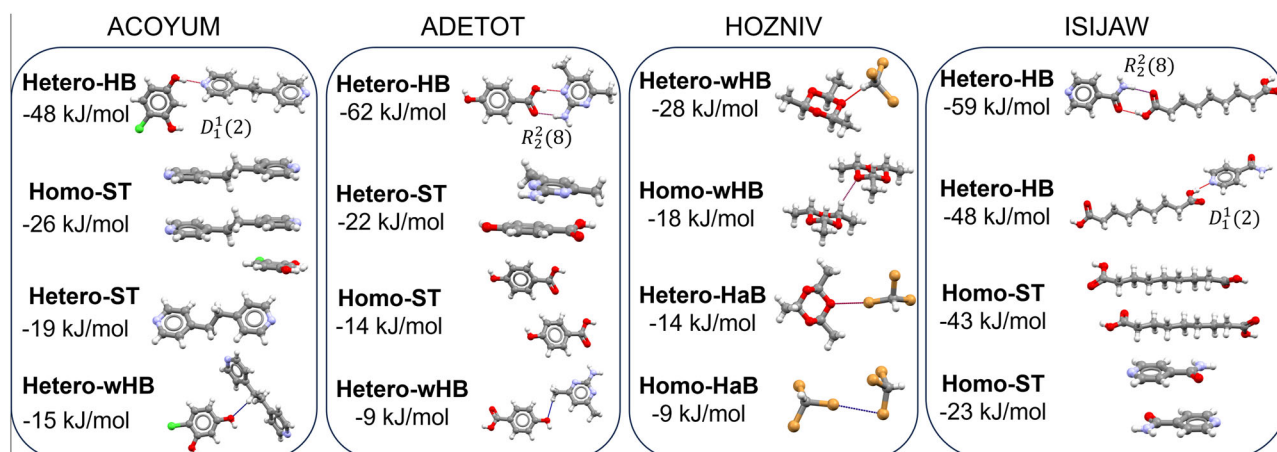
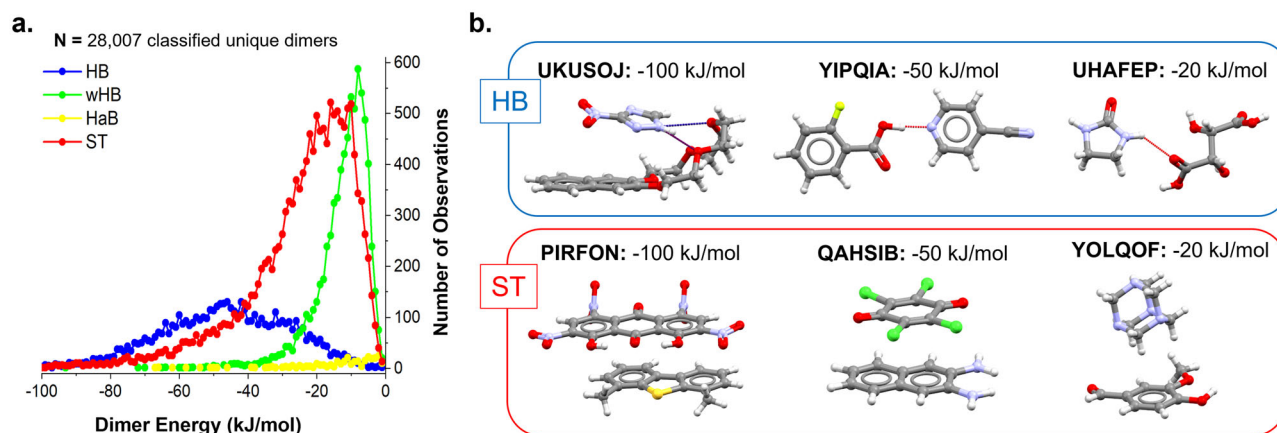


