Acta Crystallographica Section B
Structural Science
ISSN 0108-7681

Simple model designed to generate new crystal structures derived from a mother phase; application to molecular compounds

Claire Gervais and Gérard Coquerel*

Unité de Croissance Cristalline, de Chromatographie et de Modélisation Moléculaire, UPRES EA 2659, IRCOF, Université de Rouen, 76821 Mont-Saint-Aignan CEDEX, France

Correspondence e-mail: gerard.coquerel@univ-rouen.fr

Received 8 February 2002
Accepted 20 May 2002

The basic principles of a model predicting new lattices from a known crystal structure are described. The first of the two-step procedure consists of extracting one- or two-dimensional periodic fragments (PF) from the mother structure. In the second step, symmetry operators are added to the PFs in order to generate one or several new three-dimensional lattices consistent with the 230 space groups. Most of the examples are related to polymorphism, but relationships between racemic compounds and enantiomers, twinning and lamellar epitaxy phenomena are also exemplified.

1. Introduction

The study of crystalline solids is of considerable interest in both theoretical and practical approaches because different crystal structures of a given compound have different physical and thermodynamical properties (*e.g.* in terms of solubility, dissolution rate, crystal morphology, compressibility or thermodynamic stability). The understanding of polymorphism (*i.e.* why a given molecule adopts a particular packing arrangement) and of the relationships between enantiomer and racemic compound structures could develop knowledge in various fields such as nucleation, crystal growth, molecular and supramolecular recognition, or stereoselective crystallization of chiral compounds. From a practical point of view, crystal packing studies are essential in the pharmaceutical industry (Brittain, 1999), particularly for bioavailability and process problems (Rustichelli *et al.*, 2000; Kobayashi *et al.*, 2000; Wöstheinrich & Schmidt, 2001), but also play an important role in other fields such as pigments (Erk, 2001) or explosives (Dzyabchenko *et al.*, 1996; Molchanova *et al.*, 1999). Threlfall (1995) and Vippagunta *et al.* (2001) review polymorphism precisely and describe many methods for the examination of crystalline solids.

The search for new crystal structures commonly uses one of the following three routes:

(i) *Ab initio* prediction (Karfunkel & Leusen, 1992; Leusen, 1996; Hofmann & Lengauer, 1997; Gdanitz, 1997; Chin *et al.*, 1999). This method is based on the modelling of molecular recognition mechanisms and consists of determining (by computation) the preferential crystal packings for a given molecule with a particular conformation.

(ii) Prediction of crystal structures *via* crystal engineering. The analysis of experimental data allows the construction of supramolecular synthons, which, regenerated in the three dimensions, lead to new crystal packings (Desiraju, 1997).

(iii) Experimental screening (Threlfall, 1995). Different conditions of nucleation and crystal growth or desolvation are tested in order to initiate the appearance of new forms.

Ab initio prediction of possible molecular crystal structures is constantly in progress. For instance, there exists now a commercial product called 'Polymorph Predictor' implemented in the *Cerius*^{2} (Molecular Simulations Inc., 1999) molecular modelling environment (for a review see Gdanitz, 1997). A Monte Carlo method for predicting crystal structures of semiflexible molecules, called KAP (Kitaigorodskii's Aufbau Principle), has been developed by Perlstein (1994). One-dimensional translation aggregates (Perlstein, 1992) and two-dimensional monolayer aggregates composed of semiflexible molecules (Perlstein, 1994) have been successfully generated. In the same way, Gavezzotti (1991, 1993: program *PROMET*) predicted crystal structures of rigid molecules starting from clusters composed of 2–4 molecules related by symmetries.

Despite these developments, *ab initio* predicting models based on molecular modelling still suffer from some limitations, especially for structures containing highly flexible molecules and/or with large solute. The main limitation lies in the nature of the force fields, which do not take into account the polarization effects of molecules. Consequently, *ab initio* prediction is definitely less efficient for highly polarized structures (*i.e.* zwitterions and salts). Moreover, approximately 10% of the known structures have $Z' > 1$ and one third (Threlfall, 1995) consist of solvates or even mixed solvates (Görbitz & Hersleth, 2000), which adds complexity to the problem. Despite continuous progress, some people are doubtful about the *ab initio* predictability of new crystal structures (Gavezzotti, 1994).

Different calculated and/or experimental crystal structures for a given compound have frequently similar lattice energies or heats of sublimation. Therefore, preference for particular crystal packing might be mainly driven by external factors such as supersaturation, temperature, solvent and/or impurities in solution. For instance, the arrangement of molecules in solution (*i.e.* pre-aggregation such as the formation of dimers of alcohols) and the resulting growth units may differ with the solvent. As shown for the four polymorphs of sulfathiazole (Blagden *et al.*, 1998), the presence of an impurity may inhibit the nucleation of the stable polymorphic form, allowing the crystallization of a metastable one. For all these reasons, even if *ab initio* methods can determine several polymorphs for a given compound, predicting which of these crystal structures will appear experimentally seems to be a difficult if not an impossible task.

A more practical approach to predict crystal structures is called crystal engineering (Desiraju, 1995). Supramolecular synthons, composed of molecules linked together by specific interactions (*e.g.* hydrogen bonds, ionic bonds, π–π interactions), are extracted from known structures and are exploited to generate crystal structures with different properties. Contrary to *ab initio* prediction for which the aim is to generate a precise packing for a given component, this approach is less precise and deals more with network prediction, *i.e.* generation of architectures for a family of compounds exhibiting the same features (Moulton & Zaworotko, 2001).

In the present paper a simple model of prediction of crystal structures derived from an experimentally known structure is proposed. After a description of this 'Derived Crystal Packings' (DCP) model, its analogies with the two predicting methods described above are discussed and some examples are given. The DCP model is thoroughly applied to a family of 5-alkyl-5-aryl-hydantoins, allowing the prediction of a new crystal structure for one of the derivatives. Several other examples deal with the applicability of the DCP model to conformational polymorphism, relationships between enantiomers and racemic compounds, twinning or lamellar epitaxy. Finally, various consequences of the structural relationships between two phases related by the DCP model are discussed.

2. Description of the DCP model: theoretical examples

2.1. The basic principles of the model

The DCP model can be divided into two main steps:

(i) By means of a procedure inspired by a standard Periodic Bond Chain (PBC) analysis (Hartman & Perdok, 1955*a*,*b*), a known crystal structure (hereafter called mother phase) is energetically and structurally analysed in order to extract one- or two-dimensional periodic fragments called PFs (Fig. 1).

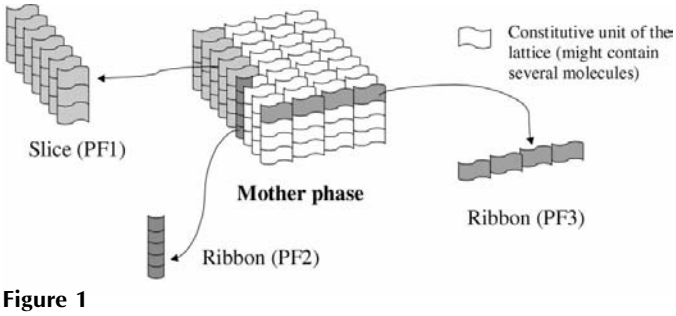


Figure 1
First step of the procedure: the mother phase is decomposed into one- or two-dimensional periodic fragments (PFs) preferentially with a low energy.

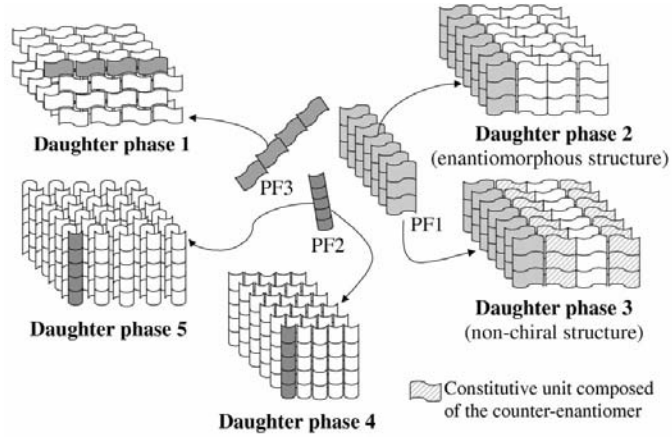


Figure 2
Second step of the procedure: for a given PF, some symmetry operators are added to generate new three-dimensional crystal packings (daughter phases).

