

Local similarity in organic crystals and the non-uniqueness of X-ray powder patterns

HEINRICH KARFUNKEL,^{a*} HEIKE WILTS,^a ZHIMIN HAO,^b ABUL IQBAL,^b JIN MIZUGUCHI^c AND ZHIJIAN WU^d

^aNovartis Services AG, R-1008, CH-4002 Basel, Switzerland, ^bCiba SC, R-1047, CH-4002 Basel, Switzerland, ^cYokohama National University, Yokohama 240, Japan, and ^dChangchun Institute of Applied Chemistry, Changchun 130022, China. E-mail: heinrich.karfunkel@sn.novartis.com

(Received 1 April 1998; accepted 29 April 1999)

Abstract

Two new concepts for molecular solids, 'local similarity' and 'boundary-preserving isometry', are defined mathematically and a theorem which relates these concepts is formulated. 'Locally similar' solids possess an identical short-range structure and a 'boundary-preserving isometry' is a new mathematical operation on a finite region of a solid that transforms mathematically a given solid to a locally similar one. It is shown further that the existence of such a 'boundary-preserving isometry' in a given solid has infinitely many 'locally similar' solids as a consequence. Chemical implications, referring to the similarity of X-ray powder patterns and patent registration, are discussed as well. These theoretical concepts, which are first introduced in a schematic manner, are proved to exist in nature by the elucidation of the crystal structure of some diketopyrrolopyrrole (DPP) derivatives with surprisingly similar powder patterns. Although the available powder patterns were not indexable, the underlying crystals could be elucidated by using the new technique of *ab initio* prediction of possible polymorphs and a subsequent Rietveld refinement. Further *ab initio* packing calculations on other molecules reveal that 'local crystal similarity' is not restricted to DPP derivatives and should also be exhibited by other molecules such as quinacridones. The 'boundary-preserving isometry' is presented as a predictive tool for crystal engineering purposes and attempts to detect it in crystals of the Cambridge Structural Database (CSD) are reported.

1. Introduction

The structural similarity between two crystals composed of different but basically isometric molecules has been discussed in terms of various concepts such as isomorphism, isotypism, isostructuralism, homostructuralism *etc.* (Kálmán *et al.*, 1993, and references therein; Rutherford, 1997). This diversity of concepts is further enlarged by the addition of further qualifiers such as 'approximate' (*e.g.* approximate isomorphism) or 'main-part' (*e.g.* main-part isostructuralism). In this article we will introduce another similarity concept termed 'local similarity' or, for reasons discussed later,

'synthomorphism', which refers to a special class of molecular solids. In view of the many existing similarity concepts the reader may well ask whether a new concept is necessary at all. We will show in the following that the new class of molecular solids (*i.e.* we intend to also treat systems which are not crystals in the strict mathematical sense) cannot be described properly by the existing concepts. After introducing the concepts of 'local similarity' and 'boundary-preserving isometry' in a general manner, the relation between 'locally similar solids' and their X-ray powder patterns will be discussed. This is, in fact, a reversal of the chronology of events (mainly for didactical reasons) since these studies began with the discovery of pigment crystals, composed of different molecules possessing practically identical X-ray powder patterns. Only after elucidating some of these crystal structures by computational methods based on *ab initio* crystal structure predictions, the new structural concepts were born.

This article deals with practically oriented subjects related to chemical solids and materials design, the underlying concepts are, however, purely mathematical. These concepts, termed 'local similarity' and 'boundary-preserving isometry' are, to the best of our knowledge, new in the domain of mathematical crystallography. To underline the mathematical nature of these concepts and their independence of chemistry, we have chosen a schematic, non-chemical presentation using 'pre-Columbian motifs' like drafts for their introduction. This, we hope, will free the reader from any chemical ballast at the beginning and will allow him to focus his attention on simple three-dimensional geometry.

2. Local similarity and boundary-preserving isometries

Let us begin with a simple schematic example shown in Fig. 1: given three different three-dimensional objects ('molecules') composed of a common central part (indicated by the 'T'-shaped geometrical form in Fig. 1) and a combination of two different substituents. They are referred to in Fig. 1 (top) as 'molecules' *a*, *b* and *c*. These three-dimensional geometrical objects (or molecules) give rise to two three-dimensional crystals (assume for simplicity, that the third crystallographic

axis is perpendicular to the drawing plane), indicated in Fig. 1 as ‘crystals’ (I) and (II). Crystal (I) contains only molecule *a* and crystal (II) contains molecules *b* and *c* in a 1:1 proportion. The unit cells are indicated by the grey background and the cell dimension of (II) in the horizontal direction is exactly twice the corresponding cell length of (I). A careful inspection of the schematic crystals shown in Fig. 1 reveals that the two crystals have a very similar three-dimensional packing. The contacts between the ‘atoms’, indicated schematically in Fig. 1 by the position of the edges of the geometrical objects, are identical in both crystals. We now pose three questions to be answered in the following discussion.

- (i) Are the two crystals isomorphous?
- (ii) Can the two crystals be classified as isostructural?
- (iii) Does there exist any geometrical operation, which relates (I) to (II) and *vice versa*?

The two molecular crystals are not isomorphous since there is no possibility for an isomorphous substitution. The shape of the molecules is too different and there is no reason to assume that parts in (I), consisting of whole molecules, can be replaced by parts of (II), consisting of whole molecules, without introducing large defects.

Concerning isostructuralism things are a little bit more complicated. We will adopt the terminology of Kálmán *et al.* (1993). According to their approach, the degree of similarity between two crystals is described in terms of several continuous descriptors of isostructurality. The two descriptors we consider as most important are

$$\Pi = (a + b + c)/(a' + b' + c') - 1, \quad (1)$$

where *a*, *b*, *c* and *a'*, *b'*, *c'* are the orthogonalized lattice parameters of the related crystals, and

$$I_i(n) = |[\Sigma(\Delta R_i)^2/n]^{1/2}| \times 100, \quad (2)$$

where *n* is the number of distance differences (ΔR_i) between the crystal coordinates of equivalent atoms within the same section of the asymmetric unit of the related structures. Another descriptor related to the ‘packing coefficient increment’ has been introduced earlier by Kálmán *et al.* (1991) but, as stressed by Rutherford (1997), unfortunately omitted from the 1993 paper and will not be discussed in this article. The molecular isometricity index, $I_i(n^*)$ (Kálmán *et al.*, 1993) should be mentioned at this point as well. It has the same mathematical form as (2), but is related to the superimposed molecules. A prerequisite for using the isostructurality indices $I_i(n)$ or $I_i(n^*)$ is the definition of sets of equivalent atoms in the two crystals.

A reasonable calculation of the similarity indices (1) and (2) for the case presented in Fig. 1 is possible if we introduce some modifications, which are consistent with the idea of similarity definition. Rather than using the lattices given we allow the usage of any multiple cell [in our case two unit cells of (I)], which minimizes Π . Thus, the Π index will be exactly 0. Extending the definition of $I_i(n)$ to the whole unit cell (rather than the asymmetric unit) now allows the calculation of $I_i(n)$, since a 1-to-1 correspondence exists between the atoms of (I) (now using the double cell as shown in Fig. 1) and (II). The value of the isostructurality index for the case presented in Fig. 1 depends on the relative population of atoms in the central common part of the molecules. If the

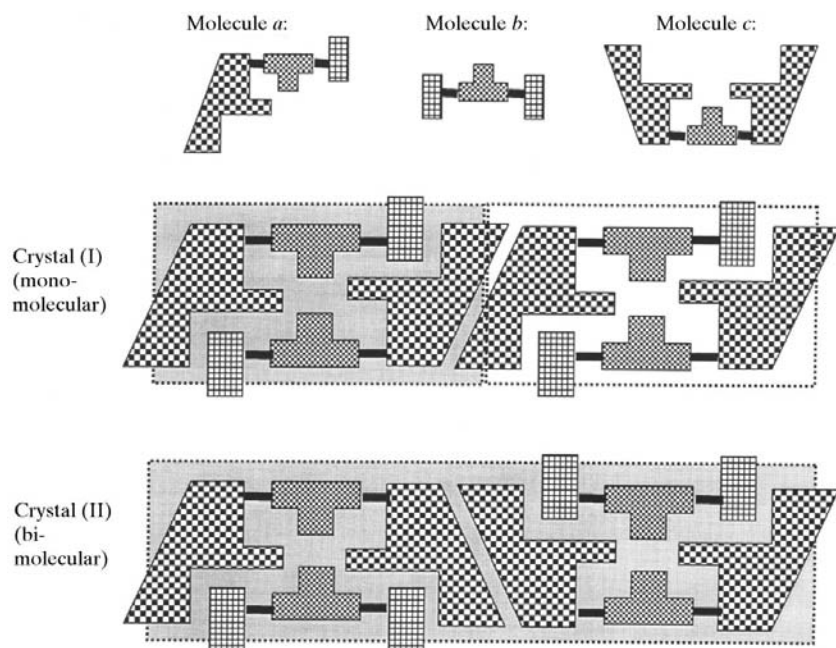


Fig. 1. A simplified scheme to illustrate the new similarity concept discussed in this article. Shown is the packing of irregular three-dimensional geometrical objects termed molecule *a*, *b* and *c* (top), which possess a common central part (here a ‘T’-shaped body) and combinations of two different substituents. The object molecule *a* forms a three-dimensional periodic solid termed crystal (I), the objects molecules *b* and *c* form crystal (II). The third dimension extends in the perpendicular direction of the drawing plane. The unit cells are symbolized by a grey background.

majority of atoms populate the substituents rather than the central part, then just over half the equivalent atoms in the crystals have exactly the same coordinates, *i.e.* ΔR_i in (2) will be 0 and just under half the equivalent atoms have completely different coordinates. In that case the isostructurality index $I_i(n)$ will be approximately 65%, which is in our opinion definitely too low to accurately reflect the degree of similarity between the crystals. To summarize the answer to question (ii): after introducing some modifications to the mathematical definition of isostructurality indices, it is possible to calculate these indices for the presented case. However, the isostructurality index $I_i(n)$ might be too low to accurately reflect the degree of similarity between the given crystals.

