Acta Crystallographica Section B
Structural Science
ISSN 0108-7681

Failures of fractional crystallization: ordered co-crystals of isomers and near isomers

Steven P. Kelley,[a] László Fábián[b] and Carolyn Pratt Brock[a]*

[a]Department of Chemistry, University of Kentucky, Lexington, KY 40506-0055, USA, and [b]Pfizer Institute for Pharmaceutical Materials Science, Cambridge Crystallographic Data Centre, 12 Union Road, Cambridge CB2 1EZ, England

Correspondence e-mail: cpbrock@uky.edu

Received 18 September 2010
Accepted 18 November 2010

A list of 270 structures of ordered co-crystals of isomers, near isomers and molecules that are almost the same has been compiled. Searches for structures containing isomers could be automated by the use of IUPAC International Chemical Identifier (InChI™) strings but searches for co-crystals of very similar molecules were more labor intensive. Compounds in which the heteromolecular $A\cdots B$ interactions are clearly better than the average of the homomolecular $A\cdots A$ and $B\cdots B$ interactions were excluded. The two largest structural classes found include co-crystals of configurational diastereomers and of quasienantiomers (or quasiracemates). These two groups overlap. There are 114 co-crystals of diastereomers and the same number of quasiracemates, with 71 structures being counted in both groups; together the groups account for 157 structures or 58% of the total. The large number of quasiracemates is strong evidence for inversion symmetry being very favorable for crystal packing. Co-crystallization of two diastereomers is especially likely if a 1,1 switch of a methyl group and an H atom, or of an inversion of a [2.2.1] or [2.2.2] cage, in one of the diastereomers would make the two molecules enantiomers.

1. Introduction

Failures of fractional crystallization[1] are rare. As a separations method fractional crystallization is so simple, powerful and inexpensive that it is the method of choice in the chemical industries. It works well for almost everything except for pairs of separable (*i.e.* resolvable) enantiomers. Its success stems from the fact that most substances crystallize as pure compounds composed of a single molecule or a single set of ions.[2]

Failures of fractional crystallization imply either the existence of an ordered solid-state compound or of a 'mixed' crystal (sometimes called a solid solution[3]). A solid-state compound (also known as a co-crystal) corresponds to a maximum in a T–X (temperature–composition) phase

[1] Fractional crystallization is a separations method based on differences in solubility. A solution containing the desired material and associated impurities is cooled or the solvent is allowed to evaporate. Crystals of the desired material, which usually has the highest concentration in the original solution, are normally deposited first, while the impurities remain in solution. The method will fail if the desired material and a co-crystal of it with one of the impurities present have similar solubilities.

[2] Sometimes crystals include a solvent molecule but solvates are not usually considered to be failures of fractional crystallization.

[3] The term solid solution suggests random disorder of the components. The less specific term mixed crystal is probably preferable when it is known that two different molecules occupy, on average, the same crystallographic site, but it is not known whether there is any correlation of occupancies in one or more directions. The situation is further complicated by the use of the term mixed crystal by some authors to describe an ordered co-crystal. In this paper the term mixed crystal always implies whole-molecule disorder.

diagram and has a fixed stoichiometric ratio; such compounds are represented on phase diagrams by vertical lines. Mixed crystals, on the other hand, are disordered solids in which several molecules or ions of different types occupy the same site. Crystals having such whole-molecule, compositional disorder correspond to a two-dimensional region of a *T*–*X* phase diagram. As this paper is concerned with ordered compounds in which the different molecules occupy different sites, mixed crystals will not be considered further.

General interest in molecular solid-state compounds grew rapidly after the pioneering work of Etter (1990, 1991) on co-crystals. The more recent field of crystal engineering depends on the ability of chemists to design systems for which fractional crystallization will fail reliably, usually because there is a strong attractive interaction (hydrogen bonding or other donor–acceptor interaction) between the two kinds of molecules. Herbstein's (2005) two-volume book provides an overview.

Solid-state compounds are much more likely to form if there is some strong, specific attraction between the molecules. Consider the expected deposition of co-crystals from a solution containing both an acid and a base: A compound is expected because complete proton transfer leads to a salt (such as an ammonium chloride), which can (assuming a favorable arrangement of the ions) have a very low energy relative to crystals of the neutral molecules, and because incomplete proton transfer still gives a compound with hydrogen bonds that are usually better than can be formed by the acid or base alone.

Fractional crystallization is also expected to fail in the case of solutions of resolvable enantiomers. This systematic failure has long been understood as evidence that symmetry operations of the second kind, especially inversion, are very favorable for crystal packing. In the case of separable enantiomers the success of fractional crystallization is so surprising that it has been given a special name (*i.e.* spontaneous resolution).

We became interested in failures of fractional crystallization for which there was no obvious reason because we had worked with or knew of several co-crystals formed by isomers [see Fig. 1; Cambridge Structural Database (Allen, 2002; hereafter the CSD) refcodes CTHXDL, Ruysink & Vos (1974); VENZOD, Ermer *et al.* (1989); RIHLUQ, Lloyd *et al.* (2007); POVSEY02 = POVSAT,[4] Loehlin *et al.* (2008)]. We wondered how frequent such co-crystals are and under what circumstances they are likely to form. We therefore decided to look for ordered co-crystals whose existence cannot be explained easily by standard donor–acceptor interactions (*e.g.* complete or partial proton or electron transfer), or by the inclusion of small molecules (*e.g.* solvent molecules) to fill voids or to satisfy hydrogen-bond donors or acceptors. Most of the co-crystals we were looking for would be composed of isomers or of near isomers, *i.e* molecules that differ by only a few atoms.

A related goal of this project was to make a more complete list of quasiracemates (see Zhang & Curran, 2005; Lineberry *et al.*, 2008, and references therein). Quasiracemates (also called pseudoracemates) are co-crystals formed by molecules (sometimes called quasienantiomers) that would be enantiomers but for a small substitution of atoms [*e.g.* (*R*)-2-bromobutane and (*S*)-2-chlorobutane]. The existence of quasiracemates is evidence that even an approximate inversion center is very favorable for crystal packing. The task of making a list of quasiracemates has been characterized as 'difficult at best and certainly impractical' (Lineberry *et al.*, 2008).

Although we could think of no way to completely automate the necessary searches of the CSD, the availability of the IUPAC International Chemical Identifier (InChI™) text strings (Stein *et al.*, 2003, 2006) proved invaluable. These strings code both molecular connectivity and stereochemistry in a way that makes computer comparisons straightforward.

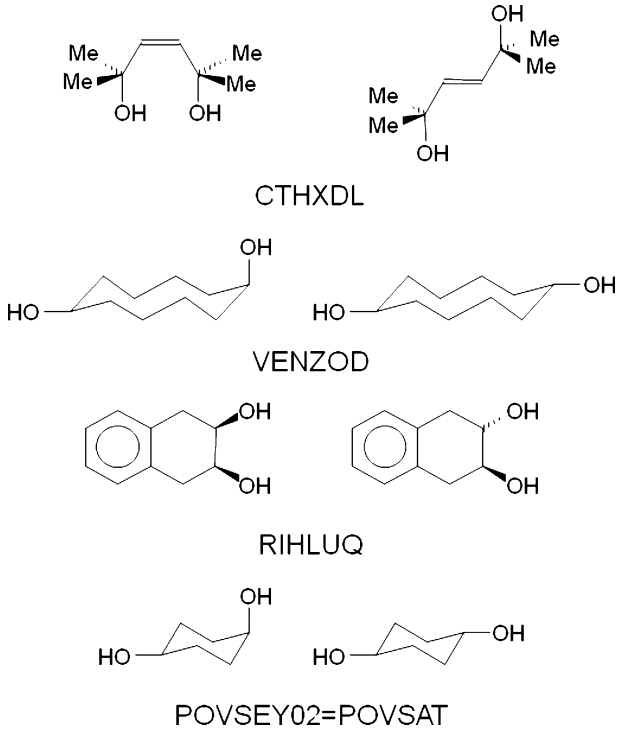

Figure 1
Examples of co-crystals known to us at the beginning of this study. All exist because the two isomers can together form a better set of hydrogen bonds and fill space more densely than they can individually. (See text for individual references.)

2. Terminology

The IUPAC Compendium of Chemical Terminology (http://goldbook.iupac.org/; see also McNaught & Wilkinson, 1997) divides isomers, which have the same chemical formulae, into two basic groups. Skeletal (or constitutional) isomers have different sets of bonds; stereoisomers have the same bonds but

[4] Note that POVSEY and POVSEY01 are a $Z' = 2$ structure containing two conformers of *trans*-cyclohexane-1,4-diol while POVSEY02 and POVSEY03 are a $Z' = 1$ structure of the 1:1 compound of *cis*- and *trans*-cyclohexane-1,4-diol. The refcode family of the 1:1 compound was changed to POVSAT in the November 2010 release of the CSD.

Table 1
Classification of co-crystals included in the final list.

Symbol	Description	Number
	Total number of different structures (not including co-crystals of tautomers related by a proton shift)	270
	Number of structures containing molecules (not including tautomers) that can interconvert faster than crystals usually grow	28 = 1 (*D*) +1 (*D SA*) + 1 (*D Q SA*) + 3 (*K*) + 13 (*R*) + 7 (*C*) + 2 (*O*)
D	Configurational diastereomers; category includes salts of diastereomeric ions and compounds of enantiomers in which the enantiomeric ratio is not 1:1	114 = 28 (*D*) + 66 (*D Q*) + 15 (*D SA*) + 5 (*D Q SA*)
DA	Molecules that are almost configurational diastereomers	1
R	Other kinds of diastereomers (*cis*/*trans* and *E*/*Z*) (13 of the pairs contain molecules like imines that can interconvert)	20 = 19 (*R*) + 1 (*R SA*)
Q	Quasiracemates (or molecules that are almost enantiomers)	114 = 36 (*Q*) + 66 (*D Q*) + 5 (*D Q SA*) + 5 (*Q K*) + 2 (*Q KA*)
K	Skeletal (or constitutional) isomers	26 = 20 (*K*) + 5 (*Q K*) + 1 (*K SA*)
KA	Molecules that are almost skeletal isomers	12 = 10 (*KA*) + 2 (*Q KA*)
SA	Molecules that are almost the same	57 = 35 (*SA*) + 5 (*D Q SA*) + 15 (*D SA*) + 1 (*K SA*) + 1 (*R SA*)
C	Conformational diastereomers (not separable)	7
O	Other	19
T	Tautomers (skeletal isomers related by a proton shift and an interchange of single and multiple bonds; not included in final count)	64

different three-dimensional structures. Both enantiomers (which are molecules[5] related by a mirror operation) and diastereomers are stereoisomers. The class of diastereomers includes configurational diastereomers (*e.g.* *R*,*R*- and *R*,*S*-2-chlorobutan-2,3-diol), *cis*/*trans* isomers (*e.g.* *cis*- and *trans*-cyclohexane-1,4-diol) and *E*/*Z* isomers (*e.g.* *E*- and *Z*-1,2-dichloroethylene). Although the general class of stereoisomers also includes conformers (*e.g.* the diaxial and diequatorial conformational diastereomers of *trans*-cyclohexane-1,4-diol), which can interconvert without breaking any bond, these isomers normally interconvert rapidly except in the solid state and so are not separable. Tautomers are readily interconvertible skeletal isomers; they are most often related by a change in position of an H atom and the rearrangement of single and multiple bonds. Tautomers may also be related in other ways (*e.g.* ring-chain tautomerism and valence tautomerism); several such structures were found [*e.g.* BAXGAI (Glen *et al.*, 1982) and SIFBIT (Poliakov & Shevchenko, 2007)]. Structures of this latter group of tautomers were included in the final list because in each case one isomer had a C—O bond not present in the other, and because there was evidence that interconversion is hindered under some conditions (*e.g.* in the absence of an acid catalyst).