Table 1 | Classification of 69,089 unique dimers (from 3082 cocrystals) into HB, HaB, ST, wHB and other unclassified interaction types with their average interaction energies

| Hierarchy | Dimer Type | All Unique Dimers (69,089) | Fraction all unique dimers | All Dimers (119,396) | Fraction all dimers | Average Interaction Energy (kJ/mol) |
|-----------|--------------|----------------------------|----------------------------|----------------------|---------------------|-------------------------------------|
| 1 | HB | 5623 | 8.1% | 10,481 | 8.8% | -48 ± 21 |
| 2 | HaB | 290 | 0.4% | 481 | 0.4% | -14 ± 14 |
| 3 | ST | 14,782 | 21.4% | 25,011 | 20.9% | -25 ± 17 |
| 4 | wHB | 7312 | 10.6% | 13,763 | 11.5% | -12 ± 7 |
| – | Unclassified | 41,082 | 59.5% | 69,660 | 58.3% | -5 ± 5 |

**Fig. 7 | Dimer interaction energy examples.** Classification and interaction energies of four dimers in four different cocrystal structures of the dataset.**Fig. 8 | Dimer interaction energies per interaction type with examples.** **a** Distribution of dimer interaction energies per interaction type (HBs, wHBs, HaBs and ST). Data on 28,007 classified unique dimers. **b** Examples of hetero HBs and STs of various energies.

The distribution of the strongest interaction per cocrystal by type is shown in Fig. 9b. This pie chart shows that in 48% of the cocrystals, a hetero-HB interaction is the strongest whilst in 20% of the cocrystals, a homo-HB interaction is the strongest. Thus, in 68% of the cocrystals, the HB dimer types are the most stabilising interaction. By contrast, ST interactions are the strongest in 29% of cocrystals, which is significantly less.

To summarise, the analysis of dimer interactions in the 3082 cocrystal structures reveals two clear observations: a) HB interactions are often the strongest interactions and b) ST interactions are often the most common. To analyse these two together, we show the distributions of HB and ST cumulative contributions to the lattice energies of cocrystals in Fig. 10. Remarkably, the distribution of the HB contribution to the cocrystal lattice energy is almost identical to that of the ST contribution (except for values at

0% which are higher for the HB as all compounds will have ST contributions but only those able to HB do). This indicates that taking strength and frequency together, HB interactions are equally as important as ST interactions in molecular cocrystals.

Conclusions

In response to the question, “Are hydrogen bonds the most important interaction in molecular cocrystals?” the answer is no. Whilst HB interactions in cocrystals are often the strongest, ST interactions are often more common. Taking strength and frequency together, both HB and ST interactions contribute equally to the stabilisation of cocrystal lattices. This is true for current (2023) 1:1 two-component cocrystals in the CSD. We remark, however, that the current cocrystal data are biased towards cocrystals

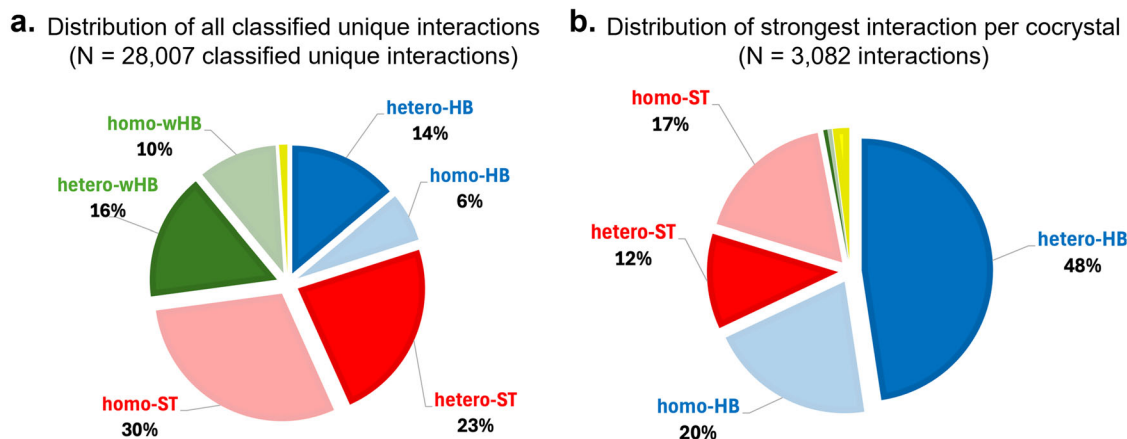


Fig. 9 | Unique and strongest interactions per cocrystal. **a** Distribution of classified unique interactions ($N = 28,007$) by type. **b** Distribution of strongest interaction per cocrystal ($N = 3082$) by type.