Before proceeding to an answer to the third question posed above, we will define the concept of 'locally similar crystals' or 'synthomorphic crystals'. Any crystal can be characterized independently of a unit-cell choice or a coordinate system by the list of all interatomic distances between element pairs and their relative frequency of occurrence. Thus, we have separate lists of C—C distances, C—N distances and so on. In practice (Karfunkel & Leusen, 1992) the distances are truncated at a certain distance value, R , and the lists of interatomic distances and their frequency of occurrence are presented by histograms, where distance ranges rather than distances are used. An example of such a histogram is shown in Fig. 2 and can be presented numerically as a (non-square) matrix, termed in this article the 'radial distribution matrix' of a crystal, truncated at the distance R .

Definition 1. Two crystal structures containing different molecules are locally similar by a degree R (R being a distance in Å) if the corresponding radial distribution matrices truncated at the distance R are identical and if

the radial distribution matrices truncated at a distance greater than R are different.

It should be emphasized that Definition 1 is an operational definition associated with a simple algorithm. For two crystals with specified atomic coordinates one can easily calculate the radial distribution matrices. Each column of these matrices corresponds to a distance range (*e.g.* column 47 in both matrices corresponds to distances between 7.1 and 7.2 Å). The columns of the two matrices are compared until non-equal columns are reached. The middle of the distance range associated with the first non-equal column is the degree of local similarity, R , in Definition 1. The calculation is independent of atom numbering, definition of equivalent atoms, definition of molecules *etc.*, which are needed in the definitions of isostructurality.

In other words, two crystals are locally similar if they are indistinguishable as long as only distances below a certain distance threshold, R , are taken into account. Thus, the two schematic crystals shown in Fig. 1 are locally similar to a certain degree R (not specified in this fictive case any further).

The answer to question (iii) posed above is rather surprising and unexpected: the two crystals shown in Fig. 1 are related by a geometrical operation, termed hereafter a 'boundary-preserving isometry', which will now be explained in detail. An isometric operation is a geometric operation, like a common symmetry operation, which leaves distances invariant. In the present context the isometric operation is applied only to a subregion of the (infinite) crystal. In Fig. 3 some examples of crystal subregions are shown schematically (slices, cylinders and spheres). These subregions are finite in one (slice), two (cylinder) or three (sphere) dimensions.

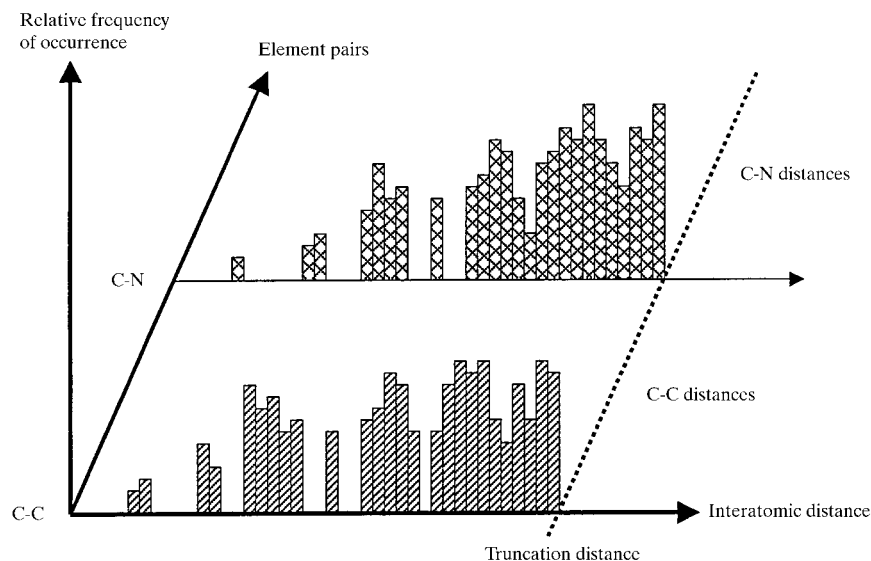


Fig. 2. A schematic presentation of the radial distribution function of a crystal. A radial distribution function (in this article) represents all the interatomic distances and their relative frequencies of occurrence. For practical purposes the distances are truncated at a certain cut-off distance. For each pair of elements a histogram is defined. The x direction corresponds to interatomic distances, the z direction corresponds to the relative frequency of these distances in the crystal. The various chemical element pairs are positioned in an arbitrary manner along the y direction. A matrix, termed the radial distribution matrix, presents the set of histograms for numerical purposes. Notice also the truncation at a specified distance.

The crucial point is that these regions are further subdivided into a boundary region (termed hereafter simply as 'boundary') and an interior region (termed hereafter simply as 'interior' and symbolized by the different shadings in Fig. 3). Owing to the boundary, there is no contact between the interior and the rest of the crystal.

Definition 2. An isometric operation applied to a subregion of a given crystal, composed of boundary and interior, is a 'boundary-preserving isometry' if:

- (i) The boundary is mapped onto itself.
- (ii) The isometric operation is not a symmetry operation of the crystal.

The fact that the boundaries are mapped onto themselves [point (i) of Definition 2] means that the rest of the crystal (not affected by the isometric operation) will not 'notice' the changes that have taken place in the interior of the transformed region. Point (ii) of Definition 2 guarantees that the result of such an isometric operation is non-trivial and that the resulting solid is indeed something new. Owing to the spatial periodicity, an infinite number of symmetry-related subregions to a given one do exist. It is, however, important to bear in mind that the 'boundary-preserving isometry' is not applied to all the symmetry-equivalent subregions, but only to a predefined subset of them (see below).

To render the answer to question (iii) more understandable, we assume that the space group of (I) in Fig. 1 is $P\bar{1}$ (with two molecules in the unit cell). Fig. 4 shows a segment of the infinite crystal (I). The chosen subregion is a slice perpendicular to the drawing plane and the subdivision of that region into boundary and interior is indicated there as well. An isometric operation is associated with a mirror plane at the middle of the slice (indicated by the dashed line) and perpendicular to the

drawing plane. The mirroring operation is certainly not a symmetry operation on the whole crystal. Applying the mirroring operation only to the indicated slice will map the 'T'-shaped geometrical figures (central part of the molecules) lying in the indicated boundaries onto themselves. The interior is, however, transformed to something new, not identical to the indicated original interior. An infinite number of translationally equivalent slices to that shown in Fig. 4 exist. If we apply this boundary-preserving isometry to every second slice we obtain exactly the 'bimolecular' crystal (II) of Fig. 1.

The secret behind the apparently trivial schemes in Fig. 1 was the existence of a boundary-preserving isometry and applying this isometric operation to every second slice will transform (I) into (II) and *vice versa*. However, this is not yet the end of the story. We may ask about the result of applying the same boundary-preserving isometry only to every n th ($n > 2$) slice of (I). The outcome will be a 'trimolecular' crystal consisting of molecules a , b and c , with molecules b and c always in a 1:1 proportion. For $n = 3$ [*i.e.* applying the isometry to every third slice of the infinite number of symmetry-equivalent slices of (I)] the outcome is a new trimolecular crystal of the molecular composition $a_1b_1c_1$ built of a slice of molecules a , followed by a slice of molecules b , followed by a slice of molecules c , and so on. We may also apply the boundary-preserving operation to the equivalent slices at random and the outcome will be a trimolecular 'defect-free' non-crystalline solid (more precisely: the resulting solid is strictly periodic in two dimensions and non-periodic in the third dimension). It should be emphasized that such a 'defect-free' solid is not a solid solution in the usual sense. To summarize: the existence of a boundary-preserving isometry in a crystal gives rise to an infinite number of different solids, crystalline as well as non-crystalline, with the same short-range packing.