The classifications used in this work are shown in Table 1. Pairs of molecules that are almost enantiomers might have been designated as EA (for enantiomers almost), but the letter Q was used instead because the term quasiracemate has been used by so many previous authors (see Lineberry *et al.*, 2008). Racemic compounds, which are composed of a pair of true enantiomers, were not considered unless the enantiomeric ratio is other than 1:1 (10 examples). Some pairs of molecules were assigned to two groups (*e.g.* *D* and *Q* for configurational diastereomers that are almost enantiomers) or even three.

3. Methodology

The search was limited to molecules that are generally classified as 'organic', although some additional elements (see below) were allowed. The study specifically excluded metal complexes in order to make the project tractable. Bonded units in organometallic structures can be difficult to define, and it can be difficult to determine whether two related metal complexes are separable or interconvert.

The goal was to find compounds that correspond to a solid-state compound, *i.e.* to an unexpected maximum of fixed composition in a *T*–*X* phase diagram. Such a maximum should occur at a composition that can be expressed as a ratio of small integers (*e.g.* 1:1, 1:2, 1:3, 2:3). Mixed crystals, in which different molecules occupy the same site with at least approximately similar occupancies, were excluded. (For a list of 30 mixed crystals of diastereomers see Plutecka *et al.*, 2010.) While it is possible that some apparently mixed crystals might be better described as ordered molecular compounds in a larger or lower-symmetry unit cell, there was no way we could investigate that possibility.

3.1. Searches

3.1.1. First search. This search was first performed on version 5.29 (November 2007) of the CSD and the January 2008 update. Search criteria were:

(i) coordinates archived in the CSD; no error flags;

(ii) at least two uncharged residues, neither of which is a 'common' (Görbitz & Hersleth, 2000) solvent;

(iii) permitted elements restricted to C, H, N, O, S, Se, Te, P, As, B, Si, Ge, F, Cl, Br, I (as well as Na^+ and K^+ counterions).

This search gave a list of *ca* 5200 hits, most of which are structures of unusual solvates or of donor–acceptor pairs. This list was screened by hand by two of us to find the modest

[5] The term 'molecules' as used here is an abbreviated form of 'molecules and molecular ions' and so does not imply electrical neutrality. A 'molecule' in this paper is a group of atoms (other than conventional polyatomic ions like NH_4^+ and SO_4^{2-}) connected by covalent bonds.

Table 2
Types of co-crystals found in the search of version 5.29 of the CSD, but excluded from the final list because the component molecules are not even approximately isomeric.

Crystals in which the second component is a standard solvent molecule and crystals that are highly disordered are not included. Approximate percentages of occurrence are given; the estimated uncertainties are several percentage points. The total number of structures classified is 3851.

Hydrogen bonded (carboxylic acid, alcohol or crown ether + an amine, oxime or pyridine; carboxylic acid + amide *etc.*) (half-neutralized acids and bases)	58%
Other donor–acceptor complexes (*e.g.* a substituted pyridine + C_6F_5I)	5%
Stacking complexes (an electon-donor + electron-acceptor molecule; stacked parts of molecules are usually planar; includes compounds of aromatic molecules in which at least most of the H atoms in one molecule are replaced by F atoms in the other)	19%
Packing complexes (*e.g.* structures including molecules like cholic acid, 1,1-bis(*p*-hydroxyphenyl)cyclohexane, 1,1,6,6-tetraphenylhexa-2,4-diyne-1,6-diol and cavitands that are known to form large numbers of inclusion compounds). Includes compounds of fullerenes with large molecules having flat or concave surfaces. Charge transfer may be important in some of the compounds.	19%

number (*ca* 150) of unexpected co-crystals of closely related molecules. In general the molecules considered to be near isomers differ by at most three non-H atoms, but we also kept lists of structures in which the difference was larger (*e.g.* a monomer plus its dimer); these last structures are mostly in the O (other) category (see Table 1). While going through the list we also classified the structures of the 'expected' compounds by type (see Table 2).

Screening was done by looking at the two-dimensional chemical line drawings stored in the CSD and at the structure as visualized with the program *MERCURY* (Macrae *et al.*, 2008). We discovered that the line drawings, which do not include any three-dimensional stereochemical information, did not always correspond to our expectations. While two line drawings are usually displayed for 'compounds' composed of two interconvertible tautomers, and sometimes for two conformers that interconvert rapidly in solution, only one line drawing is usually displayed for compounds of separable configurational diastereomers. Most compounds of diastereomers were therefore missed by the manual search. Later (see below) we worked out a semi-automatic procedure for finding compounds of these true isomers.

The compounds of tautomers related by H-atom shifts are not of direct relevance for this study but are listed in the compilation as a by-product of the survey (see also the list given by Cruz-Cabeza & Groom, 2011). Conformers were included in the final list only if a substantial barrier to interconversion seemed likely (*e.g.* axial and equatorial conformers of six-membered rings).

The manual search for co-crystals of closely related molecules was repeated using version 5.31 (November 2009) of the CSD. Another thousand entries were screened but the information in Table 2 was not updated.

3.1.2. Automated search for compounds of isomers and diastereomers. This second type of search took advantage of software for creating and comparing IUPAC International Chemical Identifier (InChI™) strings (Stein *et al.*, 2003, 2006). The advantage of this particular text representation of molecular structure is that it includes information about stereochemistry explicitly.

The procedure started with a search of the CSD [originally version 5.30 (November 2008) and later version 5.31 (November 2009)] using the following criteria:

(i) coordinates archived in the CSD; no error flags;

(ii) permitted elements restricted as described above;

(iii) either (1) two residues that have the same chemical formula for non-H atoms or (2) $Z' > 1$. The second possibility was necessary because compounds composed of configurational diastereomers had usually been entered in the CSD as $Z' > 1$ structures of a single substance.[6] The charge associated with the residues was not considered. H atoms were excluded from the test since it cannot be assumed that their positions, or even their numbers, are correct.

After this search was completed the CSD entries located were exported as mol2 files, which were split into individual model files (one covalently bonded unit per file) using a program written at the CCDC. Each file was then converted to an InChI string using version 2.2.1 of Open Babel (http://openbabel.org/). InChI strings provide a unique textual representation of chemical substances; these strings are composed of layers, which describe the substance in different levels of detail (see Fig. 2). The constituents of compounds of skeletal isomers have InChI strings that differ in the connectivity section of the main layer; the molecules in compounds of diastereomers have InChI strings that differ in the stereochemistry layer. The final list of 'hits' for this search therefore included the CSD entries composed of several residues that have the same chemical formulae (with the possible exception of H atoms), but different InChI strings.

3.2. Investigation of list entries

3.2.1. Classification of interaction type in 'expected' compounds. The distribution shown in Table 2 is necessarily approximate because while going through thousands of compounds it was impossible to spend much time thinking about them individually. The goal was to specify the type of interaction that seemed most responsible for the existence of the compound. The 'crystal engineers', however, have been clever about increasing the chances of compound formation by building in the possibility of more than one type of interaction. Some stacking compounds also have important hydrogen bonds or other donor–acceptor interactions. Molecules (*e.g.* several families of dialcohols) have been synthe-

[6] The CSD does not claim to include stereochemical details although that information can be extracted from the atomic coordinates. We found that in many cases the differences between the two configurational isomers that co-crystallize are so subtle that they would be difficult to spot without software that allows overlaying the two molecules and then rotating them together. In some cases the compound name indicates that more than one configurational diastereomer is present but in other cases it does not.

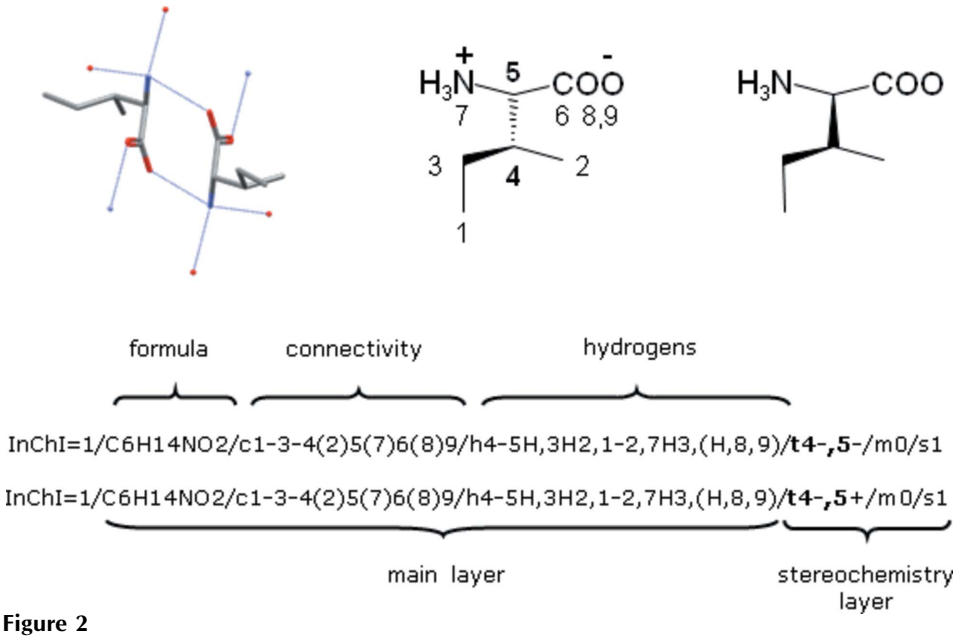

Figure 2
InChI string representations of L-isoleucine and D-*allo*-isoleucine (XADVED; Dalhus & Görbitz, 2000), which are configurational diastereomers that form a quasiracemate. The main layer, which defines the connectivity, is the same for the two isomers, while their stereochemistry layers differ. The configurations at C5 are different for the two zwitterions, but the configurations at C4 are the same. If the methyl group and the H attached to C4 were switched in one diastereomer the two zwitterions would become enantiomers.

sized so that it is unlikely that they can fill space densely while also satisfying their potential donor and/or acceptor groups; such molecules are almost guaranteed to include solvent when they crystallize. Fullerenes have been co-crystallized with molecules that have flat or concave surfaces and that also can act as electron donors or acceptors. In some sense all co-crystals are packing compounds because no solid-state compound can be formed if the molecules cannot together fill space densely.