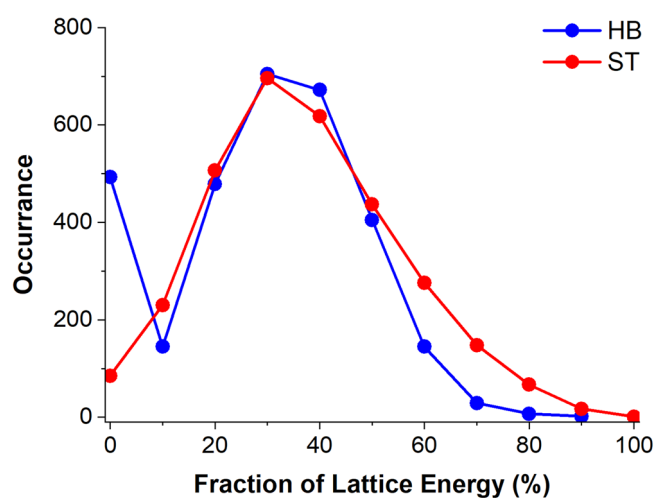


Fig. 10 | HB and ST interaction contributions to cocrystal lattice energies. Distribution of HB and ST contributions to the lattice energy in the 3082 cocrystal dataset.

containing relatively small molecules and cocrystals designed using crystal engineering concepts dominated by principles of HB pairing¹¹. Even with such a bias, HB interactions still do not dominate the stability of cocrystals. Stronger ST-interactions may indeed lead to cocrystallisation, but this will need to be proven with additional studies from the community. As we develop our understanding of stacking and T-type interactions in molecular crystals and the molecules of the future become larger⁴⁶, the HB may become a secondary role player in cocrystal design. As a community, we must recognise this and consider the incorporation of both HBs and ST interactions in overarching cocrystal design. Such a strategy will more likely lead to more successful observations of novel cocrystals with new properties.

Methods and computational details

Dataset

A recent analysis of multicomponent crystals in the CSD revealed 9464 crystal structures of cocrystals containing two main components only (A and B, where the A component is larger than B)⁴⁹. The dataset was further filtered for cocrystal structures with $Z' = 1$ and a 1:1 stoichiometry. Crystal structures with components sitting in special positions (two half molecules), containing iodine atoms, or overlapping atoms (due to disorder) were removed. This reduced the number of cocrystals in the dataset from 9464 to 3082 cocrystal structures.

Dimer and cohesion energy calculations

Dimer calculations were performed on the dataset of 3082 cocrystal structures, all of which completed the process without errors. For this, crystal structures were first retrieved from the CSD and exported as individual CIF files. Dimer energies and cohesive energies for each crystal structure were calculated using the Open Computational Chemistry (OCC) software^{58–60}. For each crystal structure, OCC generates all plausible dimers (AB, AA, and BB) within a given shell radius, set to 30 Å in our case. Using the electron densities of the isolated A and B monomers computed using the B3LYP/6-31 G** model^{61–63}, OCC then calculates the interaction energies of all unique dimers (energies of dimers related by symmetry are not recomputed) and an overall cohesive (or lattice) energy for the crystal based on a simple summation of all dimer interactions (within a 1 kJ/mol convergence) is computed. This simple method for calculating lattice energies has been shown to perform very well, with computed lattice energies with the OCC method deviating from those computed by benchmark DFT-d methods by only ~6.6 kJ/mol on average⁵⁸. Each unique dimer was classified by its composition as either homo (AA or BB) or hetero (AB). For each cocrystal structure, homo and heterodimer contributions to the cohesive energies were then calculated via independent addition of the relevant dimer interaction energies. Our model predicts lattice energies with good accuracy (<6 kJ/mol) at a very reasonable computational time, splitting the contributions into specific molecule-molecular interactions⁶⁴.