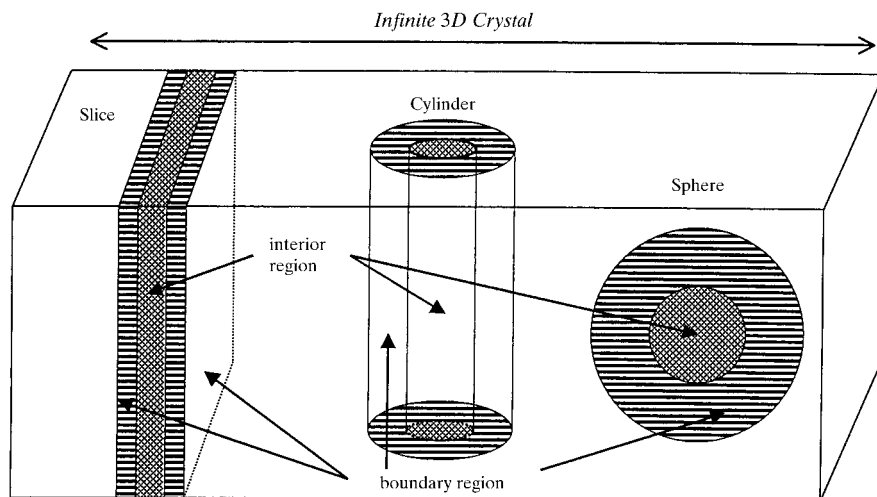


Fig. 3. Explanation of the concept of 'boundary-preserving isometry'. In an infinite crystal three types (slice, cylinder and sphere) of finite regions are schematically shown. Each region is further subdivided into an interior region and a boundary region (horizontal shading). An isometry applied to a finite region will transform only that region, leaving the rest of the crystal unchanged. In a boundary-preserving isometry the boundary regions are mapped onto themselves. If the boundary-preserving isometric operation is not a symmetry operation on the whole crystal, something new will be generated in the interior and the changes occurring there will not be 'noticed' by the rest of the crystal.

The relation between local similarity and boundary-preserving isometry is now obvious and summarized in the following theorem:

Theorem 1. A sufficient condition for local similarity of solids is the existence of a boundary-preserving isometry. The degree of local similarity is directly related to the ‘width’ of the boundary of the subregion to which the isometry is applied.

Proof. We will give a sketch of a proof rather than a strict mathematical proof. The proof is based on the subregion with the largest possible boundary. In contrast to the simple drawing shown in Figs. 3 and 4, the outer surface of the subregion to be transformed by the boundary-preserving isometry can adopt many different shapes. The choice of the surfaces defining the subregion or the interior is non-unique as well. Among all subregions which will be transformed by the boundary-preserving isometry there exists, however, a subregion whose boundary contains the boundaries of any other choice. By the ‘width’ of a boundary we mean the diameter of the largest sphere, which can still freely move between the inner and outer surface of the largest boundary. The diameter of this largest sphere is a lower bound to the truncated distance R at which the radial distribution matrices of

the original and transformed solids are still indistinguishable.

3. Local similarity and crystal engineering

We have shown that isostructuralism and local similarity are distinct concepts. Since radial distribution matrices can be defined for any solid, local similarity is a concept not restricted to crystals. On the other hand, the descriptors on which isostructuralism is based are strictly limited to periodic systems. The main conceptual difference is, however, the theorem formulated in the last section.

The radial distribution matrices as well as the descriptors of isostructuralism are *a posteriori* concepts applicable to existing data (*e.g.* for classification purposes). They cannot be used to predict anything new. The boundary-preserving isometry is, in contradistinction, a design principle that allows the prediction of new solids. Let us return to Fig. 1. The question whether (I) and (II) are locally similar or isostructural is, in our opinion, rather academic. Consider the case where only molecule *a* and (I) in Fig. 1 are shown [*i.e.* assume that (II), and molecules *b* and *c* are not yet known]. A much more relevant question would then be:

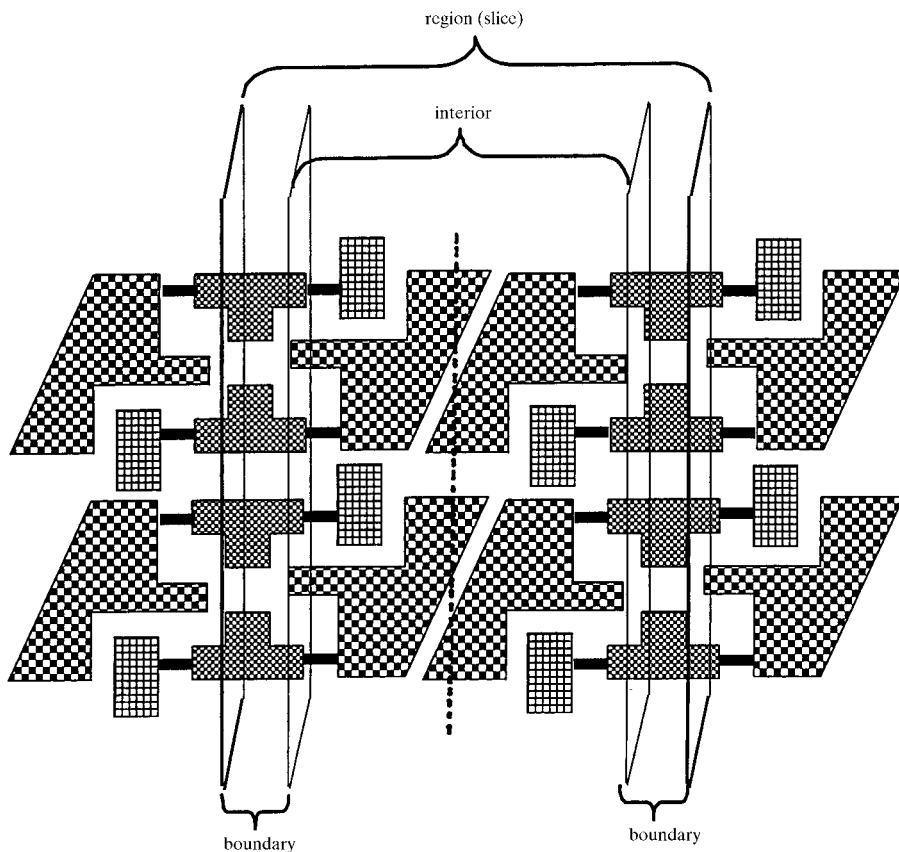


Fig. 4. The special feature of the schematic crystals shown in Fig. 1 is the existence of a boundary-preserving isometry. Crystal (I) is shown. The indicated finite region is a slice perpendicular to the drawing plane. It is further subdivided into boundary and interior. The isometry (a mirror plane indicated by the dashed line) will map the central ‘T’-shaped parts of the geometrical objects onto themselves.

do any solids exist, composed of other molecules, which are locally similar (or isostructural) to (I) and what do they look like? By detecting the boundary-preserving isometry in (I) (as shown in Fig. 4) we can now, utilizing the theorem, predict the existence of an infinite number of new solids which are locally similar to (I) [among these predictions (II) will also be present]. As we will show later with real chemical systems, this is not an academic question about chemical curiosities. A typical scenario encountered in practice: assume that the monomolecular crystal (I) in Fig. 1 is a commercially important pigment with exceptionally low solubility and other valuable properties whose production is heavily protected by patent. The detection of the boundary-preserving isometry in that pigment would be the key to new locally similar pigments with similarly low solubility and other valuable properties, but since not yet known, not protected by any patent.

4. Chemical implication of local similarity

The concept of synthons has been introduced recently (Desiraju, 1995, 1996, 1997; Reddy *et al.*, 1996) as a qualitative analysis and design tool in molecular crystal engineering. Synthons, as we interpret them, are structural units comprised of interacting groups, in a well defined orientation, belonging to molecules in close contact. A typical example of a (strong) synthon is the interaction unit generated by two planar amide groups with a double hydrogen bond between them. Synthons are the smallest structural units that contain all the information inherent in the recognition events through which molecules assemble into solids. As we have shown, locally similar solids are indistinguishable in terms of their short-range structure, although they may differ in their physical properties such as colour or melting point. It is therefore obvious that locally similar crystals have exactly the same synthons (not qualitatively but in a quantitative sense). For this reason we propose the term *synthomorphism* to be used rather than the currently used term 'local similarity'. However, during the rest of this article we will nonetheless still adhere to the term *local similarity*.

Another chemical implication of locally similar solids are the X-ray powder patterns when the degree of local similarity, R , is high (say, above 8 Å). The prevailing opinion is that each crystalline solid has its own characteristic X-ray powder pattern which may be used as a 'fingerprint' for its identification (West, 1984). Examples of crystalline solids which contradict this view are not only interesting from an academic point of view, but may have far-reaching commercial consequences, as encountered in the patent law (Rosenberg, 1975). An accurate definition of a chemical invention requires unique characterization of the objects in terms of structure. In the case of solids such as pigments, a recitation of certain inherent properties (such as the powder

pattern) may be the only way to express a complex chemical composition. Claims directed to such compositions have been referred to as 'fingerprint' claims (Rosenberg, 1975).

The X-ray powder pattern of a crystal depends on all interatomic distances, but the large distances contribute less to the intensities of the peaks. Since locally similar crystals have the same set of interatomic distances for distances less than the threshold distance R (or the degree of local similarity) we may expect that the powder patterns of crystals with a high degree of local similarity will be similar as well. For sufficiently large R values they may, from the practical point of view (*e.g.* for patenting purposes), be even identical.