(ii) In the second step, symmetry operators are added to the PFs, generating new three-dimensional crystal lattices called daughter phases (Fig. 2). If the PF is a ribbon (*i.e.* one-dimensional), two independent symmetry operators (associated with translations not parallel to that contained in the PF) are successively added to generate three-dimensional periodic structures. If the PF is a slice (*i.e.* two-dimensional), the addition of one symmetry operator (with a translation not parallel to those of the PF) is sufficient. The procedure is reversible and the mother phase can be reconstructed from the various daughter phases derived from it.

During construction particular attention is paid to two main criteria, which are: generation of a structure coherent with one of the 230 space groups and close packing (*i.e.* cohesion between the PFs). If necessary, the daughter phases are minimized using the Open Force Field module of the *Cerius*[2] software [force field: *Dreiding*2.21; Mayo *et al.*, 1990; Ewald summation for both van der Waals and coulombic interactions, ESP charges derived from *Gaussian*94 (Frisch *et al.*, 1995), basis set B3LYP, method 6-31G**]. The stability of the daughter phases is checked by computing their lattice energy, which should not differ by more than a few kJ mol^{-1} from that of the mother structure. When possible, experimental X-ray powder diffraction (XRPD) patterns, obtained under various conditions of nucleation and crystal growth, are compared with the calculated XRPD patterns of the new forms.

In the first step the same philosophy as that developed in crystal engineering is employed, since experimental features (such as the hydrogen-bond network, ionic bonds between salts or even hydrophobic layers) are extracted from the mother phase and used to generate new structures. However, the supramolecular synthons employed in crystal engineering are not necessarily periodic and may not contain symmetry operators or translations. In the second step, the successive addition of translational repeat distances to generate daughter phases is similar to that of the *ab initio* procedure developed by Perlstein (1992, 1994). Nevertheless, while the *ab initio* model consists of predicting crystal structures starting from a molecule or a finite collection of molecules, the DCP model has a more practical approach and allows the study of various phenomena as will be seen in the several examples treated further.

2.2. Choice of the PF

A periodic fragment can be described as a subunit composed of a part or the totality of the unit cell, which is periodically repeated according to one or two translations of the mother structure. For instance, for a structure in space group $P2_1/c$ with two molecules A and B in the asymmetric unit (*i.e.* $Z' = 2$), a possible subunit can be composed of the four equivalents of the molecule A. As shown in Fig. 1 the addition of one or two translations to the subunit gives a one-dimensional PF (ribbon) and a two-dimensional PF (slice), respectively.

The number of possibilities of PFs is quite large and increases with respect to the number of equivalent positions and to the number of molecules in the asymmetric unit of the mother phase (Z'). However, from more than 25 examples treated so far, the probability of obtaining relatively stable daughter phases is higher if the starting PF has a low energy. Therefore, PFs with interactions such as ionic bonds, hydrogen bonding or π–π interactions between two phenyl groups are chosen in this order of preference. Up to now, PFs are selected in a qualitative way (energies of the PFs are not calculated), but a PBC analysis (Hartman & Perdok, 1955*a*,*b*) should be employed in order to classify the PFs and extract those with the lowest energies. Indeed, one- and two-dimensional PFs can be regarded as possessing at least one and two major PBCs of the mother structure, respectively.

2.3. Theoretical examples

Hereafter, a right-handed molecule will be denoted by an arrow in an unbroken line and a left-handed molecule by an arrow in a dotted line. The nomenclature used to describe the location of the added symmetry operators is based on the crystallographic parameters of the mother phase. For instance, the addition of a twofold screw axis normal to PF at $y = 0$ refers to the cell parameter y of the mother phase.

2.3.1. Influence of the location of the added symmetry operator(s). A given PF can lead to different daughter phases, depending on the symmetry operators applied. A theoretical example is given in Fig. 3. Starting from a mother phase with space group $P2_1/c$, a PF containing the twofold screw axis, the glide plane **c**, inversion centres and two translations of the mother phase (T_y and T_z) is extracted and two daughter phases with space groups *Pnma* and *Pbca* are generated. In both constructions the same symmetry operator is applied (*i.e.* twofold screw axis), but at different locations. The construction of these two daughter phases can also be obtained for example from a ribbon containing only the translation T_z. In that case, two twofold screw axes at different locations have to be added to the PF (a general example of construction of a daughter phase from a one-dimensional PF is detailed in the supplementary materials). Therefore, a daughter phase constructed from a two-dimensional PF can always be generated starting from a one-dimensional ribbon contained in this slice. However, preference is given to the largest part of the mother structure, *i.e.* the two-dimensional PF.

2.3.2. Theoretical considerations on the number of molecules in the asymmetric unit (Z'). Some symmetries contained in the PF may disappear in the generated daughter phase, if they are not consistent with the added symmetry operator(s). Consider a mother phase with space group $P2_1/c$ and $Z'_m = 1$ (see the supplementary materials for construction[1]). The PF is composed of the four equivalents of the unit cell and the two translations T_y and T_z. The addition of a twofold screw axis normal to the yz slice and located in a general position (*i.e.* at $z \neq 0$, $\frac{1}{4}$, $\frac{1}{2}$ or $\frac{3}{4}$) suppresses the first twofold screw axis and the inversion centres of the PF.

[1] Supplementary data for this paper are available from the IUCr electronic archives (Reference: BK0114). Services for accessing these data are described at the back of the journal.

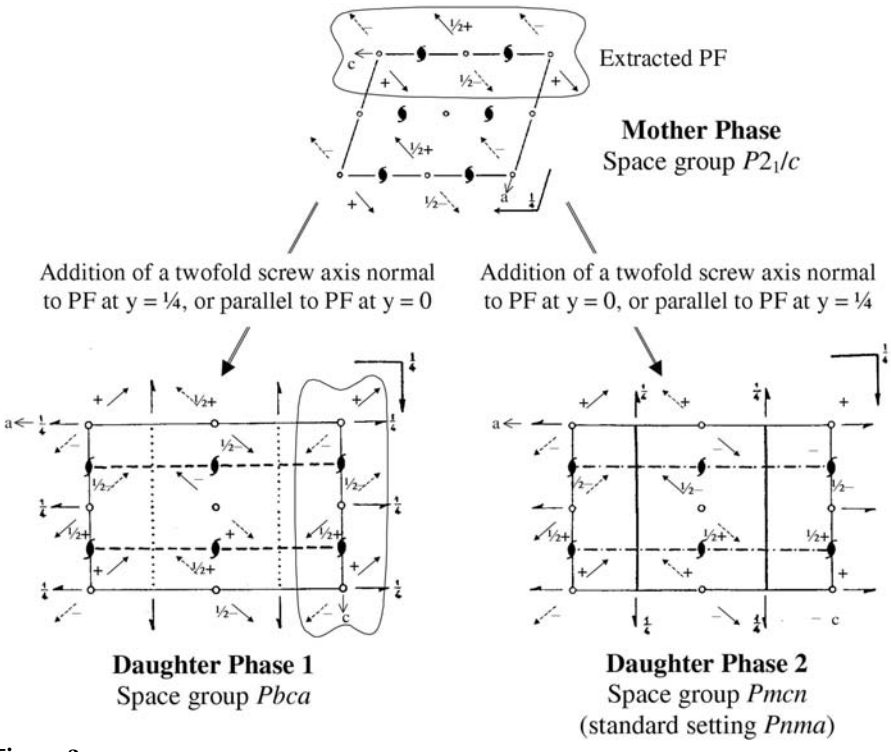


Figure 3
Construction of two structures with different space groups from a single two-dimensional PF extracted from a $P2_1/c$ structure, by means of the addition of a twofold screw axis at different locations. The same procedure can be applied to a one-dimensional PF. In that case, two added symmetries are necessary to generate the two daughter phases.