Crystals containing cyclodextrins and other cavitands are good examples of the classification problem. These co-crystals, which often include more than one kind of guest molecule, could be classified as packing compounds (because of the necessity of filling the cavitand cavity), as hydrogen-bonded compounds (because many of the cavity-filling molecules form hydrogen bonds to the host molecule) or (if the guests are all small relative to the main molecule) as solvates. The classification shown in Table 2 is only approximate, but it does provide a list of the types of interactions that are likely to lead to the formation of co-crystals and a guide to their prevalence. Some additional information about how the classifications were made is included with the supplementary material.[7]

3.2.2. Final list of 'unexpected' compounds. Structures on the final list of co-crystals were checked carefully; in almost all cases the original publication was consulted. The new overlay function in *MERCURY*[8] proved invaluable, especially for seeing how co-crystallized diastereomers differ. The CIFs for some structures were retrieved so that displacement ellipsoids could be examined for evidence of disorder.

In the case of multiple determinations of the same structure only one example was retained in the list, with the choice based on the temperature and precision of the determination, date of publication and accessibility of the journal. Multiple solvates of a basic structure type (*e.g.* GANROD and its variants DEHZIA, DEMLOX, GANQUI, GANRAP, GANRET and GANRIX; Manoj *et al.*, 2006) were treated as multiple determinations of the same structure, but the excluded solvates are listed in a different part of the deposited spreadsheet.

There were a few other rejections. The compound with refcode AWAJOX (Fernandes & Levendis, 2004), which contains a monomer and its photodimer, was rejected because it is the product of a single-crystal-to-single-crystal reaction of a reactant crystal (ZZZNQS08) that has $Z' = 3$ but with only two of the molecules in a mutual orientation favorable for photoreaction. The structure DERCUZ (Zouev *et al.*, 2006) was rejected for similar reasons, but LEZMIM, a co-crystal of a reactant and its photoproduct, was retained because the partially reacted material was recrystallized from solution before diffraction data were collected (Jones *et al.*, 1994). Compounds that were considered but eventually excluded from the final list were retained in a separate part of the spreadsheet so that the work of finding them would not be lost.

The generated list is necessarily incomplete. Since the search for compounds other than those of exact skeletal isomers and diastereomers required so much human intervention, omissions are a near certainty. There is also the possibility that ordered compounds of diastereomers *etc.* were missed by the original authors, who might have refined a disordered model in a non-Sohnke space group (*i.e.* in a group including an improper symmetry operation such as inversion). Walker *et al.* (1999) described two ordered, well characterized quasiracemates composed of configurational diastereomers (LIPYUE, see Fig. 3, and LIPZEP; both with space group $P1$) and demonstrated just how easy it would be to make the mistake of doing the refinement in a space group with too many symmetry elements. The diastereomers in these two structures are enantiomeric at the more important chiral center(s) but homochiral at the —CH(Me)— center. Both structures can be refined very well in $P\bar{1}$ with $Z' = 1$ if H/Me disorder is allowed, but other information indicated that the

[7] Supplementary data for this paper are available from the IUCr electronic archives (Reference: GP5040). Services for accessing these data are described at the back of the journal.
[8] This function became generally available in the November 2009 release.

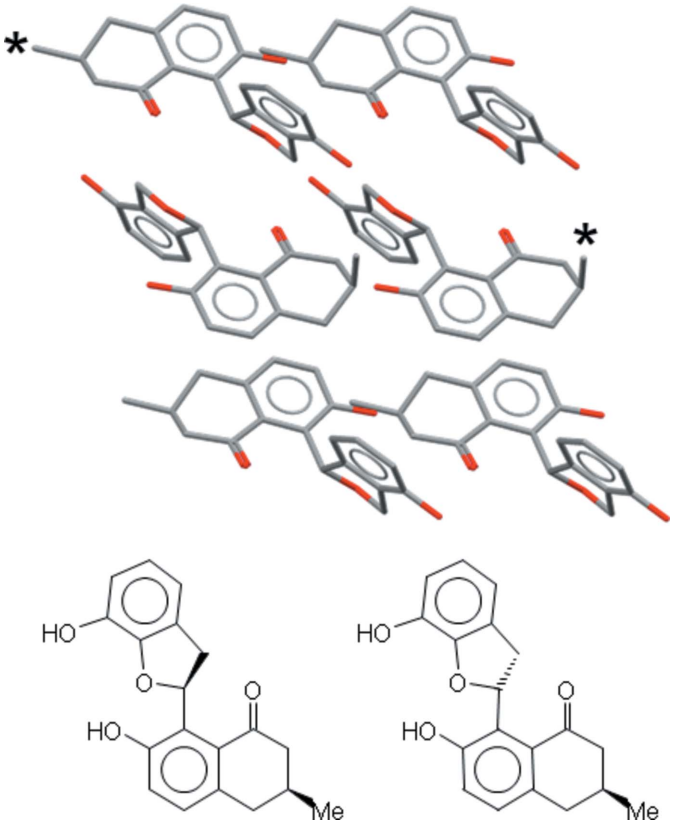

Figure 3
Part of the structure (molecules shown have centroids with $0 \leq x \leq 1$) of the natural product and quasiracemate LIPYUE (Walker *et al.*, 1999; space group *P*1). The pseudoinversion symmetry is broken by the two methyl groups, which are marked by asterisks.

samples each had to be composed of two single configurational diastereomers rather than of two diastereomeric pairs of enantiomers.

Ordered salts resulting from failed diastereomeric resolutions [*e.g.* $(\text{D}^+\text{L}^-)\cdot(\text{L}^+\text{L}^-) = \text{D}^+\cdot\text{L}^+\cdot 2\text{L}^-$;[9] sometimes known as double salts] were also included; in these 13 cases the material is a compound of two diastereomeric salts rather than of two molecules. Also included were ordered structures in which unequal numbers of enantiomers are present. These 10 products of so-called 'unbalanced crystallization' (Albano *et al.*, 1969; Cai *et al.*, 2001), which can be viewed as co-crystals of racemic compounds and pure enantiomers, do not fit easily into a classification scheme developed for molecules rather than for crystals, so we simply grouped them with the more conventional diastereomeric compounds.

3.3. Complications

3.3.1. One- *versus* two-component systems. We made judgements about whether the compound components were stable to interconversion over the time required for crystal growth. If the components do not interconvert then there are two different solutes, and their ratio in solution can be varied. If interconversion is more rapid, then there are two forms of a single solute, and their ratio in solution (or in the melt) is determined by their energy difference. A crystal containing two different molecular forms that can interconvert (*e.g.* diaxial and diequatorial *trans*-1,4-diethynylcyclohexane-1,4-diol; CEMQOA, Bilton *et al.*, 1999) is not a compound in the thermodynamic sense.

Distinguishing between one- and two-component systems requires knowledge of chemical reactivity. The two molecules in CIDSEN (see Fig. 4) are configurational diastereomers, but the published paper (Valente & Schomaker, 1984) says that the two forms interconvert in solution by a hemiketal inversion. The distinction between one- and two-component systems also depends on the temperature range considered.

Structures containing two conformers were identified only if the CSD entry shows two different chemical line drawings. Two drawings are normally shown if the CSD editors concluded that a significant energy barrier separates the two forms (as is the case in some conformers) or if the atomic connectivity is different (as it is in tautomers).

In the end we decided that structures composed of molecules that can interconvert should be retained in the list if the two forms are separated by a non-trivial energy barrier, but a special notation that the constituents can interconvert was added. CIDSEN was retained, as were structures like POVSEY (Steiner & Saenger, 1998) and SAKYUY (Seiler & Dunitz, 1989; see Fig. 4) that contain two conformers related

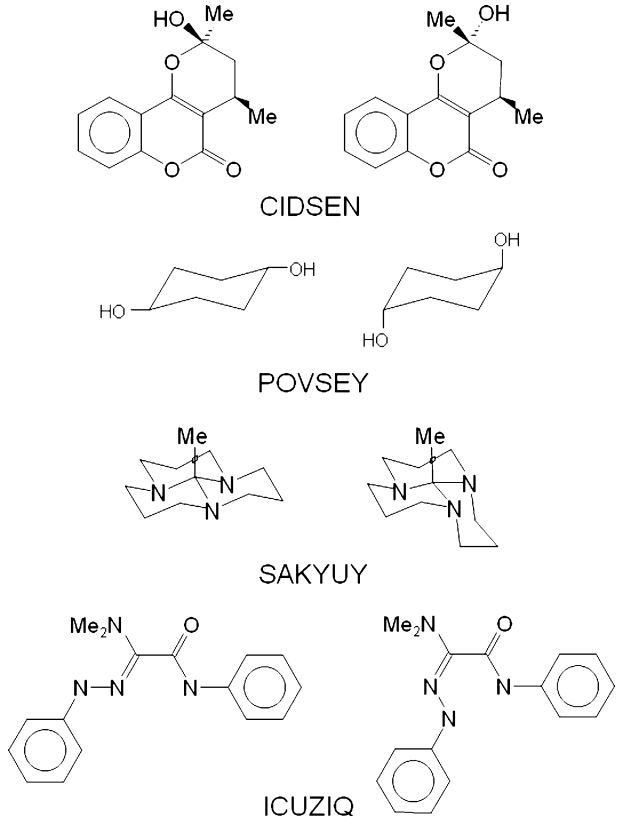

Figure 4
Examples of 'co-crystals' composed of molecules that can interconvert in solution. (See text for individual references.)

[9] Where the D and L in D^+, L^+, D^- and L^- refer to the optical rotations of the unprotonated amine and undissociated acid.

by a ring inversion. The ten structures containing both *E* and *Z* imines [see *e.g.* ICUZIQ (Frohberg *et al.*, 2006) in Fig. 4] were listed in category *R* (see Table 1), but were marked as being able to interconvert. Sulfoxides [$R-S(=O)-R'$], which have S atoms that are tetrahedral because of the lone pair, do not usually racemize at room temperature and so were included as diastereomers without comment. Tautomers related by an H-atom shift were not, however, retained in the final list because the kinetic barrier to interconversion is expected to be very low.

3.3.2. Disorder. Structures were not considered at all if complete disorder of the components had been found. It was less clear, however, whether structures having some compositional disorder but two very different molecular sites should be excluded from the list.

Several structures were found in which there is disorder at one molecular site, but not at the other. In the centrosymmetric structure MODBEM (Ansari *et al.*, 2002) one of the two sites is ordered but at the other there is a 61:39 mixture of diastereomers; the overall composition is therefore 82:18. In GINXIL (Czaplik *et al.*, 2007) one site is ordered but there is a 50:50 disorder at the other. In VEFMEZ (Gültekin & Hökelek, 2006) two of the three molecules are one enantiomer and one molecule is the other, but judging from the displacement ellipsoids, one of the two homochiral sites is partly occupied by a diastereomeric impurity (different configuration at one of the three stereocenters). The impurity level could be as high as 17%.