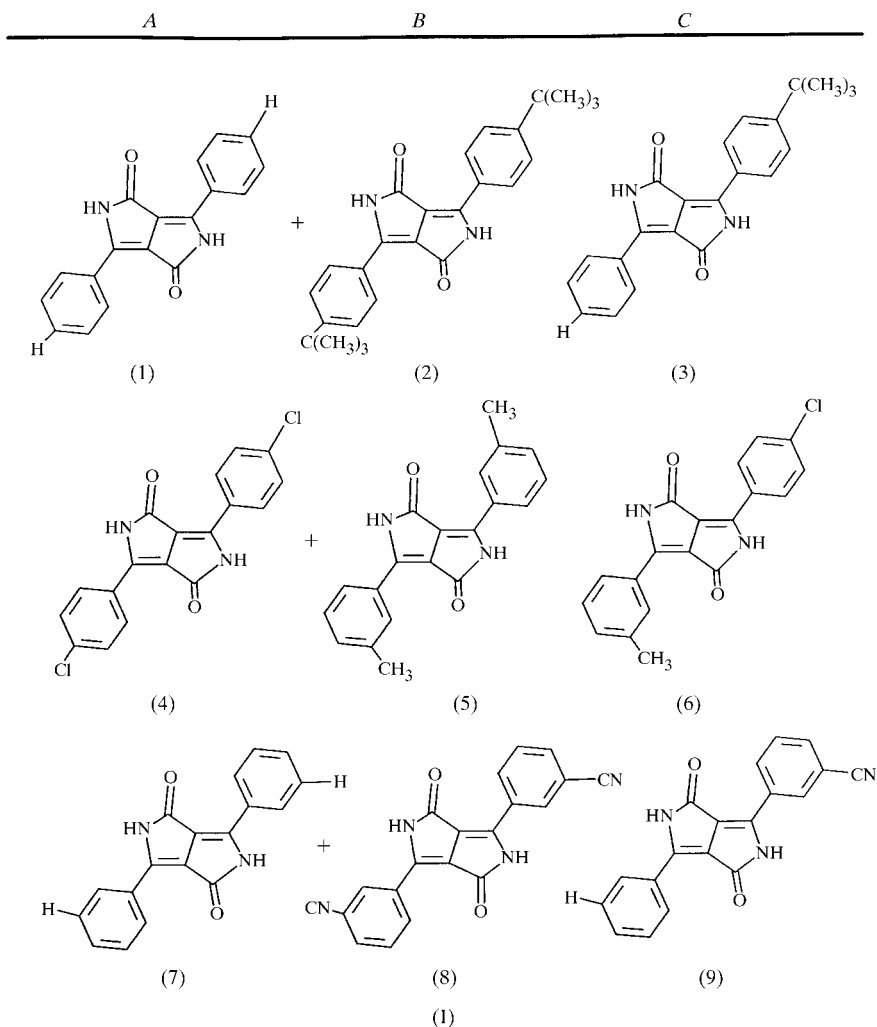
5. DPP pigments as an example of locally similar crystals

In two recent patents (Hao *et al.*, 1996; Mizuguchi *et al.*, 1997) it has been noted that some bimolecular mixtures of symmetrically substituted 1,4-diketopyrrolopyrrols (DPP), as shown in the first two columns of Scheme (I) (designated *A* and *B*), will form crystals containing both molecules in a 1:1 proportion [represented by '+' in Scheme (I)]. Thus, molecules (1) and (2) will form bimolecular crystals whose X-ray powder pattern differs from the overlay of the patterns derived from the crystals of the pure substances. The crystal structures of the pure symmetric molecules (1) [or (7)], (4) and (2) have been published elsewhere (Mizuguchi *et al.*, 1992, 1993; Mizuguchi, 1998), but the structures of the mixed crystals composed of (1)+(2) *etc.* as well as the structures of the crystals containing the asymmetrically substituted molecules (3) *etc.* have not yet been published.

The remarkable phenomenon is that the powder patterns of these bimolecular crystals [(1)+(2) *etc.*] are almost identical to the powder patterns of the monomolecular crystal consisting of the asymmetrically substituted DPP molecules of column C of Scheme (I). In the following we will focus our attention on the molecules in the first row of Scheme (I) (*i.e.* with H and *tert*-butyl as substituents). In Fig. 5, three measured X-ray powder patterns of the bimolecular crystal comprising the symmetrically substituted molecules (1) and (2) (lower row in Fig. 5: a' , b' and c') and of the monomolecular crystal of molecule 3 (upper row in Fig. 5: a , b and c) are shown. A list of relative positions and relative intensities of the main peaks is reported elsewhere (Hao *et al.*, 1996). The three measurements (a), (b), (c) or (a'), (b'), (c') refer to different samples of the same pigments and give us a qualitative estimate of the experimental error. From Fig. 5 it is obvious that the variations of the powder patterns between the different samples of the same crystal are comparable to the differences between the two crystals (1)+(2) and (3). It is true that an experienced investigator may still detect some systematic differences between the two sets of powder patterns (*e.g.* the two resolved peaks just below

15° in *b'* and *c'*). However, from the practical viewpoint of patent registration, where only the few strongest peaks are documented, such minor differences might be ignored. Whether the two crystals have, up to the experimental error, practically the same X-ray powder pattern, although they contain different molecular species, or whether they are practically different, is, in our opinion too subjective to be discussed here. As a matter of fact, such a degree of similarity has been classified in the cited patent (Hao *et al.*, 1996) as 'practically identical' and this fact was for us a sufficient motivation for further investigations. Simulated powder patterns (see below) of predicted crystals may well clarify this point later.

The only way to gain detailed structural information about the crystals is the computational *ab initio* prediction of these crystals based on molecular structure only. Although we still encounter in recent reviews (MacDonald & Whitesides, 1994; Wolff, 1996) some scepticism toward such methods, the *ab initio* prediction of possible crystal structures of polar organic molecules has been shown (Karfunkel & Gdanitz, 1992; Karfunkel *et al.*, 1993; Payne *et al.*, 1997) to be reliable if performed properly. The general outline of the method is as follows: initially, high precision *ab initio* quantum chemical calculations (usually Hartree-Fock or density functional calculations with 6-31G** basis functions) are performed on the molecules. Based on these results, an



Since single crystals were not available for these pigments, the routine X-ray structure determination methods could not be applied. Even an indexing of the lines of the powder patterns, from which the cell constants could be deduced, was not possible owing to the insufficient number of well separated lines.

individual force field for each molecule is derived in which atomic charges appear to be of uppermost importance. In a second step a large number (several thousands) of 'crude' crystals structures is generated for each relevant space group using Monte Carlo simulated annealing methods (Gdanitz, 1992). In a third step the

many 'crude' structures are divided into a few (say dozens) groups, or clusters, and out of each cluster one representative structure is chosen for further processing. Finally, a full (force field) energy minimization with respect to all degrees of freedom of the crystal is performed on the chosen 'crude' cluster representatives. Out of the resulting force field optimized structures, we chose those few with the lowest total energy to represent the possible crystal structures in which given molecules can be packed. The packing procedure following the quantum chemical calculations has been implemented as a fully automated process in the *Cerius2* modelling package. The final crystal structure elucidation makes use of the available low-resolution X-ray powder pattern and is based on the assumption that the crystal under investigation should correspond to one of the predicted low-energy crystals. The simulation of the powder patterns of all predicted crystals should allow the identification of the corresponding crystal. Owing to factors such as the approximate nature of the force field

and the neglect of temperature, a corresponding crystal, should it indeed be among the predicted ones, will deviate to some extent from the investigated crystal. Thus, we cannot expect that the simulated and measured powder patterns will be identical, but we hope that the degree of similarity is sufficient for identification purposes. After having identified a corresponding crystal, a Rietveld refinement (Young, 1995) is performed using the predicted crystal as a starting structure. A recent example of a crystal structure elucidation when neither single crystals nor indexable powder patterns were available is given elsewhere (Karfunkel *et al.*, 1996).

Applying the outlined computational procedure to the crystals containing the centrosymmetric molecules (1)+(2) and the non-centrosymmetric molecule (3), and using any of the measured powder pattern shown in Fig. 5, yielded the crystal structures shown in Fig. 6 with the corresponding fractional coordinates reported in Table 1.