Moreover, it generates a glide plane **n**, leading to a daughter phase with space group $P2_1cn$ (standard setting: $Pna2_1$) with $Z'_d = 2$. The two molecules in the asymmetric unit are related by non-crystallographic symmetry, unless minimization destroys it. Therefore, the DCP model allows one to construct a daughter phase possessing supplementary degrees of freedom compared with the mother phase (*i.e.* $Z'_d > Z'_m$).

Conversely, starting from a mother phase having more than one molecule in the asymmetric unit ($Z'_m > 1$), periodic fragments containing only a part of the constituents of the asymmetric unit can be extracted, leading to several daughter phases with $Z'_d < Z'_m$ (examples are treated in the supplementary materials). In that case, if a PBC analysis is performed to choose the PFs, sub-slices containing the different molecules of the asymmetric unit have to be defined (Petit *et al.*, 1994).

3. Structural description of the 5-alkyl-5-aryl-hydantoins and prediction of a new crystal structure of the 5-(4′-chlorophenyl)-5-methylhydantoin

In this section the DCP model is applied as a tool to analyse a family of 5-alkyl-5-aryl-hydantoin structures (with alkyl = methyl or ethyl and aryl = *para*-substituted phenyl, Fig. 4). These compounds have an unusual high propensity to crystallize as conglomerates (Coquerel *et al.*, 1993). With two exceptions for 72H and pF61H, which crystallize as racemic compounds (space group *Pbca*), the majority of the derivatives belong to the space groups $P2_1$ and $P2_12_12_1$ (approximately 25% and 40–50%, respectively, of chiral molecules crystallize in these two space groups).

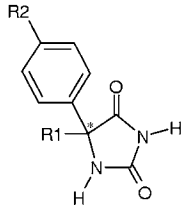
The structures are analysed by a standard PBC analysis. It appears that, in each structure, the entire hydrogen-bond network is contained in a homochiral ribbon characterized by a twofold screw axis running along **b**. Depending on the crystal packing, two types of ribbons (hereafter PFa and PFb) can be observed:

(i) For every compound forming a conglomerate, the periodic fragment PFa is characterized by a translation associated to the twofold screw axis close to 6.23 Å (±0.02 Å). The imidazole groups of the molecules included in PFa are coplanar and all their heteroatoms participate in the hydrogen-bond network (Fig. 5, left).

(ii) In the crystal structures of pF61H and 72H (space group *Pbca*), a different type of periodic fragment (PFb) is observed. The translation T_y associated with the twofold screw axis varies with the compound (7.20 Å for pF61H and 7.57 Å for 72H). The imidazole groups are not strictly coplanar to each other and one of the two O atoms is not involved in the hydrogen-bond network (Fig. 5, right). PFb is also encountered in 5-ethyl-5-methyl-hydantoin which crystallizes as a stable conglomerate, but with frequent lamellar epitaxy (Beilles *et al.*, 2001).

The construction of several daughter phases has been investigated for 5-(4′-methylphenyl)-5-methylhydantoin (71H). This compound crystallizes as a stable conglomerate ($P2_12_12_1$) and a metastable form of this conglomerate ($P2_1$). Monotropic behaviour in the range T = (293 K − melting point) is observed.

Let us consider, first, the construction of several polymorphic forms of the $P2_1$ structure, chosen as the mother phase. For this purpose, three slices with the lowest energy



Name(*)	R_1	R_2	Type of ribbon
61H	CH_3	H	PFa
71H	CH_3	CH_3	PFa
pF61H	CH_3	F	PFb
pCl61H	CH_3	Cl	PFa
pBr61H	CH_3	Br	PFa
pMeO61H	CH_3	OCH_3	PFa
81H	CH_3	CH_2CH_3	PFa
62H	CH_2CH_3	H	PFa
72H	CH_2CH_3	CH_3	PFb

Figure 4
Series of the 5-alkyl-5-aryl-hydantoins. (*) The two-digit number corresponds to the number of C atoms in the two substituents of the asymmetric carbon (except for pMeO61H).

have been selected from the monoclinic form (Fig. 6). These are slices (100), (10$\bar{1}$) and (001). As PFa provides the main part of the energy of the mother phase, the three slices contain this common feature to which a supplementary translation from the mother structure is associated. In each case, a twofold screw axis normal to PFa and located at $y = 0$ is added in order to generate structures with space group $P2_12_12_1$. A structural analysis of the three orthorhombic daughter phases shows that:

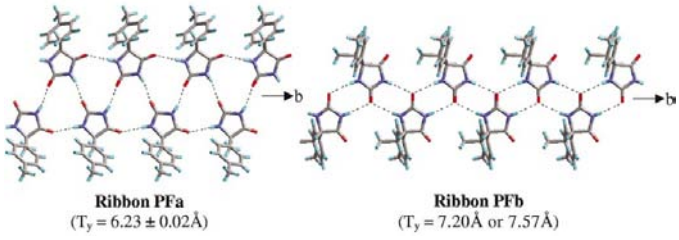


Figure 5
Description of the two hydrogen-bond ribbons encountered in the crystal structures of the 5-alkyl-5-aryl-hydantoins.

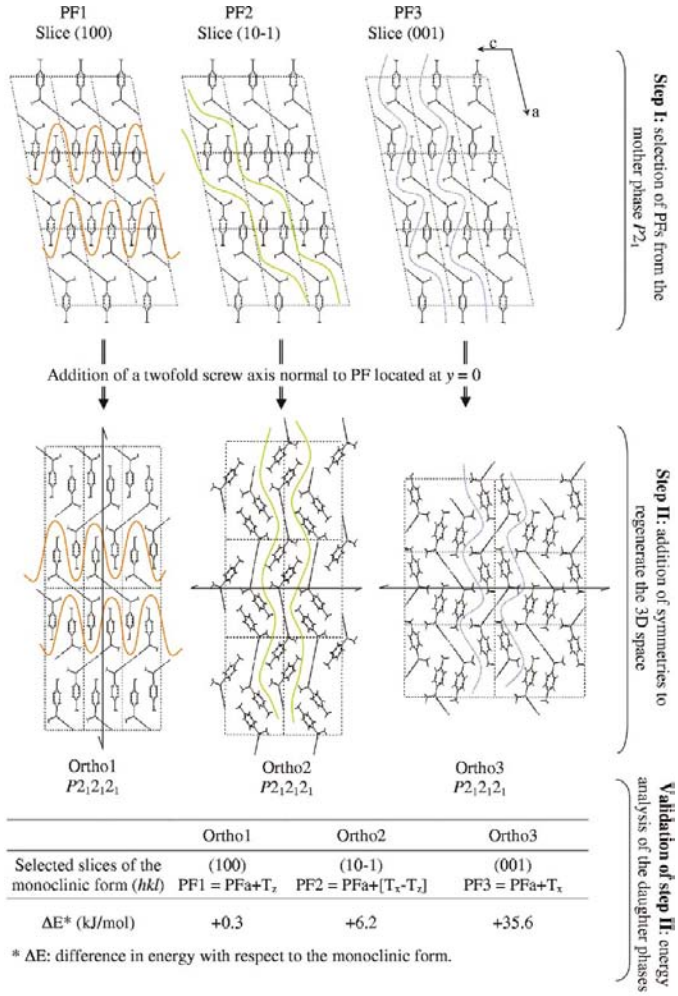


Figure 6
Construction of three orthorhombic structures from three slices of the monoclinic structure (case of 71H).

Table 1
Classification of the experimental crystal structures of the 5-alkyl-5-aryl-hydantoins.

For the general developed structure, see Fig. 4.

Type of structure	Name	a (Å)	b (Å)	c (Å)	β angle (°)
Mono ($P2_1$)	61H	10.81	6.21	7.36	108.07
	71H	11.72	6.25	7.35	103.79
Ortho1 ($P2_12_12_1$)	71H	22.75	6.25	7.34	90
	pCl61H	22.73	6.22	7.21	90
	pBr61H	22.95	6.22	7.35	90
	pMeO61H	23.59	6.22	7.18	90
	81H	24.63	6.25	7.31	90
Ortho2 ($P2_12_12_1$)	62H	10.70	6.23	15.55	90
Ortho3 ($P2_12_12_1$)	No experimental structure observed				
Ortho Rac (*Pbca*)	pF61H	25.78	7.13	11.48	90
	72H	25.42	7.57	11.63	90

(i) Ortho1 is similar to the experimental orthorhombic form of 71H, and isostructural to the experimental structures of compounds 81H, pCl61H, pBr61H and pMeO61H.