None of these three structures is included in the final list. The mixed crystal of *cis*- and *trans*- (*i.e.* *R*,*S* and either *R*,*R* and *S*,*S*) 2,3-tetralindiol (RIHLUQ, see Fig. 1) was, however, retained because there is thermodynamic evidence (see Lloyd *et al.*, 2007) of a 1:1 stoichiometric compound between *cis* molecules and enantiomerically pure *trans* molecules.

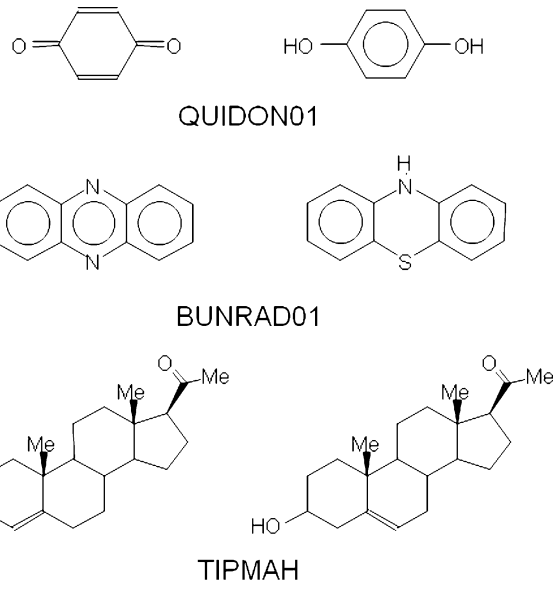

Figure 5
Examples of co-crystals of molecules that form such good donor–acceptor pairs that they were excluded from the list: QUIDON (Sakurai, 1968), BUNRAD01 (Stezowski *et al.*, 1983) and TIPMAH (Lancaster *et al.*, 2007).

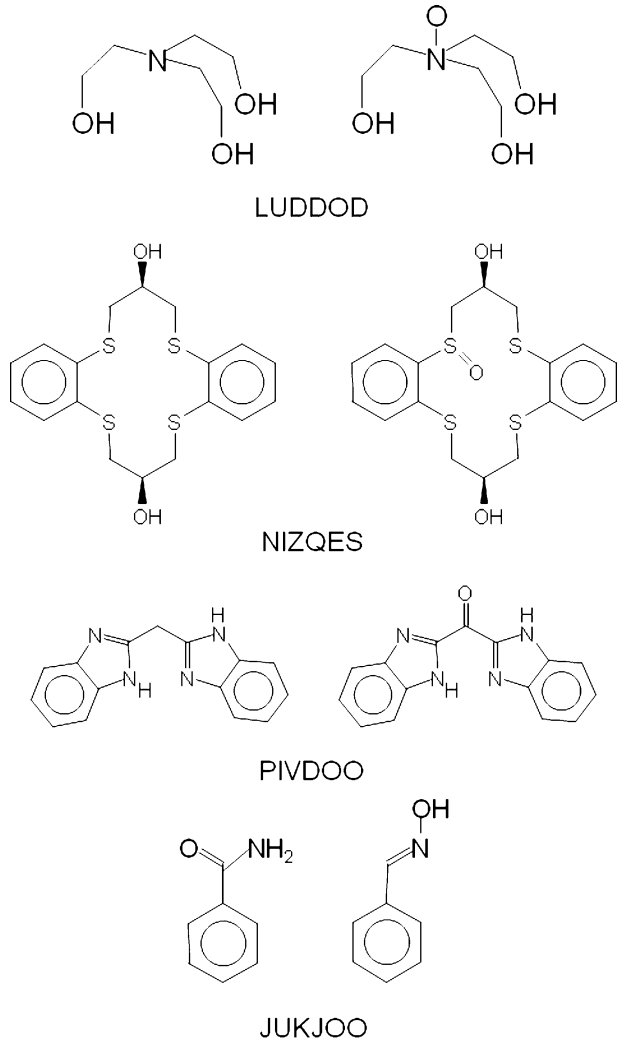

Figure 6
Examples of co-crystals in which there is potential for good donor–acceptor interactions but which were retained in the list (see text for individual references).

More minor conformational disorder (*e.g.* rotations of $-CF_3$ groups, up–down disorder of a $-CH_2-$ group in a cyclopentane ring) was ignored because disorder of this type does not raise any question about the identity of the molecule.

3.3.3. Donor–acceptor interactions. Since the goal was to make a list of unexpected compounds it was necessary to eliminate structures in which the $A\cdots B$ interactions (those between the different molecules) are more favorable than $A\cdots A$ and $B\cdots B$ interactions are likely to be.[10] Several examples of such eliminated molecule pairs are given in Fig. 5. The structures deleted for this reason all have intermolecular bonds $X-H\cdots Y$, with X and Y = O or N.

[10] Donor–acceptor interactions were inferred if the nonbonded atom–atom distance is at least 0.02 Å shorter than the sum of the van der Waals radii. A few slightly longer distances were noted if the interactions seemed to matter. Distances that seemed less important (*e.g.* somewhat short $H\cdots H$ contacts in a structure determined near 100 K) were ignored.

Structures were, however, retained if either the $A\cdots A$ or the $B\cdots B$ interactions were expected to be about as favorable as the $A\cdots B$ interactions. Consider the molecule pairs shown in Fig. 6. The compound of triethanolamine and triethanolamine oxide (LUDDOD; Kemmitt *et al.*, 2002) was not expected because the oxidation of the commonly used $N(CH_2CH_2OH)_3$ ligand was not anticipated and because in the structure of pure triethanolamine (BAFTAD10; Mootz *et al.*, 1989) all hydroxyl groups act as both hydrogen-bond donors and acceptors. Furthermore, a structure of the pure oxidized molecule can be imagined in which there are columns of molecules in which the three hydroxyl groups of one molecule form hydrogen bonds to the NO group of the next (as they do in LUDDOD). The co-crystal was therefore unexpected.

The cases of NIZQES (Comba *et al.*, 1997) and PIVDOO (Yao *et al.*, 2007), both of which also include a molecule and its oxidation product (see Fig. 6), are similar. The oxidation creates a better acceptor group in *B* than is available in *A* so that hydrogen bonds in the *A*–*B* co-crystal are expected to be better than those in pure *A* but similar to those in pure *B*. The co-crystal would not, however, have been predictable because the $A\cdots B$ interactions are no better than the $B\cdots B$ interactions.

The skeletal isomers benzamide and (*E*)-benzaldehyde oxime (JUKJOO; see Fig. 6) were synthesized (Maurin *et al.*, 1993) with the idea that they might form a hydrogen-bonded co-crystal. On the other hand, the paper indicates that the compound was not easily obtained. Furthermore, crystals of the pure amide (BZAMID) are almost certainly slightly denser than those of JUKJOO[11] and the hydrogen bonds in the amide seem to be better than in the co-crystal (chains of linked dimers are found in both). JUKJOO was therefore retained in the list. A crystal of the pure oxime (CAHDAQ; Jerslev, 1983) is a few percent less dense than the co-crystal and contains tetramers held together by $OH\cdots N$ bonds.

Deciding whether a co-crystal was engineered, is an *a posteriori* rationalized result of crystallization experiments or was serendipitous was sometimes a difficult call, but only a few of the co-crystals of diastereomers [*i.e.* KIGDOT (Benedetti *et al.*, 1990) and XADVED (Dalhus & Görbitz, 2000; see Fig. 2)] seem to have been designed. On the other hand, most of the compounds that are quasiracemates must be regarded as designed because they were made intentionally.

3.4. Spreadsheet

A spreadsheet available with the supplementary material includes the following:

(i) refcode, year of publication, temperature of determination, conventional *R* factor, space group, Z' for the compound (usually 1);

(ii) compound description, ratio of chemical components and their chemical formulae, difference between the two units in number of non-H atoms, number of chiral centers if components are configuratonal diastereomers, identity of any included solvent;

(iii) some information about imposed symmetry and any disorder;

(iv) some information about any hydrogen bonds in the structure;

(v) additional comments;

(vi) classification of the structure (see Table 1).

We were pleased to discover that in almost all cases the original authors had recognized and reported that the structure contains two different molecules.

4. Results

4.1. Unexpected compounds

The number of unexpected co-crystals of isomers (see Table 1) is modest but significant. Major failures of fractional crystallization are rare but do occur.

Of the 270 co-crystals viewed as at least largely unexpected 167 have exactly the same number of atoms. The 167 include 134 co-crystals of diastereomers (categories *D* and *R*), 26 of skeletal isomers (*K*) and 7 of axial/equatorial conformers (*C*). There are then 270 − 134 − 26 − 7 − 19(other) = 84 co-crystals that are not in the *D*, *R*, *K*, *C* (all isomers) or *O* (other) categories, but that are in one or more of the categories *DA*, *KA*, *SA* and *Q* (*i.e.* almost isomers). Of those 84, 20 have the same number, but not necessarily exactly the same types, of non-H atoms. There are 36 more for which the difference is one (*e.g.* H *versus* Me or OH), and 14 for which the difference is 2. In eight more the difference is 3, in four it is 4 and in two it is 5. These counts do not include the O category because the differences in molecule sizes in that group are larger.

Most (58%) of the co-crystals in the list are composed of either configurational diastereomers or quasienantiomers or of molecules that fit both descriptions.

While 62% of the 114 co-crystals of configurational diastereomers (*D*) can be considered to be co-crystals of quasienantiomers (*Q*), the rest clearly cannot. The same percentage of the 114 quasiracemates are compounds of diastereomers. There is substantial overlap of the groups *D* and *Q*, but 38% of each group is not part of the other.

Of the 114 quasiracemates all but one have approximate inversion symmetry. The one that does not (MIYGAC; Kooijman *et al.*, 2002; see Fig. 7) is a quasiracemate composed of diastereomers that has approximate, local glide operations only.

The formation of compounds of configurational diastereomers seems to be particularly likely if a 1,1 switch of a Me group and an H atom in one molecule would make the pair the same (10 compounds) or enantiomers [26 compounds including XADVED (see Fig. 2) and MIYGAC (see Fig. 7)]. The difference in size of the Me group and H atom is small enough that other considerations (*e.g.* the possibility of

[11] The molecular volume of BZAMID is 150.4 Å^3 at 123 K (Ruble & Galvao, 1995) and 152.0 Å^3 at 173 K (Kobayashi *et al.*, 2003), while the molecular volume of JUKJOO is 156.1 Å^3 at 223 K (Maurin *et al.*, 1993).