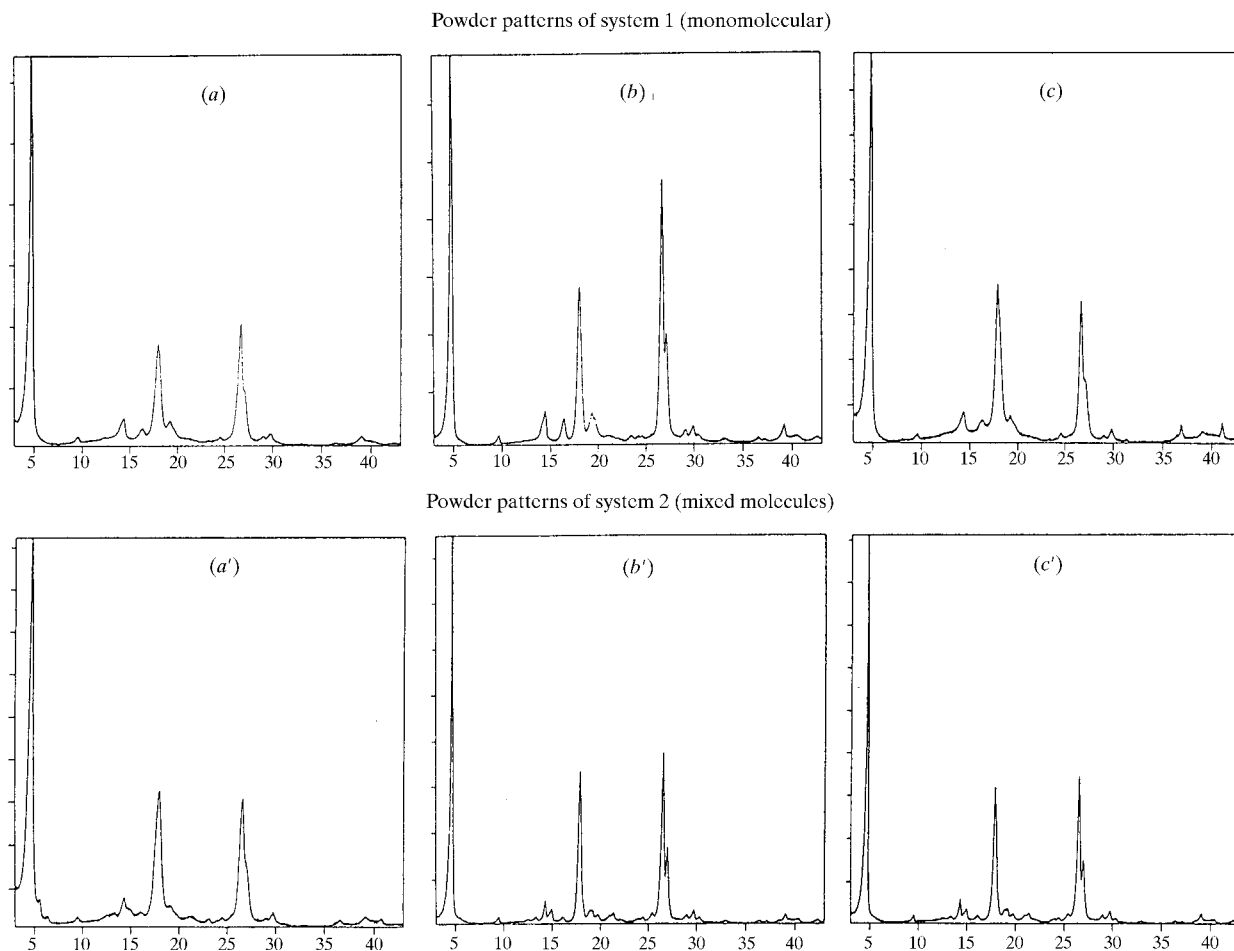


Fig. 5. Powder patterns measured for three different samples of DPP pigments (intensities versus 2θ). In the upper row the powder patterns of three samples (a), (b) and (c) of the crystal containing only molecule (3) are shown. In the lower part three powder patterns of samples (a'), (b') and (c') from the bimolecular crystal containing molecules (1) and (2) in a 1:1 ratio are given.

The degree of agreement between the simulated and measured powder patterns was not overwhelming, but in view of the low resolution and the poor crystallinity of the samples not surprising. Simulated powder patterns of the two elucidated structures are shown in Fig. 7.

Except an instrumental broadening taken into account to reflect a real situation, the two simulated patterns are 'noise free'. Up to a small discrepancy at 15° we consider the two simulated patterns as practically identical.

Are the two elucidated pigment crystals locally similar? From the atomic coordinates, the radial distribution matrices of both crystals can be easily computed. Up to a distance $R = 13.7 \text{ \AA}$ these matrices are identical, which means that the two crystals are identical as long as we consider only interatomic distances below this value. Above this cut-off distance one can detect differences resulting from H-H distances. If one considers only distances between non-H atoms, differences between the radial distribution matrices will occur only at distances well above 14 \AA . Thus, the degree of local similarity defined by the radial distribution matrices is rather large. The two crystals are related by a boundary-preserving isometry, as shown in Fig. 8. The isometric operation is a mirroring at a plane indicated by the dashed line, which exchanges H atoms and *tert*-butyl groups, but maps the phenyl rings onto themselves.

In the proceeding sections, we have emphasized the predictive advantage of the concept of boundary-preserving isometry over *a posteriori* concepts such as the radial distribution matrices or the descriptors of isostructurality. Assuming for an instance that bimolecular crystals were unknown, we would have now, after having detected the boundary-preserving isometry in the monomolecular pigment crystal, deduced the bimolecular pigment crystal. Moreover, we can now predict the existence of trimolecular pigments (crystals as well as solids which are not crystals in the strict mathematical sense) comprising molecules (1), (2) and (3) with the composition $(1)_n(2)_n(3)_m$, with m being any number between 0 and infinity.

The insoluble DPP pigment can be generated from soluble precursors or so-called 'latent pigments' (Zambounis *et al.*, 1997) by heat treatment or hydrolysis so that the generation of the free pigment molecules and the crystallization process take place simultaneously. Thus, solids of the random type with the composition $(1)_n(2)_n(3)_m$ mentioned above will be generated when the precursors of the three molecules (1), (2) and (3) in actual proportions (which depend on the kinetics of the decomposition) are hydrolysed simultaneously in the same reaction vessel. In an earlier patent on mixed DPP pigments (Rochat *et al.*, 1983), a process for the production of mixed pigments has been disclosed. The final pigment is made from a solution of a mixture of

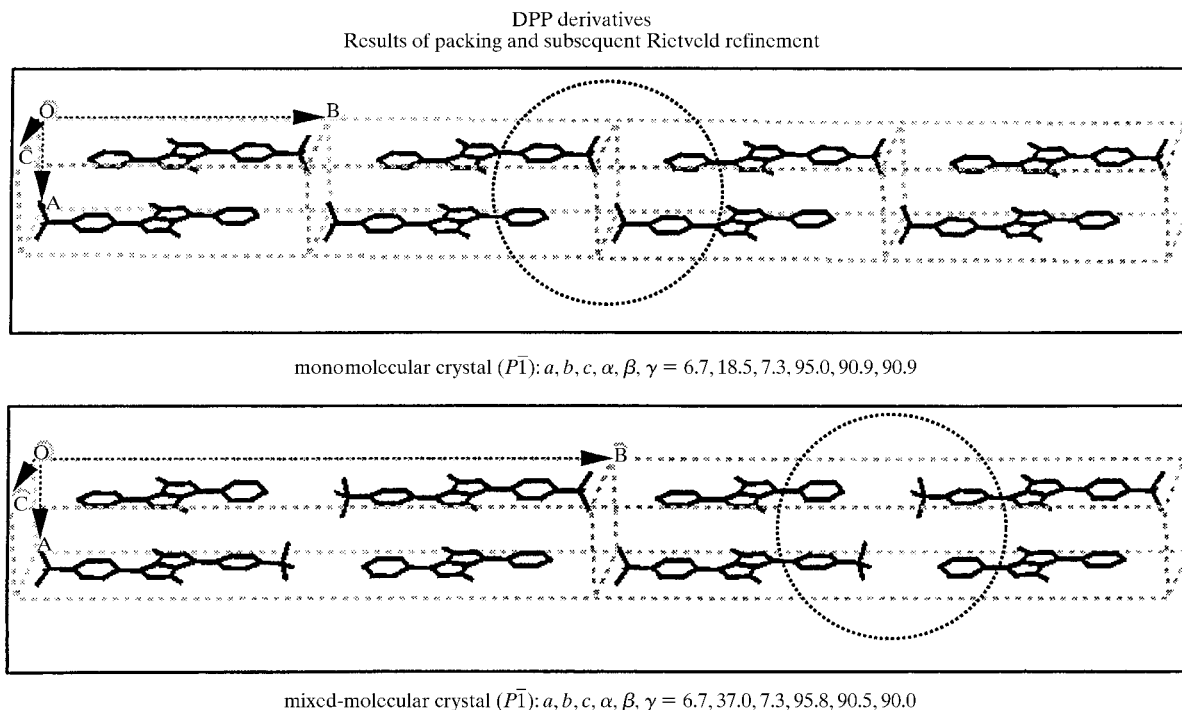


Fig. 6. Plot of the elucidated monomolecular (upper part) and bimolecular (lower part) crystals. In both cases the space group is $P\bar{1}$. Notice that the unit cell of the bimolecular crystal is exactly twice the size of the monomolecular cell. The circles symbolize the distance threshold for interatomic distances ($\sim 13.7 \text{ \AA}$), below which the two crystals are indistinguishable.

independently synthesized latent pigments by co-hydrolysis in a common reaction vessel. The disclosed products had excellent pigmentation properties not expected for mixtures of different pure pigment crystals. In the physical mixture of two pure crystalline pigments, the components show their individual behaviour upon exposure to light, often resulting in marked changes of hue as one pigment component fades more rapidly than the other. Other undesirable properties of a physical pigment mixture result from phase separation in the embedding polymer matrix owing to the different migration characteristics of the individual pigments. The local crystal similarity exhibited by the DPP pigments and the concept of boundary-preserving isometry presented above deliver a logical and consistent expla-

nation for these unexpected properties of what has been disclosed as 'mixtures' in the cited patent. We can also reverse our argumentation and state that the unexpected properties of the disclosed pigment mixture provide strong evidence for the existence of the predicted $(1)_n(2)_n(3)_m$ solids.