(ii) Ortho2 corresponds to the crystal packing of compound 62H.

(iii) Ortho3 has not been encountered so far.

The daughter phases are optimized and their lattice energies are computed (Fig. 6). The resulting energy values confirm lower lattice energies for the two experimental forms (*i.e.* Ortho1 and Mono).

In the following construction, a racemic compound with space group *Pbca* has been constructed starting from the ribbon PFa of the experimental orthorhombic structure of 71H (see supplementary materials for construction). A first glide plane **a** is added in $y = \frac{1}{4}$ to generate a two-dimensional slice composed of alternating homochiral ribbons of opposite handedness. Then, a second glide plane **b** parallel to the slice is added to generate a structure belonging to the space group *Pcab* (standard setting: *Pbca*). The daughter phase is optimized with the same procedure as that described above. Although the generated daughter phase is not competitive with the experimental structures in term of energy (*i.e.* +9.1 and +9.4 kJ per molecule with respect to Ortho1 and Mono structures, respectively), its crystal packing is closely similar to that encountered in the experimental racemic compound of pF61H and 72H, and contains the periodic fragment PFb. The transformation of the hydrogen-bond ribbon PFa to a ribbon of type PFb during minimization can be explained by the interactions between the phenyl groups of two adjacent PFs. In the chiral crystal packings (*i.e.* $P2_12_12_1$ and $P2_1$), the phenyl groups of two adjacent PFa are related by a twofold screw axis, leading to π-stacking interactions (Fig. 7*a*). In the racemic compound constructed from PFa, the two adjacent PFs are related by a glide plane, which creates sterical hindrances (Fig. 7*b*). During minimization, a rotation of the molecules in the PF is observed. In the minimized structure, although one heteroatom is no more involved in the hydrogen-bond network of the new periodic fragment (PFb), all the sterical bumps have disappeared (Fig. 7*c*).

Therefore, all the experimental crystal packings of the 5-alkyl-5-aryl hydantoins are connected to each other by

Table 2
Energy difference after minimization between the five types of crystal packings related by the DCP model (in kJ mol^{-1}).

The structure possessing the lowest lattice energy is taken as the reference. The experimental structures are in bold. The competitive crystal structures which could appear experimentally under suitable conditions are underlined.

	61H	71H	pF61H	pCl61H	pBr61H	pMeO61H	81H	62H	72H
Mono	**+0.38**	**0**	+0.2	+0.05	+0.1	+11.4	+6.9	+9.0	+9.2
Ortho1	0	**+0.3**	0	**0**	**0**	**0**	**0**	+11.2	+6.7
Ortho2	+$\bar{5}$.1	+6.2	+$\bar{6}$.1	+9.3	+12.7	+17.1	+14.6	**0**	+0.04
Ortho3	+16.8	+35.6	+24.1	+26.1	+22.8	+29.8	+27.1	+7.6	+20.8
Ortho Rac	+19.5	+9.4	**+8.8**	+14.3	+20.4	+50.0	+20.7	+8.3	**0**

means of the DCP model (Table 1). To confirm these structural relationships, the five types of crystal packings are constructed for each hydantoin. Starting from the experimental structure (mother phase), the daughter phases are constructed by using the same procedures as those described for 71H. The monoclinic form is realised by adding a simple translation to PF1 (see Fig. 6). The angle between T_y (translation associated with the twofold screw axis) and the added translation defines the monoclinic angle. For 72H, it is worth noting that the construction of the crystal packings has been realised before determination of the experimental structure by means of single-crystal X-ray diffraction. Therefore, the mother phase employed for the generation of the daughter phases has been derived from the monoclinic form of 71H (molecules of 71H have been replaced by molecules of 72H).

The daughter phases and the mother phase are then optimized. The resulting energies are summarized in Table 2. For pF61H, problems of parameterization for the F atom did not allow any conclusion on the relative stability between all the structures and therefore on the possible experimental appearance of a metastable conglomerate.

Although the structure with the lowest energy is not necessarily the experimental one (Lommerse *et al.*, 2000), energy ranking gives a list of possible candidates for experimentally observable phases. For 61H, pCl61H and pBr61H, competitive polymorphs in terms of energy with the experimental one are predicted. For 72H, the appearance of a metastable conglomerate or the crystallization of the pure enantiomer into the Ortho2 crystal packing is expected. For 61H, 71H, 62H, pCl61H and pBr61H, the melting peaks observed in differential scanning calorimetry (DSC) exhibit a shoulder. This phenomenon indicates polymorphism or the existence of a metastable racemic compound. Nevertheless, up to now, only the XRPD patterns of 71H and pCl61H exhibit extra identifiable peaks when considering the mother phase.

Various attempts to find experimentally the predicted monoclinic form of pCl61H have been performed. In 1-methyl-2-pyrrolidinone (NMP), a mixture of the orthorhombic form (type Ortho1) and of a new form has been obtained. After checking the chemical purity and by comparing the XRPD patterns of the mixture with the simulated patterns of the mother phase (Ortho1) and the monoclinic daughter phase, it

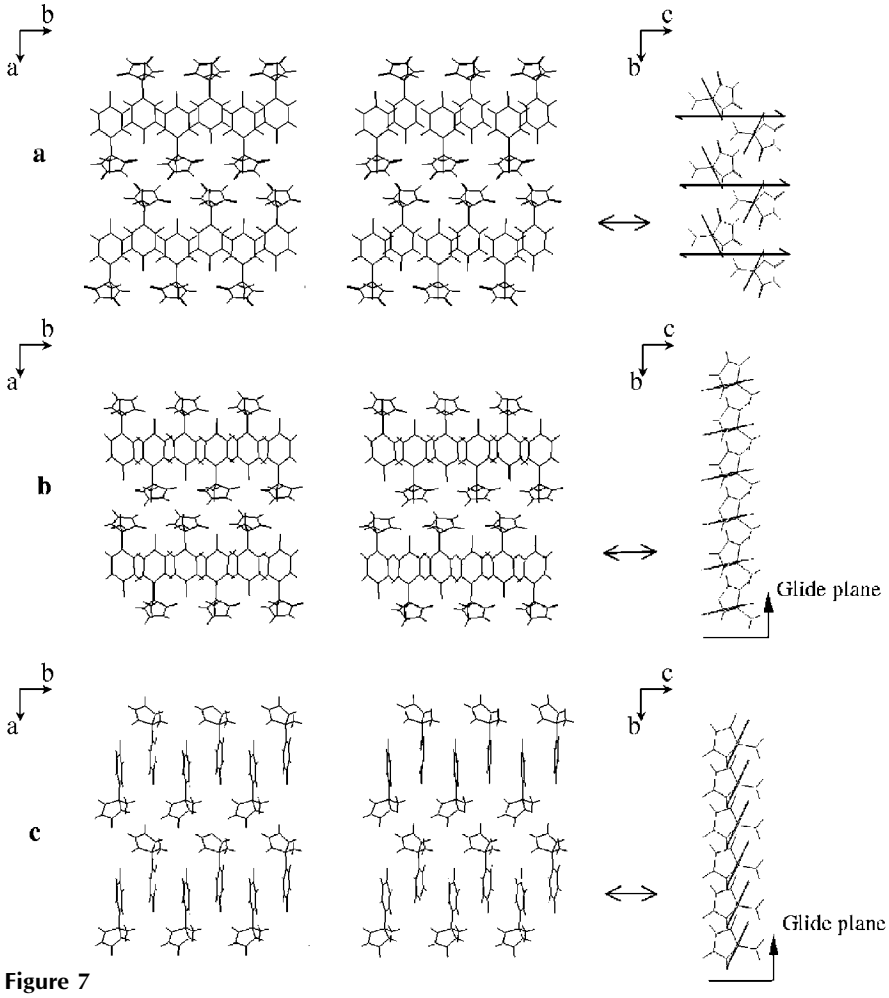
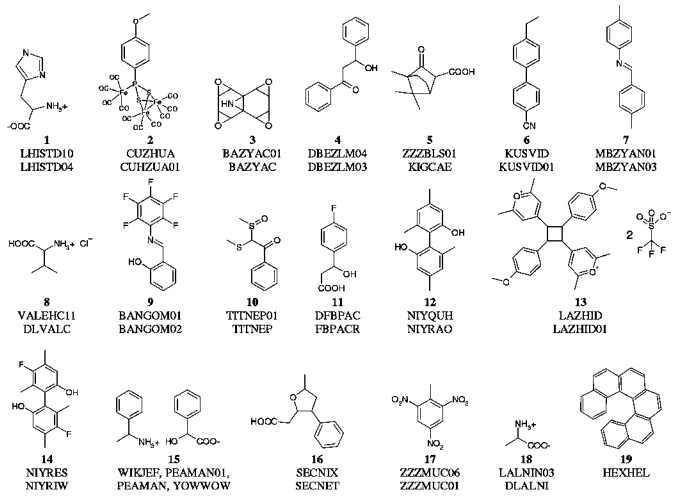


Figure 7
Comparison of the different stackings of PFa and PFb for pF61H (left: stereoviews along the *c* axis; right: projection along the *a* axis of the stacking of the phenyl groups). (*a*) type Ortho1. (*b*) Racemic compound constructed from PFa. (*c*) Experimental racemic compound (PFb).