Data analysis

For data analysis, unique dimers with an interaction energy of less than -1 kJ/mol were considered. This reduced the number of unique dimers in the dataset from almost 1.5 million to 69,089 dimers. The CSD Python API software was used to retrieve structural information from the dimers such as the number of contacts, HBs, graph set notation, SMILES, and the nature of atoms involved in the dimer interaction.

A distance between two atoms was classified as a contact if it was less than the sum of their van der Waals (vdW) radii plus 0.2 Å. Contacts were considered HBs if they involved a hydrogen bond donor ($D = O, N, \text{ or } S$), a polarised hydrogen atom and a hydrogen bond acceptor ($A = O, N, \text{ or } S$), where the distance between $H \cdots A$ must be less than the sum of their vdW radius and the $D-H \cdots A$ angle larger than 145°. When an HB interaction was found, its graph set was calculated using the method implemented in Mercury⁶⁵. Contacts were classified as HaBs when they involved a halogen donor (F, Cl, or Br) and a halogen bond acceptor (F, Cl, Br, O, N, S) atoms. The remaining contacts were then classified as either wHBs for $H \cdots X$ (where X is an O, N, S, or halogen atom) or STs for $X \cdots Y$ contacts where X and Y could be any heavy atom or hydrogen but excluding $H \cdots H$ contacts. This records $CH \cdots C$ contacts, $C \cdots C$ contacts, and halogen $\cdots C$ contacts typical of stacking or T-interactions.

Data availability

All data generated or analysed during this study are included in this published article (and its supplementary information file, named Supplementary Data 1).

Code availability

The software used to compute the dimer data, Open Computational Chemistry, is freely available from Github ([peterspackman.github.io/occ/](https://github.com/peterspackman/occ/)).

Received: 23 August 2024; Accepted: 26 November 2024;