6. Is local similarity a unique feature of DPP derivatives?

More than a dozen cases of non-unique X-ray powder patterns are known or can be deduced from the patent literature (Hao *et al.*, 1996; Mizuguchi *et al.*, 1997; Rochat *et al.*, 1983). All are DPP derivatives consisting of pairs of symmetrically substituted and the corresponding asymmetrically substituted biphenylpyrrolopyrroles. We know no other example. Is it sheer coincidence exhibited only by DPP derivatives? Can the dogma of unique powder patterns, although weakened, still be used for practical patenting purposes? To answer this question we performed packing calculations, as described above, on a number of molecules that might give rise to similar packing as the DPP derivatives. Three representative molecules are shown in Scheme (II). Molecules (5)–(10) are homologues of DAWSUP

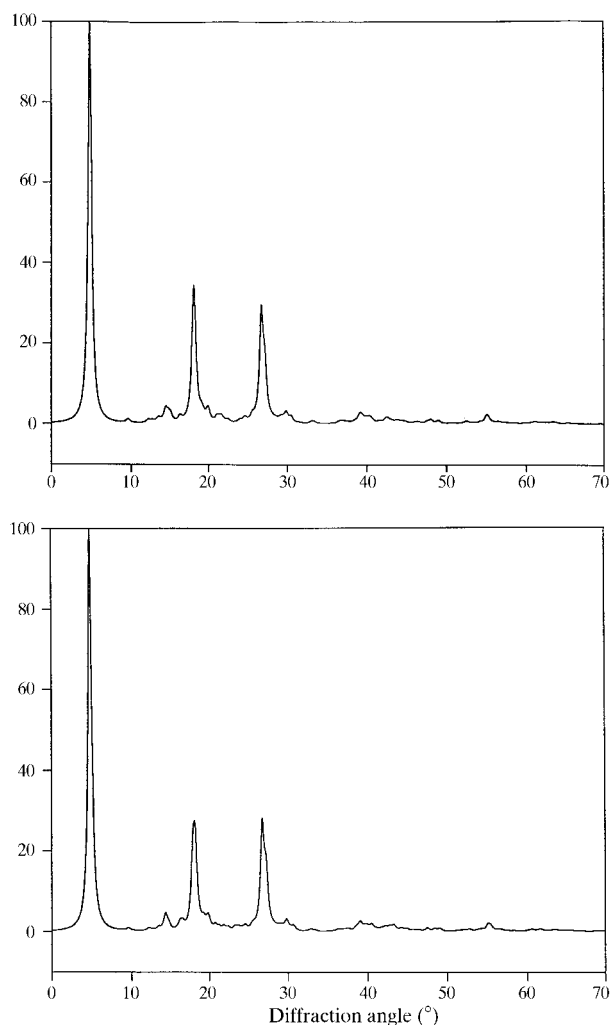


Fig. 7. The simulated X-ray powder pattern of the Rietveld refined crystals [molecules (1)+(2) upper part and molecule (3) lower part]. The simulation was made with the *Cerius2* program assuming some instrumental broadening. The assumed X-ray source is copper (wavelength 1.54178 Å), peak shape Lorentzian. This diagram should be compared to the measured one in Fig. 5.

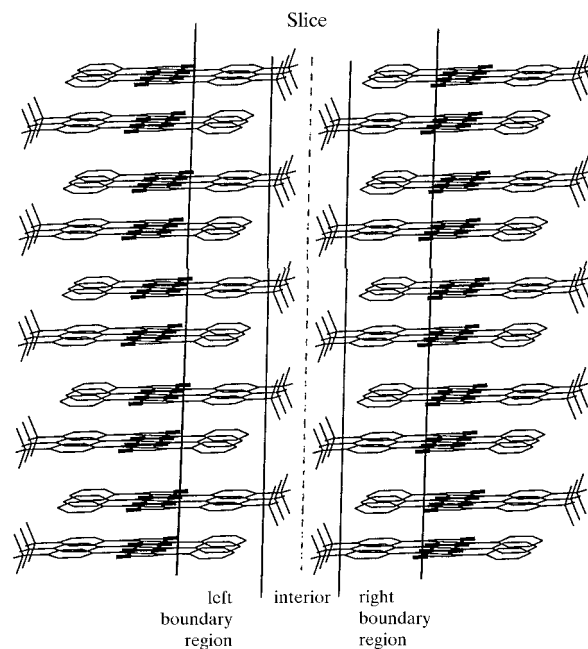
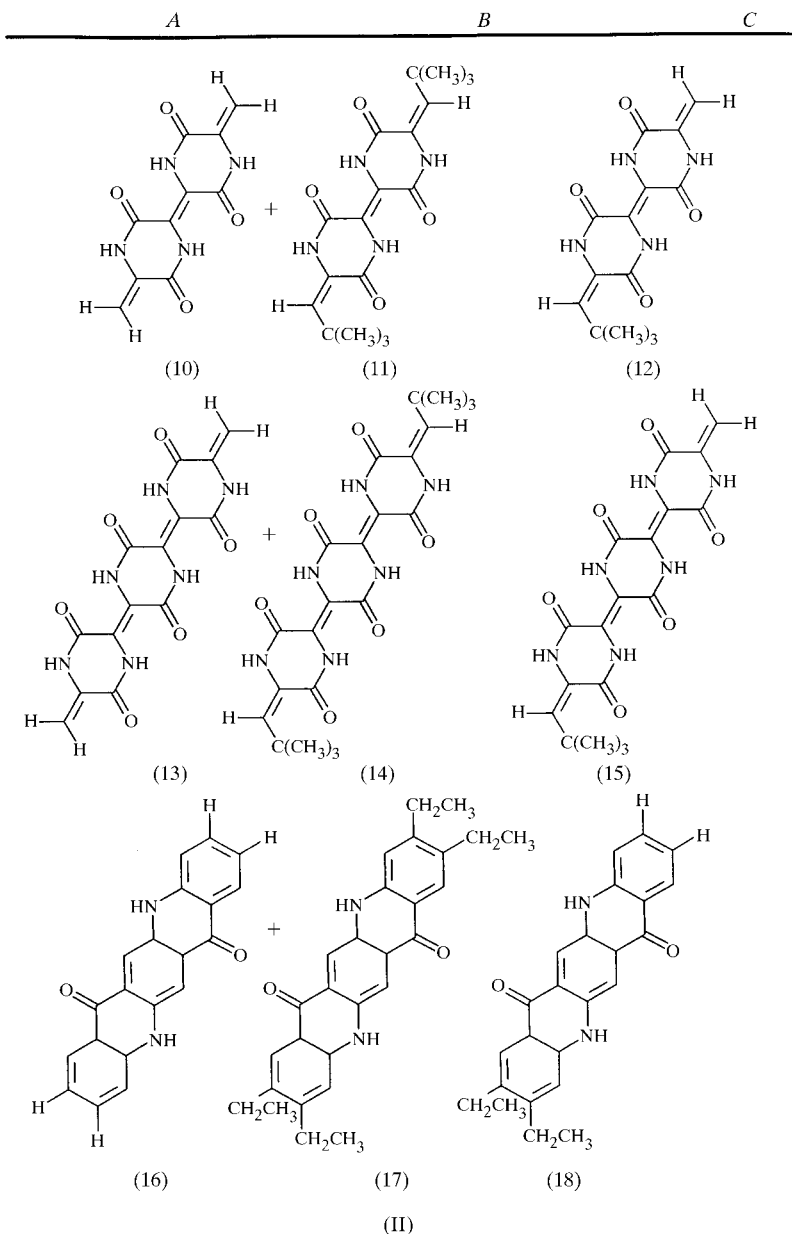


Fig. 8. Use of the Rietveld refined crystal of molecule (3) to explain the sufficient condition for local similarity. Here, a slice of the crystal is shown. An isometric operation which is not a symmetry operation on the whole ($P\bar{1}$) crystal is a mirror operation about a plane in the middle of the slice (dashed line). When the mirror operation is applied only to the indicated slice, the boundary regions of the slice are mapped onto themselves, whereas the interior region will transform a pair of the same molecule to two different molecules.

(Cambridge Structural Database code) treated as in Karfunkel *et al.* (1993). Molecules (16)–(19) are quina-ridone derivatives. Packing calculations were made for the bimolecular crystals containing a pair of symmetrically substituted molecules [first two columns in

most stable modifications there were pairs of crystals exhibiting local similarity. The packing in all cases was similar to the elucidated DPP crystals above. The crystals are built up of slices in which the coplanar molecules are connected in one direction by very strong multiple



Scheme (II)] and the monomolecular crystal of the asymmetrically substituted molecule (third column). Whether the bimolecular crystal will be formed or whether the molecules prefer to crystallize as two distinct monomolecular crystals could not be answered by the packing calculations, but there is strong evidence based on the analogy to known DPP crystals that the bimolecular crystal is the preferred form. Among the

hydrogen bonds and in the other direction they are stacked in a parallel manner (see Fig. 8). The slices are connected (face to face) in the third direction *via* hydrophobic interactions. The extent of non-uniqueness of the powder patterns is shown in Fig. 9 for the first two molecular systems, where the simulated powder patterns and the corresponding difference plot (curve at the bottom of each box in Fig. 9) are shown. As one can see,

the powder patterns for the first molecular system (left-hand side of Fig. 9) consisting of molecules with only two six-membered rings, are similar but not identical. When the degree of local similarity is increased as for the larger molecules with three six-membered rings (right hand side of Fig. 9), the powder patterns become practically indistinguishable under normal experimental conditions.