Table 3
Developed formulae of the molecules retrieved from the CSD (upper part) and details on the construction procedure (lower part).

The first refcode is the mother phase (space group sg1), the second one is the daughter phase generated by the addition of the symmetry operator(s) (space group sg2).



No.	sg1	sg2	PF	Additional symmetry operator(s)
(1)	$P2_12_12_1$	$P2_1$	Slice	Translation
(2)	$P2_1/c$	$Pbca$	Slice	2_1 axis normal to PF at $y = \frac{1}{4}$
(3), (13)	$P2_1/c$	$Pbca$	Ribbon	(i) 2_1 axis normal to PF at $y = \frac{1}{4}$ (ii) 2_1 axis parallel to PF at $y = 0$
(4)	$P2_1/c$	$Pbca$	Slice	Glide plane a at $y = \frac{1}{4}$
(5), (6), (7), (8), (9)	$P2_1$	$P2_1/c$	Slice	Glide plane c normal to PF at $y = \frac{1}{4}$
(10)	$C2$	$C2/c$	Slice	Glide plane c normal to PF at $y = \frac{1}{4}$
(11)	$P2_1$	$Pna2_1$	Slice	Glide plane c normal to PF at $x = \frac{1}{4}$
(12)	$P2_12_12_1$	$P2_1/c$	Ribbon	(i) Glide plane a normal to PF at $z = 0$ (ii) Translation
(14)	$P2_12_12_1$	$P2_1/c$	Slice	Glide plane c normal to PF at $x = 0$
(15)	$P2_1$, $Z' = 2$	$P2_1$, $Z' = 1$	Slice	Translation
	$P2_1$ $Z' = 1$	$P2_12_12_1$	Slice	2_1 axis normal to PF at $y = \frac{1}{4}$
	$P2_12_12_1$	$Pca2_1$	Slice	Glide plane a normal to PF at $y = \frac{1}{4}$
(16)	$P2_1$	$P2_1/c$	Slice	Glide plane c normal to PF at x undefined
(17)	$P2_1/b$	$Pca2_1$	Slice	2_1 axis normal to PF at $z = \frac{1}{4}$
(18)	$P2_12_12_1$	$Pna2_1$	Slice	Glide plane b normal to PF at $x = 0$
(19)	$P2_12_12_1$		Slice	Glide plane a normal to PF at $y = \frac{1}{4}$

seems that this new crystal variety can be identified as the predicted monoclinic crystal structure (Fig. 8).

4. Illustration of the applicability of the DCP model to various fields

In order to assess the prediction capability of the DCP model and its practical utility in the study of various phenomena, couples of known structures are retrieved from the Cambridge Structural Database (CSD; http://www.ccdc.cam.ac.uk/). For this purpose a specially written program is used. It resides on an alphanumeric comparison of compound names in two files, both containing a particular space group. Although this method is not exhaustive (the same compound can be differently named in the CSD), comparison between enantiomers and racemic compounds or between two diastereomers can be performed, since the configurations of the asymmetric centres (*i.e.* +, −, *R*, *S*, D and L symbols) are not differentiated. Up to now we have focused on the following space groups, which are among the most populated (see Brock & Dunitz, 1994; Belsky *et al.*, 1995): $P2_1$, $C2$, $P2_12_12_1$, $P2_1/c$, $C2/c$, $Pca2_1$, $Pna2_1$ and $Pbca$. The crystal structures are then compared in view to determine if one or more common PF exists. Details concerning the experimental structures retained and the constructions of the derived phases starting from their parents are summarized in Table 3 [compounds (1)–(19)]. Only a few examples are detailed, the remainder being described in the supplementary materials.

4.1. Case of conformational polymorphism

Dibenzoylmethane [compound (4), Table 3] crystallizes in three polymorphic forms, *i.e.* a $P2_1/c$ form, a $Pbca$ form and a metastable $Pbca$ form (Etter *et al.*, 1987). Starting from the $P2_1/c$ structure, it is possible to generate a daughter phase corresponding to the experimental metastable $Pbca$ (see supplementary materials for construction). Nevertheless, the generated crystal structure has to be minimized, particularly at the interface between the PFs. The minimization (performed under the same conditions as those used for the

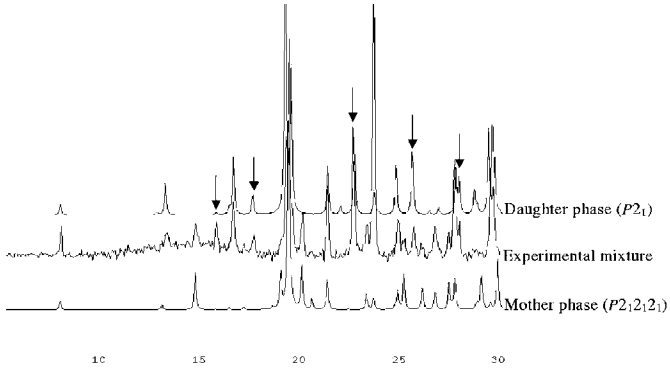


Figure 8
Comparison between simulated XRPD patterns of the mother phase of pCl61H, of the daughter phase and the experimental pattern obtained in NMP. Arrows indicate the non-identifiable peaks of the experimental XRPD pattern, corresponding to the daughter phase.

hydantoins) results in a change of the conformation of the molecules and more precisely in a modification of the torsion angle involving the phenyl groups situated in the outer region of the PFs.

Therefore, the minimization procedure can allow the disappearance of sterical bumps between two adjacent PFs (and/or the formation of interactions between molecules of two adjacent PFs, such as hydrogen bonding or π-stacking), leading to daughter phases with slightly different PFs from that of the mother phase. For instance, in the minimized racemic compound of 71H, the whole PF differs from that of the mother phase by a rotation of the molecules in the PF and a modification of the hydrogen-bond network, but the conformation of the molecules is conserved (see PFa and PFb, Fig. 5). In dibenzoylmethane the shape of the PF is globally conserved, but the conformation of the molecules has been sensitively modified. Therefore, the model is not restricted to the analysis and to the prediction of structures having the same fixed, rigid common periodic fragment and may deal (to some extent) with conformational polymorphism. Nevertheless, the minimization procedure has some limitations. For instance, the calculated XRPD patterns of the minimized daughter phase and of the minimized experimental *Pbca* of dibenzoylmethane are similar. However, there exists a non-negligible difference with the calculated XRPD pattern of the non-minimized experimental structure (Fig. 9), which highlights the fact that a small change in the lattice or especially in the asymmetric unit can lead to important changes in the XRPD pattern (a shift of the peaks or a modification of the intensities of the peaks, respectively).

4.2. Relationships between enantiomers and racemic compound structures: a case study: α-methylbenzylamine mandelate

α-Methylbenzylamine and mandelic acids are well known compounds, frequently employed as resolving agents in classical resolution. The diastereomeric salts of α-methylbenzylamine mandelate (15) crystallize in various forms (Table 3). Hereafter, π and π' represent (*R*,*R*)-(15) and (*S*,*S*)-(15), respectively; ν and ν' represent (*R*,*S*)-(15) and (*S*,*R*)-(15), respectively. The structural analysis of the diastereomeric salts shows that:

(i) Except the ν salt, which crystallizes with space group $P1$ ($Z' = 4$), the four other structures exhibit a packing composed of alternate hydrophilic and hydrophobic slices.