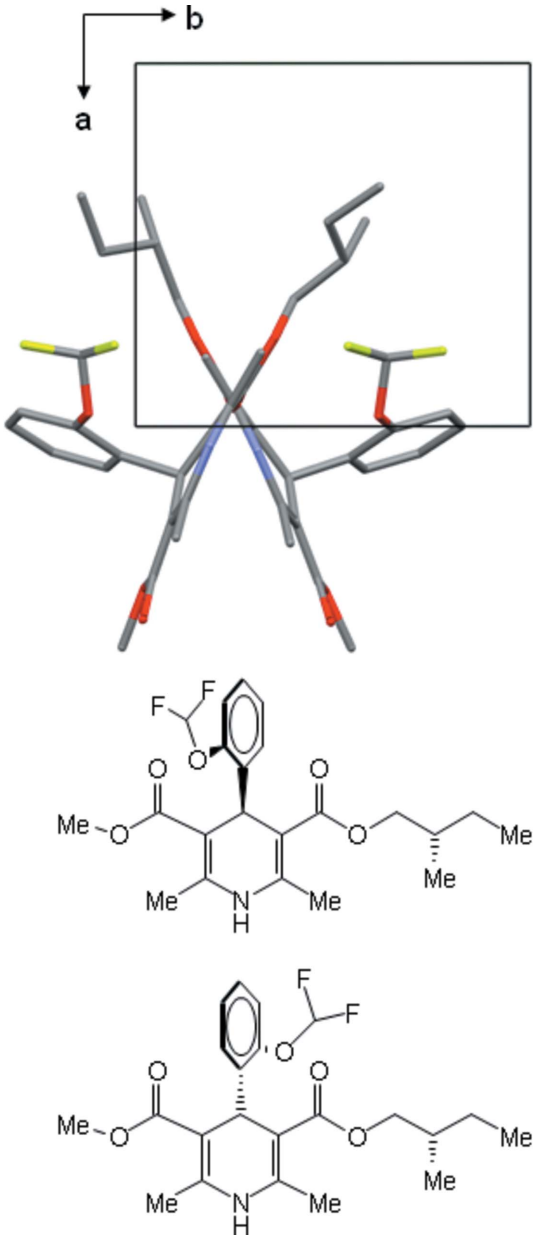

Figure 7
Projection down **c** of the structure of MIYGAC (Kooijman *et al.*, 2002; space group *P*1). The configurations of the 2-methylbutyl groups are both *S*, but the configuration at the other asymmetric C atom is *R* in one molecule and *S* in the other. The two molecules are related by a local, pseudo-*c*-glide operation perpendicular to **b**, but the 2-methylbutyl groups cannot be related by a glide.

approximate inversion symmetry) probably determine the solid phase found.[12] Another 12 compounds are formed from molecules that would be the same or enantiomeric if a 1,1 switch of an H atom and an OH or NH_3^+ group were made in one of them.

[12] Addadi *et al.* (1977) and Green *et al.* (1979) found that the H and Me parts of *sec*-butyl substituents are often disordered. The implication is that an enantiomerically pure *sec*-butyl group in each of an otherwise racemic pair of molecules is unlikely to be an effective enough chiral discriminator to allow separation by fractional crystallization.

Table 3
Space-group statistics for the co-crystals included in the final list.

Space group	Number	Percentage (%)
All quasiracemates (114 total)		
#1 (*P*1)	33	29
#4 ($P2_1$)	60	53
#5 (*C*2)	11	10
#18 ($P2_12_12$)	1	1
#19 ($P2_12_12_1$)	9	8
Total	114	
All other co-crystals (156 total)		
#1 (*P*1)	10	6
#2 ($P\bar{1}$)	49	31
#4 ($P2_1$)	24	15
#5 (*C*2)	5	3
#7 (*Pc*)	2	1
#9 (*Cc*)	1	1
#14 ($P2_1/c$)	41	26
#15 (*C*2/*c*)	7	4
#18 ($P2_12_12$)	2	1
#19 ($P2_12_12_1$)	10	6
#29 ($Pca2_1$)	1	1
#80 ($I4_1$)	2	1
#96 ($P4_32_12$)	1	1
#146 (*R*3)	1	1
Total	156	

There are 14 compounds of molecules that would be enantiomers if a [2.2.1] (*e.g.* camphor or bornyl) or [2.2.2] cage in one of the molecules were inverted. The van der Waals surfaces of these cage substituents are roughly spherical so that the difference between the two absolute configurations is not important sterically. In these 14 structures the two diastereomers are very nearly enantiomers.

The percentage of compounds of separable skeletal isomers and near isomers (14%) is relatively small, and is even smaller (11%) if the skeletal isomers that are also quasiracemates are not counted.

Most rare are the double salts associated with failed diastereomeric resolutions (13; 5%).

Hydrogen bonds seem to favor the formation of co-crystals; in 70% of the structures there is at least one bond $X-H\cdots Y$, where X and Y are either N or O. If short $C-H\cdots O{=}$ interactions are also counted the percentage rises to 85% (94% for the quasiracemates).

The overall percentage of co-crystals that seem to have been designed is 23% (63 structures), but all but two of the 44 quasiracemates that are not diastereomers (classes *Q*, *QK*, and *Q KA*) were made intentionally. The number of co-crystals that are natural products (with both components usually having been isolated together) is 20, of which 13 are quasiracemates, 11 of which are composed of pairs of diastereomers.

4.2. Space-group statistics

The space-group counts for the 114 quasiracemates (see Table 3) show that the two most common groups are *P*1 and $P2_1$. This result is no surprise because adding an inversion center to these groups gives $P\bar{1}$ and $P2_1/c$, which are the two most frequent groups overall. The ratio of quasiracemate

structures in $P2_1$ to $P1$ is 1.8, which is very similar to the corresponding ratio (1.9) of $P2_1/c$ to $P\bar{1}$ for all structures in the CSD (Brock & Dunitz, 1994). In this way the quasiracemates also mimic true racemic compounds.

The statistics for the remaining 156 structures are complicated by the fact that the fraction of these structures in Sohnke space groups (35%) is higher than for the CSD as a whole (18%) because many of the crystals are composed of enantiomerically pure material (as are most of the compounds of diastereomers). It is possible, however, to compare the ratios of structures in non-Sohnke space groups. In the CSD as a whole $P2_1/c$ is more frequent than $P\bar{1}$ by a factor of 1.9, but in this list of co-crystals there are more structures in $P\bar{1}$ than in $P2_1/c$. The total number of molecules in the unit cell may be a factor; in a $P\bar{1}$ co-crystal that number is four, which is the same as in a $P2_1/c$ structure of a pure single-component material. Brock & Dunitz (1994) found that there are rarely more than four orientations of the molecular inertia tensor, which is unchanged by inversion, and in most structures there are no more than two.

4.3. 'Predictable' compounds

More than half of the 'expected' compounds (*i.e.* those that are not included in the final list; see Table 2) include hydrogen bonds.

Among the compounds excluded from the final list are several types of co-crystals worth attention. Co-crystals containing conjugate acid–base pairs (*e.g.* *HA* plus Na^+A^- or an amine *B* plus an ammonium salt BH^+Cl^-) occur frequently enough in the CSD that they must have a structural advantage – perhaps the increased hydrogen-bond strentgh of the $HA\cdots A^-$ and $BH^+\cdots B$ bonds. Compounds that appear to contain both neutral molecules and ion pairs (*e.g.* IDUXEK, a polyiodide salt of a partially oxidized tetrathiofulvalene derivative; Iyoda *et al.*, 2001) are also notable. Among the tautomers there are a number of examples of co-crystallization of a molecule and its zwitterion (*e.g.* the orthorhombic form of *ortho*-aminobenzoic acid; AMBACO07; Brown & Ehrenberg, 1985).

5. Discussion

5.1. Basis of fractional crystallization

Fractional crystallization is reliable because ordered crystals are usually more energetically favorable than disordered crystals in which two different molecules occupy the same site, and because crystals with smaller asymmetric units are usually more favorable than crystals with larger asymmetric units. These trends are a result of the general correlation of greater density with lower energy (Dunitz *et al.*, 2000).[13]

If two different molecules occupy, on average, the same site then the unit cell must be large enough to accommodate the larger molecule as well as the smaller molecule; the packing efficiency is therefore necessarily lowered. The increased entropy associated with the disorder does not usually lower the free energy enough to offset the unfavorable energy change.

On the other hand, if two different molecules form an ordered co-crystal with a larger unit cell (approximately doubled in the case of a 1:1 co-crystal) then the number of different nonbonded contacts is also increased. The larger the number of different nonbonded contacts there are in a crystal the more difficult it is to optimize them. If there are two molecules of different types in the unit cell, or if there are two (or more) molecules of one type in quite different orientations, then there is certain to be at least one direction (and probably many directions) along which the molecules have different lengths. If the lengths are different then their optimum spacings are different. The possibility that one optimum spacing might be an almost integral multiple of the other is low. Consequently, minimizing crystal energy normally means minimizing the size of the asymmetric unit (see also Section 6.2 of Lloyd *et al.*, 2007).[14]

The tendency for the size of the asymmetric unit to be minimized probably explains why diastereomic resolutions are usually successful. While the difference in the way the enantiomers interact with the resolving counterion may explain *which* diastereomeric salt is less soluble, the general success of the method is most likely a consequence of the tendency for the asymmetric unit to be as small as possible.

5.2. Why spontaneous resolution is rare

Unlike other co-crystals, co-crystals formed from a pair of separable enantiomers are the rule rather than the exception. The resulting racemic compounds are different from other co-crystals because if the enantiomers are symmetry related then compound formation does not raise the number of intermolecular contacts that must be optimized. Furthermore, the possibility of improper symmetry greatly increases the possible range of mutual orientations of adjacent molecules. In a Sohnke group that has only proper symmetry elements the adjacent molecules in a $Z' = 1$ structure must be related by a rotation of $2\pi/n$, where $n = 1, 2, 3$ or 6. This restriction is very strong indeed (see also Section 4 of Patrick & Brock, 2006), and may explain why $Z' > 1$ structures are so much more frequent in Sohnke groups than in non-Sohnke groups (Brock & Dunitz, 1994).[15] Improper symmetry is so favorable that in many $Z' > 2$ structures of pure enantiomers the independent

[13] While this generalization is most applicable to polymorphs (*i.e.* to different structures of the same molecule), the molecules that make up the co-crystals considered here are so similar that we expect the generalization to hold. Exceptions to the correlation of higher density with lower energy often involve strong donor–acceptor interactions like hydrogen bonds.

[14] The entropy differences between ordered co-crystals and their pure components are seldom important because the entropy of all perfectly ordered crystals is 0 at $T = 0$ K, because the number of normal modes is the same for both sets of crystals, and because the entropy of mixing macroscopic crystals is insignificant (see Brock *et al.*, 1991).

[15] It might seem that the number of apparent $Z' > 1$, Sohnke-group structures identified in this work as actually being co-crystals of configurational diastereomers is large enough that previously published frequencies of $Z' > 1$ structures in Sohnke groups need revision. Brock & Dunitz (1994) counted 93 of 220 $P1$ structures and 272 of 1628 $P2_1$ structures as having $Z' > 1$. We find 32 $P1$ structures and 52 $P2_1$ structures that are co-crystals of diastereomers, but only 4 and 7 of those structures were published by the time the searches reported for that 1994 publication were completed.

molecules are related by approximate inversion (or possibly glide) symmetry (see page 932 of Marsh, 1999).