The results of the packing calculations suggest that local similarity of crystals could be exhibited by many elongated molecules bearing substituents at the ends. The main reason for the lack of examples is that nobody looked for such a phenomenon. The predicted or elucidated crystals, presented so far, satisfy the sufficient condition of Theorem 1 in a rather strict manner, *i.e.* the atom positions of the boundary regions deviate from their counterpart by 0.1–0.2 Å after applying the boundary-preserving isometric transformation.

With the present search software supplied by the Cambridge Crystallographic Data Centre it is not yet possible to perform a systematic search on the crystal database for candidates satisfying the sufficient condition for local crystal similarity. Two groups of industrially important pigments, together with some other

molecules, are shown in Scheme (III). We consider first the azobenzenes (19).

Crystal structures of molecules with various substituents have been published by Maginn *et al.* (1993). These donor–acceptor molecules have been shown to adopt a dimer arrangement with parallel molecular planes. The bulky *R* substituents prevent, however, structures that fulfil the minimum condition (Theorem 1) in a strict manner. The most promising candidate would be that with $R^1 = R^2 = R^3 = R^4 = \text{H}$, but for this molecule no crystal data has been published. The crystal for the molecule with $R_1 = \text{methyl}$, $R^2 = R^3 = R^4 = \text{H}$ is close to fulfilling the condition, as can be deduced from Fig. 10. The boundary regions of the slice are very narrow and only a quantitative analysis shows that the condition is not fulfilled. Another interesting class of industrially important compounds are the squaraines (20). Many unsymmetrically substituted squaraines are mentioned by Law & Bailey (1993), but no crystal data for these compounds have been published. However, for the symmetric molecule (21) or similar molecules, for which crystal data exist (Farnum *et al.*, 1974; Bernstein & Goldstein, 1988), the relative orientation of the molecules suggests that the unsymmetrical molecules (20)

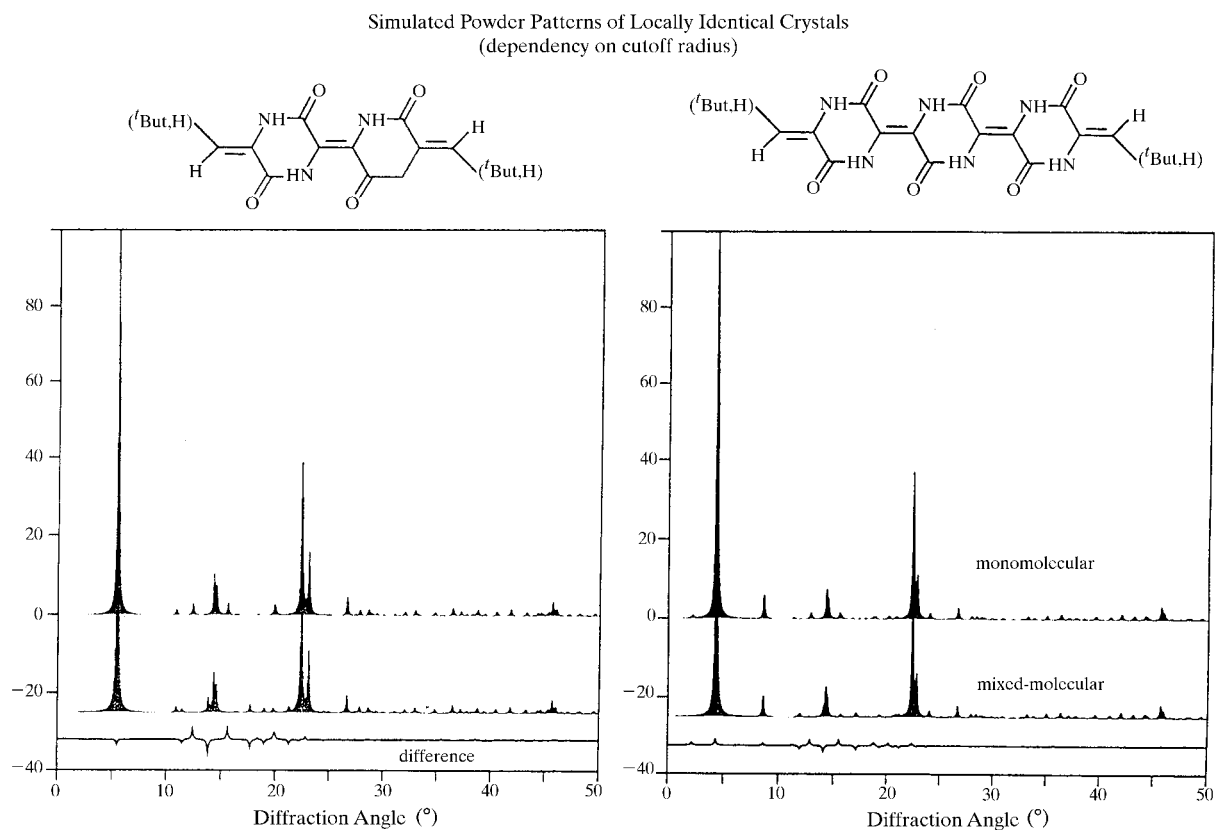
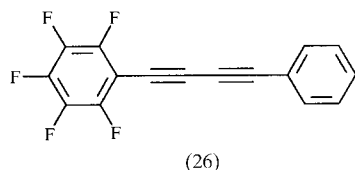
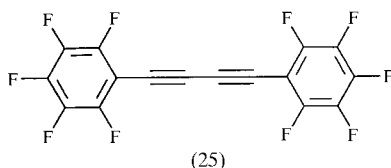
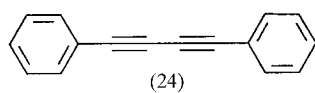
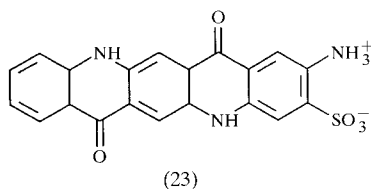
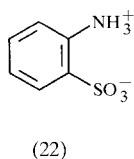
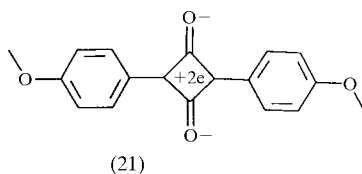
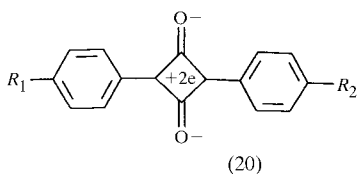
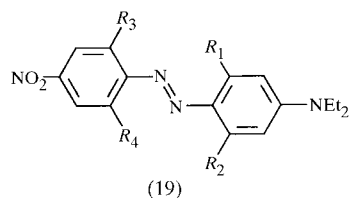


Fig. 9. Simulated powder patterns of two pairs of locally similar crystals predicted by packing calculations on crystals derived from molecules (10)–(15). In each box the powder patterns and the difference plot are shown. The two pairs differ by the truncation distance *R* or the degree of local similarity.

may fulfil the condition since the intermolecular interactions have been shown to be governed by strong interactions between the four-membered rings (Law & Bailey, 1993). Another known crystal, where the condition is not strictly (but nearly) fulfilled is orthanilic acid, (22) (Hall & Maslen, 1967). Using the concept of synthons we may conclude that the quinacridone derivative (23) might form crystals which fulfil the sufficient condition for local similarity. In fact, (23) has the synthons of the previously predicted quinacridone derivative (18) and the synthons of orthanilic acid (22).



(III)

As a last example we would like to mention the recent article of Coates *et al.* (1997), dealing with the interesting case of molecules (24)–(26). The packing of the molecules in the bimolecular crystal consisting of (24)+(25) is very similar to the packing of the monomolecular crystal of (26) and both crystals are triclinic (space group $P\bar{1}$).

7. Summary

Two new and closely related concepts for solids, termed ‘local similarity’ (or ‘synthomorphism’) and ‘boundary-preserving isometry’, have been introduced in a general

manner. Precise mathematical definitions (Definitions 1 and 2) for these concepts have been supplied. The relation between these new concepts and isostructuralism has been discussed. The structures of ‘locally similar’ solids are indistinguishable if only short-range interatomic distances below a certain threshold value are taken into account. A ‘boundary-preserving isometry’ transforms mathematically a solid to a new solid with the same short-range structure. A theorem (Theorem 1), which relates ‘local similarity’ to ‘boundary-preserving isometry’, has been formulated

and the sketch of a proof has been given. Thus, the existence of a ‘boundary-preserving isometry’ is a sufficient condition for ‘local similarity’ and allows one to predict the potential existence of infinitely many solids, locally similar to a given solid. The chemical implication of (high degree) local similarity in connection with the similarity of X-ray powder patterns have been stated. The manifestation of the presented theoretical concepts in nature is given by a series of pigments derived from DPP. For some pigments, poor quality X-ray powder patterns do exist. These powder patterns were not indexable so that a new technique based on the

Table 1. Fractional coordinates of Rietveld refined crystals of (1)+(2) and (3)

Molecules (1)+(2), space group $P\bar{1}$; lattice: 6.684 36.996 7.278 95.500 90.500 89.855. Molecule (3), space group $P\bar{1}$; lattice: 6.673 18.543 7.296 95.300 90.900 90.930.