(ii) The two polymorphs of the π salt [Cambridge Structural Database (Allen, 2002) refcodes PEAMAN (Brianso *et al.*, 1979) and PEAMAN01 (Larsen & de Diego, 1993)] and the racemic compound YOWWOW (de Diego, 1995) possess the same hydrophilic slice PF1 (Fig. 10, left).

(iii) The mixed crystal (WIKJEF; de Diego, 1994) and the monoclinic form of the π salt (PEAMAN01) possess the same hydrophobic slice PF2 (Fig. 10, right).

All these structural observations can be highlighted by applying the DCP model to these crystalline structures. It appears that, except the $P1$ structure of the ν salt, all the crystal structures are related to each other (Fig. 11). The periodic fragments are PF1 and PF2, respectively, the hydrophilic and hydrophobic slices described above. Starting from the mixed salt WIKJEF ($P2_1$, $Z' = 2$), PF2 is extracted. A simple translation regenerates a $P2_1$ structure containing only the (*S*,*S*) pair (*i.e.* $Z' = 1$) and corresponds to PEAMAN01. From PEAMAN01 the crystal packing of PEAMAN ($P2_12_12_1$) is generated, by extracting PF1 and adding a twofold screw axis normal to it, at $x = \frac{1}{4}$. Finally, the racemic compound YOWWOW ($Pca2_1$) is constructed by the addition of a glide plane **a** to PF1, which can be extracted either from PEAMAN or from PEAMAN01 (constructions are detailed in the supplementary materials).

In order to understand why the pure ν salt does not crystallize with alternate slices, we attempted to realise a $P2_1$ structure for the ν salt, starting from the hydrophobic slice of WIKJEF, composed of (*R*,*S*)-(15) only. It appears that the hydroxyl group of the mandelate is not able to participate in the hydrogen-bond network, leading to hydrogen-bond ribbons rather than slices. Therefore, it seems that the ν salt cannot crystallize in the same packing as that of the π salt and adopts a less favourable one, which could explain the large difference in solubility and melting point between the two salts and therefore the frequent use of α-methybenzylamine to resolve mandelic acid (Ebbers *et al.*, 1997). On the contrary, the resolution of α-methylbenzylamine by mandelic acid might be difficult, considering the existence of the mixed crystal.

Therefore, the DCP model has been successfully applied to the construction of a daughter phase with $Z'_d \neq Z'_m$ and allowed the analysis of the structural relationships between enantiomer and racemic compound structures and more generally between enantiomorphous and non-enantiomorphous structures [for other examples, see compounds (5)–(12), in supplementary materials].

Therefore, the DCP model may be helpful in preferential crystallization (PC). In the case of stable conglomerates, the DCP model can compute possible unstable racemic

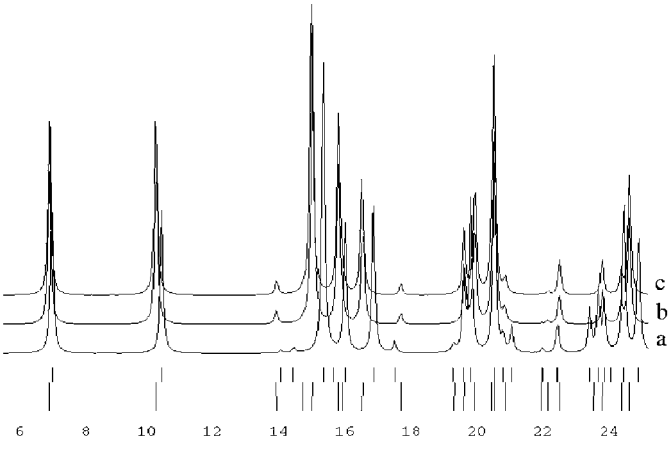


Figure 9
Comparison between the simulated XRPD patterns of (*a*) the experimental metastable *Pbca*; (*b*) the minimized experimental metastable *Pbca*; (*c*) the minimized predicted form of dibenzoylmethane [compound (4)].

compounds, which may impair the performances of PC (Houllemare-Druot & Coquerel, 1998; Dufour *et al.*, 2001). The closer their stability to their corresponding eutectic mixture, the more detrimental the impact on PC. Common faces between enantiomer and racemic compound crystals lead to similar heterochiral and homochiral interaction energies at the crystal face – mother liquor interface. Therefore, the desorption of the counter-enantiomer docked on these faces is slowed down. It results in a weak entrainment effect (*i.e.* the yield of the PC is limited), with at extreme cases its inhibition.

In the same way, it can also be applied to two diastereomers. For instance, the diastereomer *RRS* of the (5-methyl-3-phenyltetrahydrofuran-2-yl)acetic acid (16) crystallizes with space group $P2_1$. By applying the DCP model, it is possible to generate a daughter phase $P2_1/c$, closely similar to the crystalline structure of the diastereomers *RRR* and *SSS* (details of the construction are in the supplementary materials). This relation can be explained by the presence in both structures of the same type of ribbons containing the entire hydrogen-bond network. The asymmetric centre that differs in the two diastereomers does not participate in the hydrogen-bond network, but plays a role in the packing of the periodic fragment.

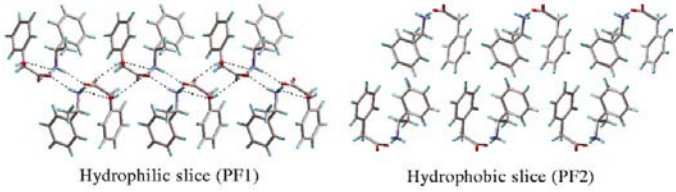


Figure 10
Periodic fragments encountered in the crystal structures of α-methylbenzylamine mandelate [compound (15)]. Left: hydrophilic slice (PF1) common to PEAMAN, PEAMAN01 and YOWWOW structures. Right: hydrophobic slice (PF2) common to PEAMAN01 and WIKJEF structures.

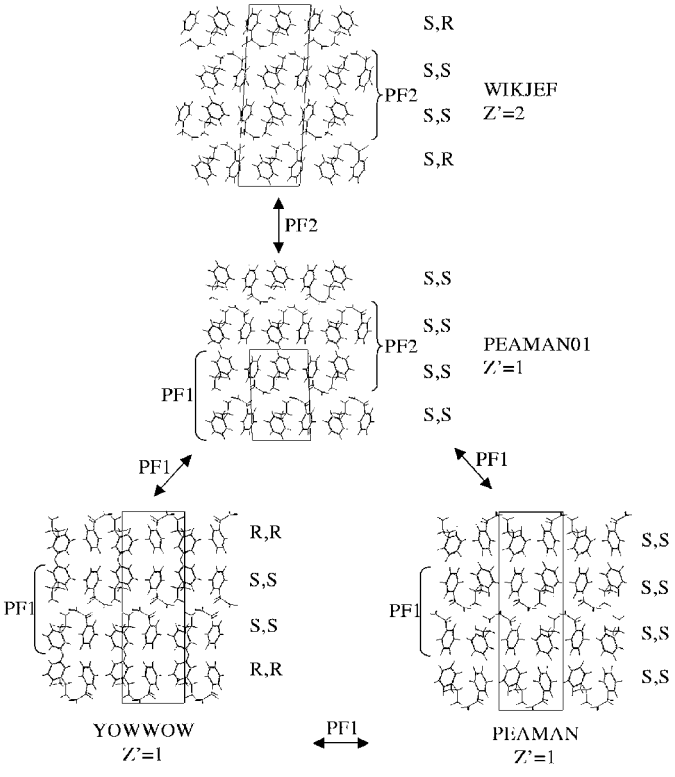


Figure 11
Application of the DCP model to the diastereomeric salts of α-methylbenzylamine mandelate [compound (15)].