5.3. The existence of quasiracemates

Quasiracemates exist because the advantage of approximate inversion symmetry (or, much more rarely, approximate glide symmetry) is considerably more important than the disadvantage of doubling the size of the asymmetric unit. The advantages of removing the requirement that molecules be related by rotations of $2\pi/n$, n = 1, 2, 3 or 6, are more important than the disadvantages of raising the number of different intermolecular contacts. The experiments of Wheeler with carboxylic acids and amides (Wheeler *et al.*, 2008; Breen *et al.*, 2008; Lineberry *et al.*, 2008, and references therein), and those of Görbitz with amino acids (Dalhus & Görbitz, 1999*a*,*b*,*c*, 2000; Görbitz *et al.*, 2009) suggest that single crystals of quasiracemates are not difficult to grow, at least as long as the constituent molecules form good hydrogen bonds. Some inorganic examples have been characterized by Englert (Englert *et al.*, 2002; Calmuschi *et al.*, 2004, and references therein).

5.4. The role of hydrogen bonding

Hydrogen bonds are well known to generally promote the formation of co-crystals, many of which are composed of hydrogen-bond donors and acceptors (see Table 2). Hydrogen bonding seems to be even more important to the formation of co-crystals composed of molecules that are isomers, near isomers or nearly the same, perhaps partly because few of the molecule pairs form stacking complexes.

Hydrogen bonds are especially important in co-crystals because the intermolecular bonds are so adjustable; the distances and angles of intermolecular hydrogen bonds can be varied quite a lot at a small energy cost. The hydrogen-bonding interactions then simplify the problem of optimizing the much larger number of interatomic distances in a co-crystal. Almost all the quasiracemates we found that were designed intentionally contain either a carboxylic acid or an amide group. The two quasiracemates that do not are dialcohols.

The structure of VASKUV (Yurdakul *et al.*, 1998; Fig. 8) is illustrative. The presence of the two diastereomers in the centrosymmetric ($P\overline{1}$) crystal allows the formation of hydrogen-bonded tetamers around inversion centers even while the two different molecules are related by a pseudo-translation along **a**. Clearly the two diastereomeric pairs of enantiomers form a better structure than either could by itself. It is therefore no surprise that neither diastereomer is found in any other structure in the CSD. The case of RIHLUQ (Fig. 1) is similar although structures of the two pure components are both known (see Lloyd *et al.*, 2007).

Consider compound *A* that crystallizes poorly. If a closely related impurity *B* that can facilitate hydrogen bonding is present in a significant amount at the time of attempted crystallization of *A* then the formation of at least a small amount of a co-crystal is easily explained. If the desired compound and impurity are very similar then the presence of two components does not greatly complicate the problem of optimizing spacings. If the $A\cdots B$ hydrogen bonds are better than the $A\cdots A$ (but not necessarily the $B\cdots B$) bonds the donor–acceptor interactions favor a co-crystal including the impurity.[16] Understanding compound formation, however, is not the same as predicting it. Fractional crystallization is usually effective, even when an alcohol is crystallized in the presence of an oxidation product containing a carbonyl group or when an ether is crystallized in the presence of a hydrolysis product that includes an hydroxyl group.

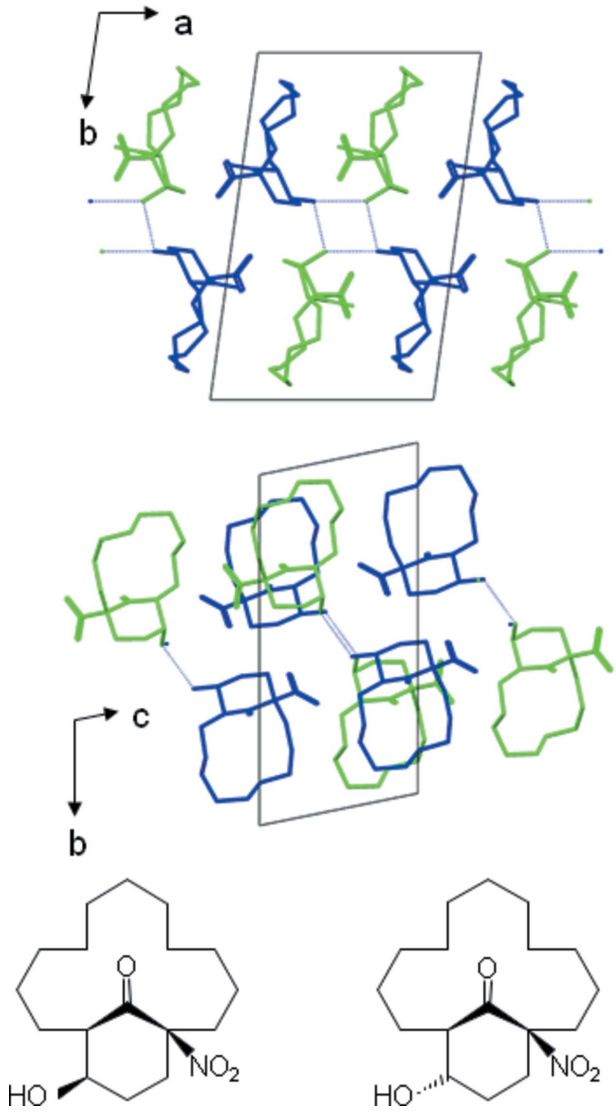

Figure 8
Crystal packing in the structure of VASKUV (Yurdakul *et al.*, 1998; space group $P\overline{1}$). The presence of the second diastereomer allows the formation of hydrogen-bonded tetramers in the presence of pseudotranslation. Projections along **c** and **a** are shown.

[16] In the case of an impurity it sometimes happens that crystals of both the pure component and the co-crystal are obtained. If the co-crystals are better formed then one of them may be chosen for structure determination even if the co-crystal is the minor component. See, *e.g.*, Brock *et al.* (1992).

The final list includes fewer than 20 co-crystals in which the functional groups of the two components can be considered to be complementary.

5.5. Solution concentration *versus* solid composition

Some authors have expressed surprise that co-crystals of one composition (almost always 1:1 or 1:2) grew from a solution of very different composition. While the composition of the solution does matter, the compositions of the eutectics of the pure materials and their co-crystal are equally important (see Fig. 9). If the co-crystal is ordered, it will be the equilibrium product if and only if the ratio of the two components of the material from which the crystal is grown is between the compositions of the two adjacent eutectics (Fig. 9*a*) or between the compositions of a peritectic and a eutectic (Fig. 9*b*). Fractional crystallization is usually successful because even if there is a stable stoichiometric compound of the desired material and one of its impurities, the composition of the eutectic of the pure material and the co-crystal usually corresponds to a higher level of the impurity than is likely to be encountered. There are, however, striking counterexamples (see, *e.g.*, Brock *et al.*, 1992). Also, the crystal form found is sometimes governed by kinetic as well as thermodynamic factors.

5.6. One- and two-component systems

If the two isomers interconvert faster than crystals grow then their ratio in solution (or the melt) is fixed and a 1:1 'co-crystal' is really a $Z' = 2$ structure of a single component. We see no reason, however, to exclude these structures from the count, at least as long as no comparison is made of melting points, heats of fusion or densities of the co-crystals and their pure components (see Brock *et al.*, 1991). We think inclusion is appropriate because it is unusual for the independent molecules in a $Z' > 1$ structure to have different conformations and/or orientations as they do in the examples listed here; it is much more common for the independent molecules to have such similar orientations and conformations that small displacements would give a unit cell of smaller volume and/or higher symmetry. Again, the more different kinds of intermolecular spacings there are (and the larger those differences are) the more difficult the problem of filling space densely.

The final list contains only a few $Z' > 1$ structures containing conformers, all of which are separated by significant energy barriers, but there are undoubtedly many more co-crystals described in the CSD containing conformers that interconvert more readily. The structure of PINCOL (2,3-dimethyl-2,3-butanediol; Jeffrey & Robbins, 1978), for example, includes both *trans* and *gauche* conformers so that a good pattern of hydrogen bonds can be formed. It would, however, have been impractical, as well as outside the focus of this study, to try to search for all such examples.

5.7. Order in the absence of inversion symmetry

In the case of the quasiracemates the crystalline order is no surprise because any disorder would be whole-molecule disorder of enantiomers. Such disorder is rare and would have, in any case, caused the co-crystal to be excluded from the list.

Order in the structures of pairs of molecules that are almost the same (category SA) is more of a surprise but it does occur. Although the two molecules in KAYNED (Trikha *et al.*, 1990; see Fig. 10) are very nearly the same, the N—C—C═O torsion angles in the two molecules differ by 167° so that they are clearly different. In KEMYIK (Soriano-Garcia *et al.*, 1989) the two molecular conformations are essentially the same but there are no indications of any short contacts or voids in the structure and there is no pseudosymmetry. In RALYOT (Sarmah *et al.*, 2005) there is an extra methyl group in one molecule. Pairs of molecules in the *Cc* unit cell are related by pseudo-inversion and by pseudo-twofold rotation, but the two types of pseudosymmetry elements are separated by 0.246 and 0.254 along **a** (a difference of 0.22 Å). In TIVMIU [$Ph_2As(=X)(OH)$, X = O, S; Silaghi-Dumitrescu *et al.*, 1996] the difference in hydrogen-bonding ability of the O and S atoms is sufficient to produce ordering.

Reports of structures of co-crystals in which there is no obvious reason for the existence of the compound are, however, always worth critical examination. Consider

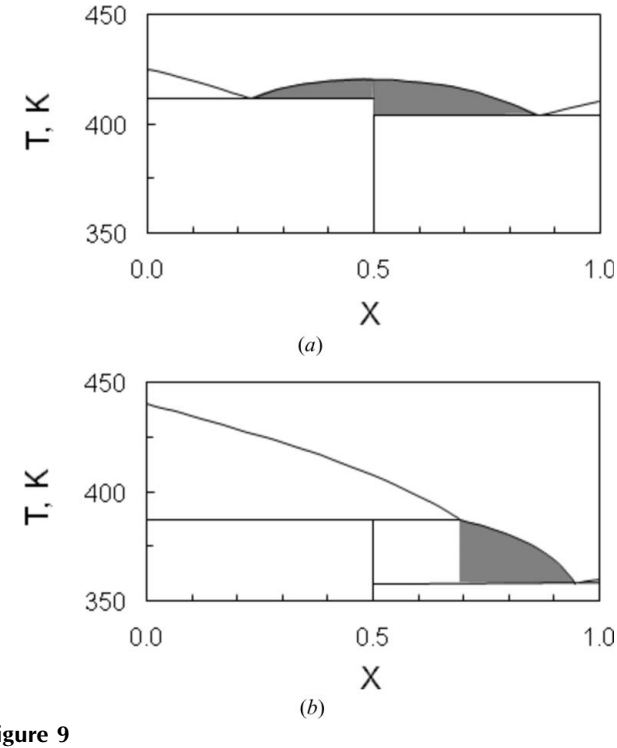

Figure 9
Schematic solid–liquid phase diagrams drawn assuming the components act ideally and have unexceptional heats of fusion. These curves are also indicative of what happens when crystals are grown from a two-solute solution. (*a*) The melting points of the pure components are 425 and 410 K; the melting point of the 1:1 compound is 420 K. Melt compositions above the gray regions ($0.23 < X < 0.87$) would deposit crystals of the 1:1 compound first. (*b*) The melting points of the pure components are 440 and 360 K. The compound melting point of 390 K is below the freezing-point depression curve of the higher-melting component; the freezing-point depression curve of the co-crystal does not become important until $X = 0.70$. Under equilibrium conditions crystals of the compound are deposited in the region $0.70 < X < 0.95$.