Published online: 02 December 2024

References

- Bolla, G., Sarma, B. & Nangia, A. K. Crystal Engineering of Pharmaceutical Cocrystals in the Discovery and Development of Improved Drugs. *Chem. Rev.* **122**, 11514–11603 (2022).
- Fleischman, S. G. et al. Crystal engineering of the composition of pharmaceutical phases: Multiple-component crystalline solids involving carbamazepine. *Cryst. Growth Des.* **3**, 909–919 (2003).
- Hall, A., Cruz-Cabeza, A. & Steed, J. What Has Carbamazepine Taught Crystal Engineers? *Cryst. Growth Des.* **24**, 7342–7360 (2024).
- Wang, F. et al. Molecular dynamics application of cocrystal energetic materials: A review. *Nanotechnol. Rev.* **11**, 2141–2153 (2022).
- Tariq, Q. et al. Comparative Studies of Synthesis, Performance, and Applications of Recently Developed CL-20 Based Co-crystals. *Cryst. Growth Des.* **23**, 6974–6987 (2023).
- Bolton, O. & Matzger, A. Improved Stability and Smart-Material Functionality Realized in an Energetic Cocrystal. *Angew. Chem. Int. Ed.* **50**, 8960–8963 (2011).
- Liu, H. Y., Li, Y. C. & Wang, X. D. Recent advances in organic donor-acceptor cocrystals: design, synthetic approaches, and optical applications. *CrystEngComm* **25**, 3126–3141 (2023).
- Xu, W. et al. Nucleic-acid-base photofunctional cocrystal for information security and antimicrobial applications. *Nat. Commun.* **15**, 1–9 (2024).
- Wei, C. et al. Flexible molecular crystals for optoelectronic applications. *Chem. Soc. Rev.* **53**, 3687–3713 (2024).
- Hall, A. V., Musa, O. M. & Steed, J. W. Properties and Applications of Stimuli-Responsive Diacetylenes. *Cryst. Growth Des.* **21**, 3614–3638 (2021).
- Corpinot, M. K. & Bucar, D. K. A Practical Guide to the Design of Molecular Crystals. *Cryst. Growth Des.* **19**, 1426–1453 (2019).
- Steed, J. W. 21st century developments in the understanding and control of molecular solids. *Chem. Commun.* **54**, 13175–13182 (2018).
- Wohler, F. Untersuchungen über das Chinon. *Ann. Chem. Pharm.* **51**, 145–163 (1844).
- Stahly, G. A Survey of Cocrystals Reported Prior to 2000. *Cryst. Growth Des.* **9**, 4212–4229 (2009).
- Helbing, H. *Modern Materia Medica*, 4th ed. (Lehn & Fink, 1895).
- Pfeiffer, P. *Organische Molekülverbindungen* (Ferdinand Enke, Stuttgart, 1922).
- Dickinson, R. G. & Raymond, A. L. The crystal structure of hexamethylene-tetramine. *J. Am. Chem. Soc.* **45**, 22–29 (1923).
- Kofler, L. & Kofler, A. Über Molekülverbindungen des Nicotinsäureamids. *Ber. dtsh. Chem. Ges. A/B* **76**, 718–722 (1943).
- Groom, C. R., Bruno, I. J., Lightfoot, M. P. & Ward, S. C. The Cambridge Structural Database. *Acta Crystallogr. Sect. B Struct. Sci. Cryst. Eng. Mater.* **72**, 171–179 (2016).
- Dai, X. L., Chen, J. M. & Lu, T. B. Pharmaceutical cocrystallization: an effective approach to modulate the physicochemical properties of solid-state drugs. *CrystEngComm* **20**, 5292–5316 (2018).
- Blagden, N., de Matas, M., Gavan, P. T. & York, P. Crystal engineering of active pharmaceutical ingredients to improve solubility and dissolution rates. *Adv. Drug Deliv. Rev.* **59**, 617–630 (2007).
- Sharma, B. D. & McConnell, J. F. The crystal and molecular structure of isocytosine. *Acta Crystallogr.* **19**, 797–806 (1965).
- McConnell, J. F., Sharma, B. D. & Marsh, R. E. Co-crystallization of Two Tautomers: Crystal Structure of Isocytosine. *Nature* **203**, 399–400 (1964).