4.3. Expression of the structural similarities by forming twins and epitaxies

Twinning can be the underlying expression of a derived polymorphic form. For instance, the hydantoin 71H crystallizes frequently in the monoclinic form with twins on the {100} faces, which correspond to the common PF between the monoclinic form and Ortho1 (see Fig. 6). At the twin interface, a two-dimensional orthorhombic-like packing is observed. The same type of relationship exists between the two polymorphic forms *Pbca* and $P2_1/c$ of trinitrotoluene [compound (17), see supplementary materials for construction]. The monoclinic form shows extensive twinning along the {100} planes (Gallagher & Sherwood, 1996; Gallagher *et al.*, 1997), which correspond to the common PF. It appears that, in suitable experimental conditions, the two forms crystallize simultaneously, *i.e.* as concomitant polymorphs (Bernstein *et al.*, 1999). For these two examples, twinning is a direct consequence of the similarities between the two structures and it is not surprising that the DCP model demonstrates the structural relationships between them.

In the case of epitaxy, different molecules crystallize with their lattices connected by a specific and reproducible orientation. Epitaxy between the racemic compound and the enantiomer of alanine (18) occurs on {001} faces (Weissbuch *et al.*, 1994), which correspond to the periodic fragment common to the two structures (see supplementary materials for construction and discussion).

Hexahelicene crystallizes in the chiral space group $P2_12_12_1$ (19). Surprisingly, crystals grown from a racemic mixture reveal almost no enantiomeric excess. Green & Knossow (1981) explained the formation of these particles by the epitaxial association of macroscopic homochiral lamellar fragments parallel to the {100} faces. Starting from the $P2_12_12_1$ structure, we tried to find a daughter phase that could describe the interface zone between two adjacent lamellae. For this purpose, a periodic fragment corresponding to the slice (100) of the $P2_12_12_1$ structure was extracted and a glide plane **a** normal to this PF was added to generate a daughter phase with space group $Pna2_1$ (Fig. 12). The generated daughter phase could describe exactly the epitaxy zone occurring along the [100] direction. As the relative lattice energies of both structures are closely similar (*i.e.* a difference of ∼3.3 kJ per molecule), this type of epitaxy phenomenon could occur

frequently without providing a high excess of energy (Gervais *et al.*, 2002).

4.4. Concomitant polymorphism

Therefore, the DCP model constructs daughter phases with structural features existing in the mother phase so that they are likely to have relative stabilities close to that of the mother phase (similar lattice energies). As shown above, these two factors can lead, for instance, to the presence of a metastable racemic compound, or to the formation of twinning or lamellar epitaxy, but it can also result in concomitant polymorphism (Bernstein *et al.*, 1999). Indeed, as the mother phase and the daughter phase(s) possess similar lattice energies and related structures, it is likely that under definite conditions of nucleation and crystal growth a mixture of phases occurs. For instance, fullerene C60 (Herbstein, 1996) can crystallize as a mixture of two structures having both an orientational disorder, *i.e.* a predominant cubic close-packed structure (with repetition of three layers *ABC*) and a hexagonal close-packed structure (with repetition of only two layers *AB*).

4.5. Mechanism of heterogeneous nucleation

The DCP model can be useful in the understanding of mechanisms occurring during polymorphic transitions or even during racemate/conglomerate transitions, although for the latter (involving the change from one phase $\langle R,S\rangle$ to two phases $\langle R\rangle$ and $\langle S\rangle$) the predominant role played by diffusion can lead to more complex mechanisms. Consider the transition

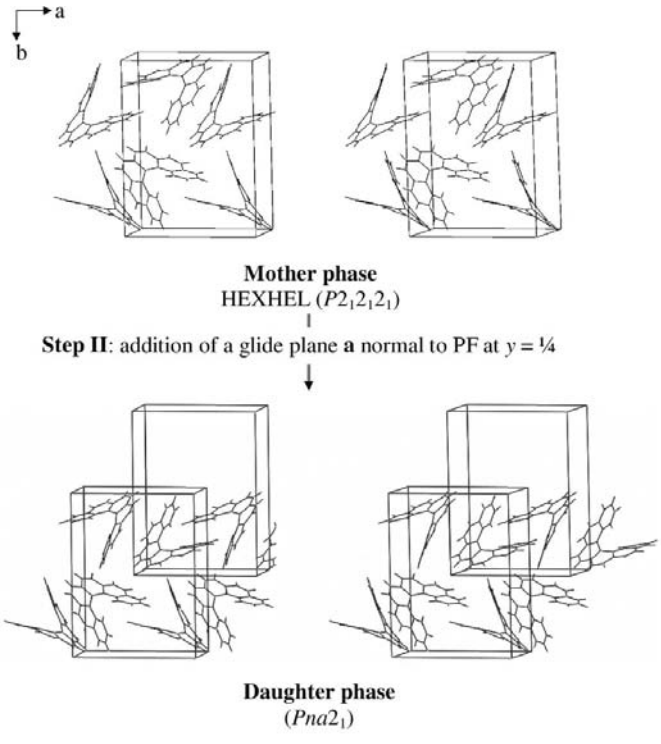


Figure 12
Construction of a $Pna2_1$ structure of hexahelicene [compound (19)]. The daughter phase corresponds to the interface of the 'twinned' crystal (in fact epitaxy) described by Green & Knossow (1981).

by means of a destructive/reconstructive process between two crystalline structures related by the DCP model, *i.e.* possessing a common PF. During the transition, provided that form I has not totally disappeared, its structural features can help in the heterogeneous nucleation of form II. In the case of a two-dimensional PF, the structures have a crystallographic plane in common where the heterogeneous nucleation of one form on another one is easily initiated. With a one-dimensional PF in common, the heterogeneous nucleation may also occur *via* a ledge-directed mechanism already pointed out (Bonafede & Ward, 1995; Ward, 2001). Nevertheless, the structural resemblance between two forms leads to similar lattice energies, so that the transition is conducted by a weak driving force (ΔG). The slow rate of the polymorphic transition between the metastable form ($P2_1$) and the stable one ($P2_12_12_1$, Ortho1) of 71H in ethanol and acetone (more than 1 month for a total transformation if no seeding is carried out) exemplifies this effect.

5. Conclusions

A two-step model predicting new crystal structures called DCP (derived crystal packings) is proposed. It consists of extracting a one- or two-dimensional periodic fragment from a mother structure and in adding new symmetry operators, which generate new three-dimensional crystal packings. For a given mother phase, several derived structures can thus be generated either by using different periodic fragments and/or by applying different sets of symmetry operators to the same periodic fragment.

It has been successfully applied to:

(i) Polymorphism: One of the predicted crystal structures of 5(4′-chlorophenyl)-5-methylhydantoin has been found experimentally. The current status of the model can withstand slight modifications of the constitutive unit (*i.e.* a small variation of the geometry of the molecules).

(ii) Enantiomorphous/non-enantiomorphous space groups (which can also include polymorphism). The interest of the method relies on the possibility of improving the understanding of the relationships between enantiomer structures and racemic compound structures. Moreover, systematic construction of enantiomorphous and non-enantiomorphous structures for a given chiral compound could predict a metastable (or unstable) racemic compound or conglomerate.

(iii) It could also explain and predict the occurrence of particular phenomena, such as lamellar epitaxy on crystals grown from a racemic solution and/or twinning.

As the DCP model is easy to apply, it could be used as a routine method after every resolution of crystal structure in order to detect possible polymorphs, metastable racemic compounds or conglomerates and to explain the possible occurrence of twinning or lamellar epitaxy.

In some cases, the single crystal used to solve the structure may not be representative of the bulk (Aakeröy & Seddon, 1993). This can be detected by comparison between the XRPD pattern of the bulk (measured with a spinning capillary set-up in order to avoid preferential orientation effects) and the

diffraction pattern simulated from the single-crystal data (unfortunately, it is often overlooked!).

The extended version of this model could focus on structures:

with $Z' > 1$ for which the possibility of generating new packings increases rapidly with the value of Z'. For this purpose, space groups having a high percentage of structures with $Z' > 1$ will be first contemplated (Brock & Dunitz, 1994);

of solvates and their possible derived sub-solvated or asolvated phases;

of different conformers of the molecule in mother and daughter phases.

In the same way, the model could be applied to the study of structural relationships between incommensurate and commensurate structures. Moreover, it could focus on the automation of the DCP model, particularly by the characterization of the PBCs in the mother phase for the extraction of the PFs.