LAPSAX (1:1 compound of $C_6F_5-N{=}CH-C_6H_5$ and $C_6F_5-CH{=}N-C_6H_5$; Zhu *et al.*, 2005). The C_6F_5 and C_6H_5 rings would be expected to stack alternately in the structures of the co-crystal and in both of its pure components; the only advantage of the co-crystal would be associated with the spatial separation of N atoms and CH groups. Examination of the displacement ellipsoids for LAPSAX, however, shows that in one molecule the ellipsoid for the N atom is suspiciously large while the ellipsoid for the adjacent C atom is somewhat small. At the very least disorder seems to have been overlooked. This structure was therefore excluded from the final list.

We also found that disorder of the two forms is common in 'co-crystals' of tautomers. The presence of two tautomeric forms is often necessary for the formation of a good set of hydrogen bonds, but if the hydrogen-bond network is not three-dimensional, which it usually is not, then there is often disorder.

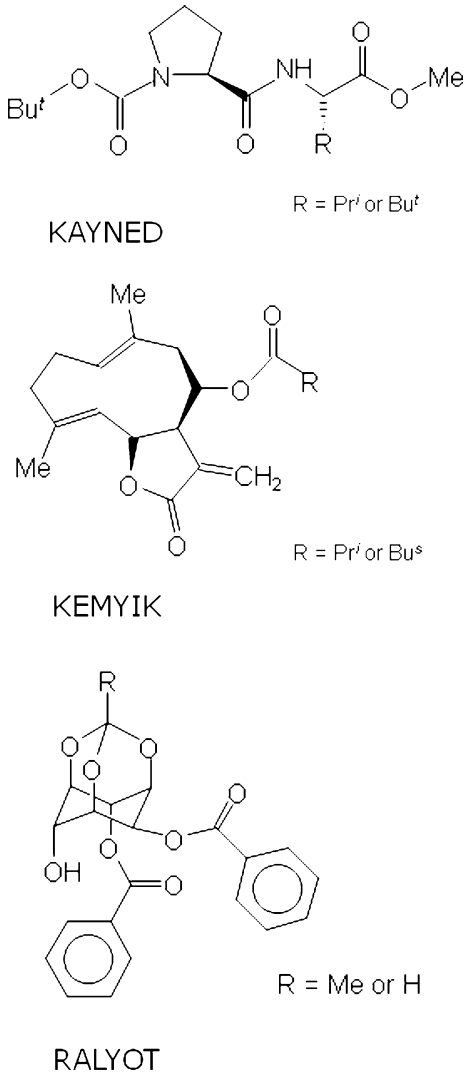

Figure 10
Examples of three co-crystals composed of molecules that are almost the same (see text for individual references).

5.8. Final comments

In some cases there is just no obvious reason why the two isomers co-crystallize. The co-crystal of *cis* and *trans* isomers shown in Fig. 11 (RIZJUG; Wasserman *et al.*, 2008) is an example. Fig. 12 shows chemical line drawings for the components of three additional surprising co-crystals. These compounds demonstrate just how unpredictable crystallization can be.

Papers describing several structures not included in the final list are well worth reading. They include reports for KILJOE (Carman *et al.*, 1999), a disordered co-crystal that has a melting point slightly higher than those of the pure isomers, which cannot be separated by recrystallization or chromatography, PDTOMS11 (Wong-Ng *et al.*, 1984), the disordered form of a co-crystal of two natural products related by hydrogenation of a double bond (the ordered form, PDTOMS10, is included in the final list), and HBTBBZ (Gafner & Herbstein, 1964), a co-crystal of C_6Br_6 and 1,2,4,5-tetrabromobenzene.

6. Summary

A list of 270 compounds of ordered co-crystals of isomers, near isomers and molecules that are almost the same has been compiled. Compounds of true isomers were located semi-automatically by comparing the IUPAC International Chemical Identifier (InChI™) strings generated from the information in the CSD entries.

Structures of co-crystals of isomers are unusual because the functional groups of the two components are usually the same while the molecules that form co-crystals usually have

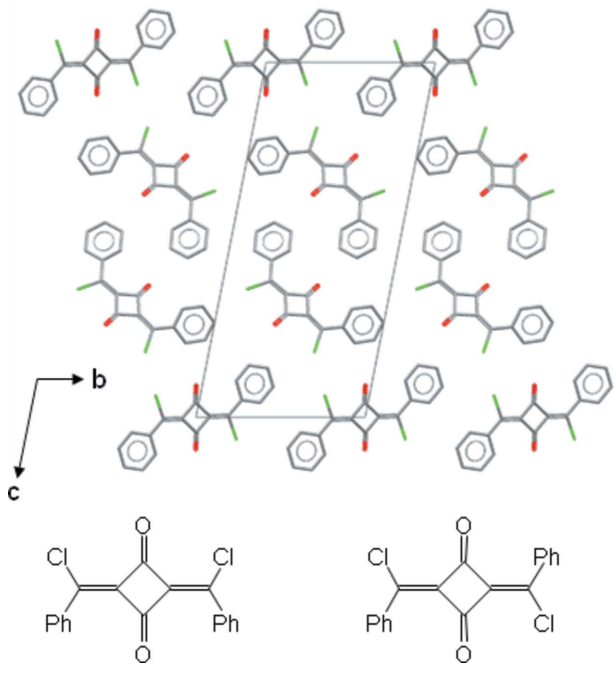

Figure 11
Projection along **a** (length 3.80 Å) of RIZJUG (Wasserman *et al.*, 2008; space group $P\bar{1}$).

complementary functional groups. In the absence of $A\cdots B$ interactions that are more attractive than the average of the $A\cdots A$ and $B\cdots B$ interactions the filling of space is the primary consideration. Co-crystals of isomers are unlikely because dense crystal packing is usually associated with minimization of the size of an ordered asymmetric unit.

The fact that crystals used for structure determinations are usually grown from reasonably pure, single-component material is also important. If more structures of crystals grown from multi-component solutions had been determined there would be more structures of co-crystals in the CSD.

The large number (114) of quasiracemates in the list of compounds is additional strong evidence that symmetry operations of the second kind, and especially inversion centers, are favorable for crystal packing. Crystals grown from a solution of quasienantiomers are likely to be quasiracemates.

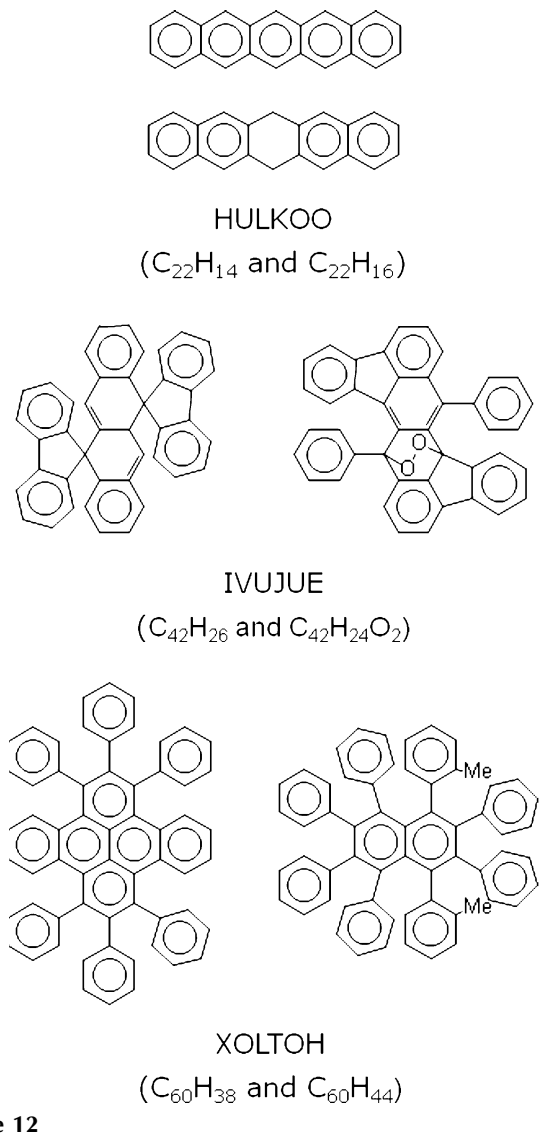

Figure 12
Examples of three unexpected co-crystals of near isomers: HULKOO (Mattheus *et al.*, 2002), IVUJUE (Harrington *et al.*, 2004) and XOLTOH (Lu *et al.*, 2002).

The pseudosymmetry in a quasiracemate compensates for the packing problems associated with the larger asymmetric unit.

Compounds of configurational diastereomers are surprisingly common. If it appears that $Z' > 1$ and if the molecule contains several chiral centers then the possibility of co-crystal formation must always be considered. If two diastereomers would be enantiomers but for the exchange of an H atom and a methyl, hydroxyl or amino substituent or but for the inversion of a [2.2.1] or [2.2.2] cage then compound formation is so likely that it should be considered predictable.

Intermolecular hydrogen bonding increases the chance of compound formation because the interactions are both attractive and adjustable. Among isomers, near isomers and molecules that are almost the same the likelihood of obtaining an $A\cdots B$ co-crystal during fractional crystallization of A is enhanced if an otherwise very similar impurity B includes functional groups such that $A\cdots B$ hydrogen bonds are more attractive than the $A\cdots A$ interactions. That said, this study shows there are more failures of fractional crystallization that cannot be explained simply than there are failures that seem to result from better hydrogen bonding.

We are grateful to Dr Frank Allen of the CCDC for his enthusiasm and good advice. László Fábián thanks Pfizer for funding *via* the Pfizer Institute for Pharmaceutical Materials Science. Steven P. Kelley was supported under the NSF REU Site program grant CHE-0552247.