- Stein, R. S. & Sutherland, G. B. B. M. Effect of Intermolecular Interactions Between CH Frequencies on the Infrared Spectra of N-paraffins and Polythene. *J. Chem. Phys.* **22**, 1993–1999 (1954).
- Kuz'mina, L. G., Bokii, N. G., Struchkov, Y. T., Ravtsov, D. N. & Rohklina, E. M. X-ray diffraction investigation of nonvalent interactions and coordination in organometallic compounds. *J. Struct. Chem.* **15**, 419–423 (1974).
- Pepinsky, R. Crystal engineering: New Concepts in Crystallography. *Phys. Rev.* **100**, 971 (1955).
- Schmidt, G. M. J. Photodimerization in the solid state. *Pure Appl. Chem.* **27**, 647–678 (1971).
- Desiraju, G. R. *Crystal engineering: the design of organic solids* (Elsevier, 1989).
- Desiraju, G. R. Supramolecular Synthons in Crystal Engineering—A New Organic Synthesis. *Angew. Chem. Int. Ed. Engl.* **34**, 2311–2327 (1995).
- Gunawardana, C. A. & Aakeroy, C. B. Co-crystal synthesis: fact, fancy, and great expectations. *Chem. Commun.* **54**, 14047–14060 (2018).
- Almarsson, O. & Zaworotko, M. J. Crystal engineering of the composition of pharmaceutical phases. Do pharmaceutical cocrystals represent a new path to improved medicines?, *Chem. Commun.* **17**, 1889–1896 (2004).
- Aakeröy, C. B., Beatty, A. M. & Helfrich, B. A. Total synthesis" supramolecular style: Design and hydrogen-bond-directed assembly of ternary supermolecules. *Angew. Chem. Int. Ed.* **40**, 3240–3242 (2001).
- Seaton, C. C., Parkin, A., Wilson, C. C. & Blagden, N. Controlling the Formation of Benzoic Acid: Isonicotinamide Molecular Complexes. *Cryst. Growth Des.* **9**, 47–56 (2009).
- Taylor, C. R. & Day, G. M. Evaluating the Energetic Driving Force for Cocrystal Formation. *Cryst. Growth Des.* **18**, 892–904 (2018).
- Cavallo, G. et al. The Halogen Bond. *Chem. Rev.* **116**, 2478–2601 (2016).
- Berger, G., Soubhye, J. & Meyer, F. Halogen bonding in polymer science: from crystal engineering to functional supramolecular polymers and materials. *Polym. Chem.* **6**, 3559–3580 (2015).
- Horiuchi, S. et al. Quantum ferroelectricity in charge-transfer complex crystals. *Nat. Commun.* **6**, 1–7 (2015).
- Black, S. & Davey, R. Synthons reveal how fluorine disrupts pi-pi stacking in halobenzene crystal structures. *CrystEngComm* **26**, 4498–4508 (2024).
- Black, S. Mixed halobenzenes also prefer face...face approaches of aromatic rings, revealing a new inversion synthon. *CrystEngComm* **25**, 3079–3087 (2023).
- Black, S. Geometry and Symmetry of Aromatic Approaches in Halobenzene Crystal Structures: Defining Two Synthons. *Cryst. Growth Des.* **21**, 6981–6991 (2021).
- Sun, L. et al. Creating Organic Functional Materials beyond Chemical Bond Synthesis by Organic Cocrystal Engineering. *J. Am. Chem. Soc.* **143**, 19243–19256 (2021).
- Pidcock, E., Sadiq, G., Stevens, J. & Willacy, R. Aromatic Interactions in the Cambridge Structural Database: Comparison of Interaction Geometries and Investigation of Molecular Descriptors as an Indicator of Strong Interactions. *Cryst. Growth Des.* **22**, 788–802 (2022).
- Devic, T. et al. The maximin principle of π -radical packings. *J. Am. Chem. Soc.* **127**, 14616–14627 (2005).
- Ostroverkhova, O. Organic Optoelectronic Materials: Mechanisms and Applications. *Chem. Rev.* **116**, 13279–13412 (2016).
- Hu, P. et al. Solvent-Dependent Stoichiometry in Perylene-7,7,8,8-Tetracyanoquinodimethane Charge Transfer Compound Single Crystals. *Cryst. Growth Des.* **14**, 6376–6382 (2014).