The authors would like to thank Morgan Pauchet for elaboration of the program of comparison of space groups in the Cambridge Structural Database, and CRIHAN (Région Haute-Normandie) for providing molecular modelling tools.

References

Aakeröy, C. B. & Seddon K. R. (1993). *Chem. Soc. Rev.* pp. 397–407.

Allen, F. H. (2002). *Acta Cryst.* B**58**, 380–388.

Beilles, S., Cardinaël, P., Ndzié, E., Petit, S. & Coquerel, G. (2001). *Chem. Eng. Sci.* **56**, 2281–2294.

Belsky, V. K., Zorkaya, O. N. & Zorky, P. M. (1995) *Acta Cryst.* A**51**, 473–481.

Bernstein, J., Davey, R. J. & Henck, J-O. (1999). *Angew. Chem. Int. Ed. Engl.* **38**, 3440–3461.

Blagden, N., Davey, R. J., Rowe, R. & Roberts, R. (1998). *Int. J. Pharm.* **172**, 169–177.

Bonafede, S. J. & Ward, M. D. (1995). *J. Am. Chem. Soc.* **117**, 7853–7961.

Brianso, M. C., Leclercq, M. & Jacques, J. (1979). *Acta Cryst.* B**35**, 2751–2753.

Brittain, H. G. (1999). Editor. *Polymorphism in Pharmaceutical Solids*. New York: Marcel Dekker.

Brock, C. P. & Dunitz, J. D. (1994). *Chem. Mater.* **6**, 1118–1127.

Chin, D. N., Palmore, G. T. R. & Whitsides, G. M. (1999). *J. Am. Chem. Soc.* **121**, 2115–2122.

Coquerel, G., Petit, M. N. & Robert, F. (1993). *Acta Cryst.* C**49**, 824–825.

Desiraju, G. R. (1995). *Angew. Chem. Int. Ed. Engl.* **34**, 2311–2327.

Desiraju, G. R. (1997). *Curr. Opin. Solid State Mater. Sci.* **2**, 451–454.

Diego, H. L. de (1994). *Acta Cryst.* C**50**, 1995–1998.

Diego, H. L. de (1995). *Acta Cryst.* C**51**, 935–937.

Dufour, F., Gervais, C., Petit, M. N., Perez, G. & Coquerel, G. (2001). *J. Chem. Soc. Perkin Trans. 2*, pp. 2022–2036.

Dzyabchenko, A. V., Pivina, T. S. & Arnautova, E. A. (1996). *J. Mol. Struct.* **378**, 67–82.

Ebbers, E. J., Plum, B. J. M., Ariaans, J. A., Kaptein, B., Broxterman, Q. B., Bruggink, A. & Zwanenburg, B. (1997). *Tetrahedron: Asymm.* **8**, 4047–4057.

Erk, P. (2001). *Curr. Opin. Solid State Mater. Sci.* **5**, 155–160.

Etter, M. C., Jahn, D. A. & Urbańczyk-Lipkowska, Z. (1987). *Acta Cryst.* C**43**, 260–263.

Frisch, M. J., Trucks, G. W., Schlegel, H. P., Gill, P. M. W., Johnson, B. G., Robb, M. A., Cheeseman, T. A. K., Petersson, G. A., Montgomery, J. A., Raghavachari, K., Al-Laham, M. A., Zakryewski, V. G., Ortiz, J. V., Foresman, J. B., Cioslowski, J., Stefanov, B. B., Nanayakkara, A., Challacombe, M., Peng, C. Y., Ayala, P. Y., Chen, W., Wong, M. W., Andres, J. L., Replogle, E. S., Gomperts, R., Martin, R. L., Fox, D. J., Binkley, J. S., Defrees, D. J., Baker, J., Stewart, J. P., Head-Gordon, M., Gonzalez, C. & Pople J. (1995). *Gaussian94.* Gaussian Inc., Pittsburgh, PA, USA.

Gallagher, H. G., Roberts, K. J., Sherwood, J. N. & Smith, L. A. (1997). *J. Mater. Chem.* **7**, 229–235.

Gallagher, H. G. & Sherwood, J. N. (1996) *J. Chem. Soc. Faraday Trans.* **92**, 2107–2116.

Gavezzotti, A. (1991). *J. Am. Chem. Soc.* **113**, 4622–4629.

Gavezzotti, A. (1993). *PROMET*. University of Milano, Italy.

Gavezzotti, A. (1994). *Acc. Chem. Res.* **27**, 309–314.

Gdanitz, R. J. (1997). *Theoretical Aspects and Computer Modeling*, edited by A. Gavezzotti, ch. 6, pp. 185–201. New York: John Wiley and Sons Ltd.

Gervais, C., Beilles, S., Cardinaël, P., Petit, S. & Coquerel, G. (2002). *J. Phys. Chem.* **106**, 646–652.

Görbitz, K. H. & Hersleth, H.-P. (2000). *Acta Cryst.* B**56**, 526–534.

Green, B. S. & Knossow, M. (1981). *Science*, **214**, 795–797.

Hartman, P. & Perdok, W. G. (1955*a*). *Acta Cryst.* **8**, 49–52.

Hartman, P. & Perdok, W. G. (1955*b*). *Acta Cryst.* **8**, 521–524.

Herbstein, F. H. (1996). *J. Mol. Struct.* **374**, 111–128.

Hofmann, D. W. M. & Lengauer, T. (1997). *Acta Cryst.* A**53**, 225–235.

Houllemare-Druot, S. & Coquerel, G. (1998). *J. Chem. Soc. Perkin Trans 2*, pp. 2211–2220.

Karfunkel, H. R. & Leusen, F. J. J. (1992). *Speedup*, **6**, 43–50.

Kobayashi, Y., Ito, S., Itai, S., Yamamoto, K. (2000). *Int. J. Pharm.* **193**, 137–146.

Larsen, S. & de Diego, H. L. (1993). *Acta Cryst.* B**49**, 303–309.

Leusen, F. J. J. (1996). *J. Crystal Growth*, **166**, 900–903.

Lommerse, J. P. M., Motherwell, W. D. S., Ammon, H. L., Dunitz, J. D., Gavezzotti, A., Hofmann, D. W. M., Leusen, F. J. J., Mooij, W. T. M., Price, S. L., Schweizer, B., Schmidt, M. U., van Eijck, B. P., Verwer, P. & Williams D. E. (2000). *Acta Cryst.* B**56**, 697–714.

Mayo, S. L., Olafson, B. D. & Goddard III, W. A. (1990). *J. Phys. Chem.* **94**, 8897–8909.

Molchanova, M. S., Pivina, T. S., Arnautova, E. A. & Zefirov, N. S. (1999). *J. Mol. Struct. (Theochem.)* **465**, 11–24.

Molecular Simulations Inc. (1999). *Cerius2*. Version 4.0. Molecular Simulations Inc., Cambridge, England.

Moulton, B. & Zaworotko, M. J. (2001). *Chem. Rev.* **101**, 1629–1658.

Perlstein, J. (1992). *J. Am. Chem. Soc.* **114**, 1955–1963.

Perlstein, J. (1994). *J. Am. Chem. Soc.* **116**, 11420–11432.

Petit, S., Coquerel, G. & Hartman, P. (1994). *J. Cryst. Growth*, **137**, 585–594.

Rustichelli, C., Gamberini, G., Ferioli, V., Gamberini, M. C., Ficarra, R. & Tommasini, S. (2000). *J. Pharm. Biomed. Anal.* **23**, 41–54.

Threlfall, T. L. (1995). *Analyst*, **120**, 2435–2460.

Vippagunta, S. R., Brittain, H. G. & Grant, D. J. W. (2001). *Adv. Drug Delivery Rev.* **48**, 3–26.

Ward, M. D. (2001). *Chem. Rev.* **101**, 1697–1725.

Weissbuch, I., Kuzmenko, I., Vaida, M., Zait, S., Leiserowitz, L. & Lahav, M. (1994). *Chem. Mater.* **6**, 1258–1268.

Wöstheinrich, K. & Schmidt, P. C. (2001). *Drug Dev. Ind. Pharm.* **27**, 481–489.