References

Addadi, L., Gati, E., Lahav, M. & Leiserowitz, L. (1977). *Isr. J. Chem.* **15**, 116–123.
Albano, V. G., Bellon, P. L. & Sansoni, M. (1969). *Chem. Commun.* pp. 899–901.
Allen, F. H. (2002). *Acta Cryst.* B**58**, 380–388.
Ansari, F., Fedorchuk, C., Parvez, M., Umbreen, S. & Saghir, S. (2002). *Acta Cryst.* E**58**, o422–o425.
Benedetti, E., Blasio, B. D., Pavone, V., Pedone, C., Fuller, W. D., Mierke, D. F. & Goodman, M. (1990). *J. Am. Chem. Soc.* **112**, 8909–8912.
Bilton, C., Howard, J. A. K., Madhavi, N. N. L., Nangia, A., Desiraju, G. R., Allen, F. H. & Wilson, C. C. (1999). *Chem. Commun.* pp. 1675–1676.
Breen, M. E., Tameze, S. L., Dougherty, W. G., Kassel, W. S. & Wheeler, K. A. (2008). *Cryst. Growth Des.* **8**, 3863–3870.
Brock, C. P. & Dunitz, J. D. (1994). *Chem. Mater.* **6**, 1118–1127.
Brock, C. P., Kwiatkowski, S., Watt, D. S. & Sayed, A. (1992). *Acta Cryst.* B**48**, 719–725.
Brock, C. P., Schweizer, W. B. & Dunitz, J. D. (1991). *J. Am. Chem. Soc.* **113**, 9811–9820.
Brown, C. J. & Ehrenberg, M. (1985). *Acta Cryst.* C**41**, 441–443.
Cai, J.-W., Hu, X.-P., Chen, C.-H. & Ji, L.-N. (2001). *Acta Cryst.* C**57**, 394–396.
Calmuschi, B., Alesi, M. & Englert, U. (2004). *J. Chem. Soc. Dalton Trans.* pp. 1852–1857.
Carman, R. M., Kennard, C. H. L., Venzke, B. N. & Smith, G. (1999). *Aust. J. Chem.* **52**, 329–332.
Comba, P., Fath, A., Kuhner, A. & Nuber, B. (1997). *J. Chem. Soc. Dalton Trans.* pp. 1889–1898.
Cruz-Cabeza, A. J. & Groom, C. R. (2011). *CrystEngComm*, doi:10.1039/C0CE00123F.

Czaplik, W. M., Neudörfli, J.-N. & von Wangelin, A. J. (2007). *Green Chem.* **9**, 1163–1165.
Dalhus, B. & Görbitz, C. H. (1999*a*). *Acta Cryst.* B**55**, 424–431.
Dalhus, B. & Görbitz, C. H. (1999*b*). *Acta Cryst.* C**55**, 1105–1112.
Dalhus, B. & Görbitz, C. H. (1999*c*). *Acta Cryst.* C**55**, 1547–1555.
Dalhus, B. & Görbitz, C. H. (2000). *Acta Cryst.* B**56**, 720–727.
Dunitz, J. D., Filippini, G. & Gavezzotti, A. (2000). *Helv. Chim. Acta*, **83**, 2317–2335.
Englert, U., Haering, A., Hu, C. & Kalf, I. (2002). *Z. Anorg. Allg. Chem.* pp. 1173–1179.
Ermer, O., Vincent, B. R. & Dunitz, J. D. (1989). *Isr. J. Chem.* **29**, 137–142.
Etter, M. C. (1990). *Acc. Chem. Res.* **23**, 120–126.
Etter, M. C. (1991). *J. Phys. Chem.* **95**, 4601–4610.
Fernandes, M. A. & Levendis, D. C. (2004). *Acta Cryst.* B**60**, 315–324.
Frohberg, P., Wagner, C., Meier, R. & Sippl, W. (2006). *Tetrahedron*, **62**, 6050–6060.
Gafner, G. & Herbstein, F. H. (1964). *J. Chem. Soc.* pp. 5290–5302.
Glen, R. C., Murray-Rust, P., Riddell, F. G., Newton, R. F. & Kay, P. B. (1982). *J. Chem. Soc. Chem. Commun.* pp. 25–26.
Görbitz, C. H. & Hersleth, H.-P. (2000). *Acta Cryst.* B**56**, 526–534.
Görbitz, C. H., Rissanen, K., Valkonen, A. & Husabø, Å. (2009). *Acta Cryst.* C**65**, o267–o272.
Green, B. S., Lahav, M. & Rabinovich, D. (1979). *Acc. Chem. Res.* **12**, 191–197.
Gültekin, Z. & Hökelek, T. (2006). *Anal. Sci.* **22**, x9–x10.
Harrington, L. E., Britten, J. F. & McGlinchey, M. J. (2004). *Org. Lett.* **6**, 787–790.
Herbstein, F. H. (2005). *Crystalline Molecular Complexes and Compounds*, 1st ed. Oxford University Press.
Iyoda, M., Hasegawa, M., Kuwatani, Y., Nishikawa, H., Fukami, K., Nagase, S. & Yamamoto, G. (2001). *Chem. Lett.* pp. 1146–1147.
Jeffrey, G. A. & Robbins, A. (1978). *Acta Cryst.* B**34**, 3817–3820.
Jerslev, B. (1983). *Acta Cryst.* C**39**, 1447–1454.
Jones, R., Scheffer, J. R., Trotter, J. & Yang, J. (1994). *Acta Cryst.* B**50**, 601–607.
Kemmitt, T., Gainsford, G. J., Steel, P. J. & Wikaira, J. (2002). *Acta Cryst.* E**58**, o851–o852.
Kobayashi, K., Sato, A., Sakamoto, S. & Yamaguchi, K. (2003). *J. Am. Chem. Soc.* **125**, 3035–3045.
Kooijman, H., Spek, A. L., Sobolev, A., Jongejan, H. & Franssen, M. C. R. (2002). *Acta Cryst.* E**58**, o532–o534.
Lancaster, R. W., Karamertzanis, P. G., Hulme, A. T., Tocher, D. A., Lewis, T. C. & Price, S. L. (2007). *J. Pharm. Sci.* **96**, 3419–3431.
Lineberry, A. M., Benjamin, E. T., Davis, R. E., Kassel, W. S. & Wheeler, K. A. (2008). *Cryst. Growth Des.* **8**, 612–619.
Lloyd, M. A., Patterson, G. E., Simpson, G. H., Duncan, L. L., King, D. P., Fu, Y., Patrick, B. O., Parkin, S. & Brock, C. P. (2007). *Acta Cryst.* B**63**, 433–447.
Loehlin, J. H., Lee, M. & Woo, C. M. (2008). *Acta Cryst.* B**64**, 583–588.
Lu, J., Zhang, J., Shen, X., Ho, D. M. & Pascal, R. A. Jr (2002). *J. Am. Chem. Soc.* **124**, 8035–8041.
Macrae, C. F., Bruno, I. J., Chisholm, J. A., Edgington, P. R., McCabe, P., Pidcock, E., Rodriguez-Monge, L., Taylor, R., van de Streek, J. & Wood, P. A. (2008). *J. Appl. Cryst.* **41**, 466–470.
Manoj, K., Gonnade, R. G., Bhadbhade, M. M. & Shashidhar, M. S. (2006). *Cryst. Growth Des.* **6**, 1485–1492.
Marsh, R. E. (1999). *Acta Cryst.* B**55**, 931–936.
Mattheus, C. C., Baas, J., Meetsma, A., Boer, J. L. de, Kloc, C., Siegrist, T. & Palstra, T. T. M. (2002). *Acta Cryst.* E**58**, o1229–o1231.
Maurin, J. K., Winnicka-Maurin, M., Paul, I. C. & Curtin, D. Y. (1993). *Acta Cryst.* B**49**, 90–96.
McNaught, A. D. & Wilkinson, A. (1997). *IUPAC Compendium of Chemical Terminology*, 2nd ed. Oxford: Blackwell Science.
Mootz, D., Brodalla, D. & Wiebcke, M. (1989). *Acta Cryst.* C**45**, 754–757.
Patrick, B. O. & Brock, C. P. (2006). *Acta Cryst.* B**62**, 488–497.
Plutecka, A., Rychlewska, U., Prusinowska, N. & Gawroński, J. (2010). *Acta Cryst.* B**66**, 678–686.
Poliakov, D. & Shevchenko, I. (2007). *Eur. J. Org. Chem.* pp. 2055–2057.
Ruble, J. R. & Galvao, A. (1995). *Acta Cryst.* B**51**, 835–838.
Ruysink, A. F. J. & Vos, A. (1974). *Acta Cryst.* B**30**, 1997–2002.
Sarmah, M. P., Gonnade, R. G., Shashidhar, M. S. & Bhadbhade, M. M. (2005). *Chem. Eur. J.* **11**, 2103–2110.
Sakurai, T. (1968). *Acta Cryst.* B**24**, 403–412.
Seiler, P. & Dunitz, J. D. (1989). *Helv. Chim. Acta*, **72**, 1125–1135.
Silaghi-Dumitrescu, L., Gibbons, M. N., Silaghi-Dumitrescu, I., Zukerman-Schpector, J., Haiduc, I. & Sowerby, D. B. (1996). *J. Organomet. Chem.* **517**, 101–106.
Soriano-Garcia, M., Guerrero, C. & Toscano, R. A. (1989). *Rev. Latinoam. Quim.* **20**, 1–4.
Stein, S. E., Heller, S. R. & Tchekhovskoi, D. (2003). *An Open Standard for Chemical Structure Representation: The IUPAC Chemical Identifier*, in Proceedings of the 2003 International Chemical Information Conference (Nimes), Infonortics, pp. 131–143; http://www.iupac.org/inchi.
Stein, S. E., Heller, S. R. & Tchekhovskoi, D. V. (2006). *The IUPAC Chemical Identifier – Technical Manual.* National Institute of Standards and Technology, Gaithersburg, Maryland, US 20899–8380.
Steiner, T. & Saenger, W. (1998). *J. Chem. Soc. Perkin Trans. 2*, pp. 371–377.
Stezowski, J. J., Stigler, R.-D., Karl, N. & Schuller, K. (1983). *Mol. Cryst. Liq. Cryst.* **94**, 243–253.
Trikha, J., Patel, H. C. & Singh, T. P. (1990). *Acta Cryst.* C**46**, 74–78.
Valente, E. J. & Schomaker, V. (1984). *Acta Cryst.* C**40**, 1068–1070.
Walker, M., Pohl, E., Herbst-Irmer, R., Gerlitz, M., Rohr, J. & Sheldrick, G. M. (1999). *Acta Cryst.* B**55**, 607–616.
Wasserman, H. H., Wasserman, E. R., Coats, S. J., Davis, R. E., Lynch, V. M. & Wiberg, K. B. (2008). *Tetrahedron Lett.* **49**, 2049–2051.
Wheeler, K. A., Grove, R. C., Davis, R. E. & Kassel, W. S. (2008). *Angew. Chem. Int. Ed. Engl.* **47**, 78–81.
Wong-Ng, W., Cheng, P.-T. & Nyburg, S. C. (1984). *Acta Cryst.* B**40**, 151–158.
Yao, H.-C., Li, M.-M., Yang, G.-S., Li, Z.-J. & Zhu, Y. (2007). *Inorg. Chim. Acta*, **360**, 3959–3964.
Yurdakul, A., Gurtner, C., Jung, E.-S., Lorenzi-Riatsch, A., Linden, A., Guggisberg, A., Bienz, S. & Hesse, M. (1998). *Helv. Chim. Acta*, **81**, 1373–1392.
Zhang, Q. & Curran, D. P. (2005). *Chem. Eur. J.* **11**, 4866–4880.
Zhu, S., Zhu, S., Jin, G. & Li, Z. (2005). *Tetrahedron Lett.* **46**, 2713–2716.
Zouev, I., Levy, T. & Kaftory, M. (2006). *Eur. J. Org. Chem.* pp. 4164–4169.