46. Vainauskas, J., Wahrhaftig-Lewis, A. & Friscic, T. Utilizing “Latent” Carbon: Repositioning Hydrogen-Bonded Synthons and Assemblies via Halogen Bonding to π -Systems. *Angew. Chem. Int. Ed.* **63**, 1–8 (2024).
47. Dunitz, J. & Gavezzotti, A. Supramolecular Synthons: Validation and Ranking of Intermolecular Interaction Energies. *Cryst. Growth Des.* **12**, 5873–5877 (2012).
48. Ding, Y., Zhao, Y. & Liu, Y. Organic cocrystals: From high-performance molecular materials to multi-functional applications. *Aggregate* **e626**, 1–26 (2024).
49. Cruz-Cabeza, A. J., Lusi, M., Wheatcroft, H. P. & Bond, A. D. The role of solvation in proton transfer reactions: implications for predicting salt/co-crystal formation using the ΔpK_a rule. *Faraday Discuss.* **235**, 446–466 (2022).
50. Cruz-Cabeza, A. J., Reutzel-Edens, S. M. & Bernstein, J. Facts and fictions about polymorphism. *Chem. Soc. Rev.* **44**, 8619–8635 (2015).
51. Fábíán, L. Cambridge Structural Database Analysis of Molecular Complementarity in Cocrystals. *Cryst. Growth Des.* **9**, 1436–1443 (2009).
52. ChemAxon, Marvin (version 23.16), www.chemaxon.com, (accessed 13/03, 2024).
53. Spackman, P. et al. CrystalExplorer: a program for Hirshfeld surface analysis, visualization and quantitative analysis of molecular crystals. *J. Appl. Crystallogr.* **54**, 1006–1011 (2021).
54. Chickos, J. S. & Gavezzotti, A. Sublimation Enthalpies of Organic Compounds: A Very Large Database with a Match to Crystal Structure Determinations and a Comparison with Lattice Energies. *Cryst. Growth Des.* **19**, 6566–6576 (2019).
55. Colombo, V., Lo Presti, L. & Gavezzotti, A. Two-component organic crystals without hydrogen bonding: structure and intermolecular interactions in bimolecular stacking. *CrystEngComm* **19**, 2413–2423 (2017).
56. Jiang, M., Zhen, C., Li, S., Zhang, X. & Hu, W. Organic Cocrystals: Recent Advances and Perspectives for Electronic and Magnetic Applications. *Front. Chem.* **9**, 1–16 (2021).
57. Hu, B. & Zhang, Q. Recent Progress in Polycyclic Aromatic Hydrocarbon-based Organic Co-crystals. *Chem. Rec.* **21**, 116–132 (2021).
58. Spackman, P. R., Walisinghe, A. J., Anderson, M. W. & Gale, J. D. CrystalClear: an open, modular protocol for predicting molecular crystal growth from solution. *Chem. Sci.* **14**, 7192–7207 (2023).
59. Mackenzie, C. F., Spackman, P. R., Jayatilaka, D. & Spackman, M. A. CrystalExplorer model energies and energy frame-works: extension to metal coordination compounds, organic salts, solvates and open-shell systems. *IUCrJ* **4**, 575–587 (2017).
60. Thomas, S. P., Spackman, P. R., Jayatilaka, D. & Spackman, M. A. Accurate Lattice Energies for Molecular Crystals from Experimental Crystal Structures. *J. Chem. Theory Comput.* **14**, 1614–1623 (2018).
61. Pritchard, B. P., Altarawy, D., Didier, B., Gibson, T. D. & Windus, T. L. New Basis Set Exchange: An Open, Up-to-Date Resource for the Molecular Sciences Community. *J. Chem. Inf. Model.* **59**, 4814–4820 (2019).
62. Feller, D. The role of databases in support of computational chemistry calculations. *J. Comput. Chem.* **17**, 1571–1586 (1996).
63. Schuchardt, K. L. et al. Basis set exchange: A community database for computational sciences. *J. Chem. Inf. Model.* **47**, 1045–1052 (2007).
64. Spackman, P., Spackman, M. & Gale, J. A transferable quantum mechanical energy model for intermolecular interactions using a single empirical parameter. *IUCrJ* **10**, 754–765 (2023).
65. Etter, M. C., Macdonald, J. C. & Bernstein, J. Graph-Set Analysis of Hydrogen-Bond Patterns in Organic-Crystals. *Acta Crystallogr. Sect. B Struct. Sci. Cryst. Eng. Mater.* **46**, 256–262 (1990).
66. Sakurai, T. The crystal structure of the triclinic modification of quinhydrone. *Acta Cryst.* **19**, 320–330 (1965).

Acknowledgements

A.J.C.C. thanks the Hamilton HPC Service of Durham University for computational support. A.V.H. thanks The Leverhulme Trust for Early Career Fellowship funding [grant number ECF-2022-196].

Author contributions

A.J.C.C. constructed the research idea and experiment design, managed data collection and interpretation, and contributed to manuscript writing. P.R.S. contributed expertise on calculations and manuscript feedback. A.V.H. contributed to the research ideas, data interpretation, and manuscript writing.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s42004-024-01380-3>.

Correspondence and requests for materials should be addressed to Aurora J. Cruz-Cabeza or Amy V. Hall.

Peer review information *Communications Chemistry* thanks Benjamin Tan and the other, anonymous, reviewers for their contribution to the peer review of this work.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024