Molecules (1)+(2)	x	y	z
O1	0.74179	0.79197	1.09711
O2	0.24614	0.70868	-0.15043
O3	0.22574	0.78333	0.58145
O4	0.76221	0.71587	0.36849
N5	0.74835	0.77563	0.48108
N6	0.23959	0.72377	0.46580
N7	0.24241	0.76861	-0.03541
N8	0.74553	0.73186	0.98518
C9	0.24566	0.73181	-0.02012
C10	0.74228	0.76860	0.96792
C11	0.71627	0.85546	0.82935
C12	0.27163	0.64458	0.11363
C13	0.23472	0.76058	0.45062
C14	0.75322	0.73888	0.49825
C15	0.27518	0.60643	0.09782
C16	0.71270	0.89361	0.84313
C17	0.24761	0.70280	0.3004
C18	0.74032	0.79692	0.64546
C19	0.23999	0.78987	0.12949
C20	0.74796	0.71030	0.82125
C21	0.74746	0.734869	0.6917
C22	0.24048	0.76501	0.25749
C23	0.24636	0.72767	0.17315
C24	0.74158	0.77234	0.77422
C25	0.22661	0.94819	0.18883
C26	0.24513	0.84862	0.31546
C27	0.74284	0.65123	0.63824
C28	0.23525	0.82870	0.14680
C29	0.75272	0.67148	0.80600
C30	0.24038	0.88683	0.33027
C31	0.74761	0.61302	0.62546
C32	0.73574	0.83566	0.66198
C33	0.25217	0.66405	0.28188
C34	0.76676	0.65225	0.96165
C35	0.22121	0.84826	-0.00766
C36	0.23899	0.64420	0.43431
C37	0.74892	0.85518	0.50834
C38	0.21964	0.88648	0.00668
C39	0.76835	0.61407	0.94936
C40	0.74337	0.89329	0.52241
C41	0.24452	0.60608	0.41821
C42	0.26176	0.58714	0.25015
C43	0.72611	0.91254	0.68963
C44	0.22947	0.90672	0.17606
C45	0.75853	0.59348	0.78089
C46	0.11094	0.96526	0.03477
C47	0.44255	0.96414	0.18823
C48	0.12369	0.96339	0.36415
C49	0.14109	0.53609	0.39165
C50	0.17473	0.53594	0.04479
C51	0.48661	0.53747	0.24996
C52	0.26600	0.54921	0.23416

Molecule (3)

O1	0.25918	0.41552	-0.12584
O2	0.23183	0.56740	0.60691
N3	0.24680	0.44785	0.49088
N4	0.25435	0.53581	-0.01039
C5	0.25745	0.46222	0.00467
C6	0.27263	0.28815	0.13731

Table 1 (cont.)

	x	y	z
C7	0.24206	0.52140	0.47590
C8	0.27650	0.21177	0.12133
C9	0.25638	0.40530	0.32536
C10	0.25038	0.57889	0.15475
C11	0.24966	0.52959	0.28271
C12	0.25631	0.45461	0.19805
C13	0.23720	0.89575	0.21553
C14	0.24254	0.69688	0.34147
C15	0.24546	0.65656	0.17229
C16	0.23781	0.77340	0.35651
C17	0.26114	0.32776	0.30649
C18	0.24432	0.69530	0.01767
C19	0.25623	0.28868	0.45960
C20	0.24280	0.77173	0.03242
C21	0.26201	0.21236	0.44319
C22	0.27133	0.17383	0.27424
C23	0.23976	0.81273	0.20231
C24	0.12978	0.92041	0.04486
C25	0.45587	0.92535	0.22906
C26	0.12021	0.92519	0.38467

ab initio prediction of possible polymorphs with a subsequent Rietveld refinement was used for the elucidation of their crystal structures. Extensive *ab initio* packing calculations on other molecules have been made and it has been shown that the phenomenon of

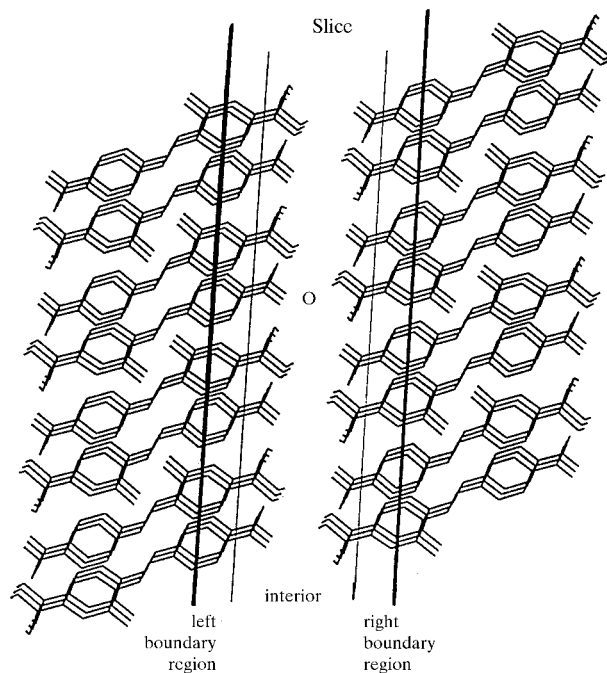


Fig. 10. A view of a slice of the crystal derived from (19) ($R^1 = \text{methyl}$, $R^2 = R^3 = R^4 = \text{H}$, space group $P2_1/c$). The sufficient condition is not strictly fulfilled, although the degree of violation is not large. The chosen view direction is deceptive and may suggest that the condition is fulfilled *via* the inversion centre symbolized by the small circle.

local similarity may not be restricted to DPP derivatives. Attempts to detect boundary-preserving isometries in crystals given in the CSD were made in an unsystematic manner. Although no example was found which fulfils the sufficient condition in a manner comparable to the elucidated DPP crystals or the presented theoretical examples based on *ab initio* packing, cases violating the condition in a moderate manner have been reported.

References

- Bernstein, J. & Goldstein, E. (1988). *Mol. Cryst. Liq. Cryst.* **164**, 213–229.
- Coates, G. W., Dunn, A. R., Henling, L. M., Dougharty, D. A. & Grubbs, R. H. (1997). *Angew. Chem.* **109**, 290–293.
- Desiraju, G. R. (1995). *Angew. Chem.* **107**, 2541–2558.
- Desiraju, G. R. (1996). *Acc. Chem. Res.* **29**, 441–449.
- Desiraju, G. R. (1997). *Science*, **278**, 404–405.
- Farnum, D. G., Neuman, M. A. & Suggs, W. T. (1974). *J. Cryst. Mol. Struct.* **4**, 199–212.
- Gdanitz, R. J. (1992). *Chem. Phys. Lett.* **190**, 391–396.
- Hall, S. R. & Maslen, E. N. (1967). *Acta Cryst.* **22**, 216–228.
- Hao, Z., Iqbal, A., Medinger, B. & Wallquist, O. (1996). European patent EP 0 704 497/A1.
- Kálmán, A., Argay, Gy., Scharfenberg-Pfeiffer, D., Höhne, E. & Ribár, B. (1991). *Acta Cryst.* **B47**, 68–77.
- Kálmán, A., Párkányi, L. & Argay, Gy. (1993). *Acta Cryst.* **B49**, 1039–1049.
- Karfunkel, H. R. & Gdanitz, R. J. (1992). *J. Comput. Chem.* **13**, 1171–1183.
- Karfunkel, H. R. & Leusen, F. J. J. (1992). *SPEEDUP J.* **6/2**, 43–50.
- Karfunkel, H. R., Leusen, F. J. J. & Gdanitz, R. J. (1993). *J. Comput.-Aided Mat. Des.* **1**, 177–185.
- Karfunkel, H. R., Wu, Z. J., Burkhard, A., Rihs, G., Sinnreich, D., Buerger, H. M. & Stanek, J. (1996). *Acta Cryst.* **B52**, 555–561.
- Law, K. Y. & Bailey, F. C. (1993). *Dyes Pigments*, **21**, 1–12.
- MacDonald, J. C. & Whitesides, G. M. (1994). *Chem. Rev.* **94**, 2383–2420.
- Maginn, S. J., Bullock, J. F. & Docherty, R. (1993). *Dyes Pigments*, **23**, 159–178.
- Mizuguchi, J. (1998). *Acta Cryst.* **C54**, 1482–1484.
- Mizuguchi, J., Grubenmann, A. & Rihs, G. (1993). *Acta Cryst.* **B49**, 1056–1060.
- Mizuguchi, J., Grubenmann, A., Wooden, G. & Rihs, G. (1992). *Acta Cryst.* **B48**, 696–700.
- Mizuguchi, J., Hao, Z., Wallquist, O. & Iqbal, A. (1997). European Patent EP 0 764 696/A2.
- Payne, R. S., Roberts, R. J. & Docherty, R. (1997). *J. Comput. Chem.* **19**, 1–20.
- Reddy, D. S., Ovchinnikov, Y. E., Shishkin, O. V., Struchkov, Y. T. & Desiraju, G. R. (1996). *J. Am. Chem. Soc.* **118**, 4085–4089.
- Rochat, A. C., Cassar, L. & Iqbal, A. (1983). European Patent EP 0 094 911/A1.
- Rosenberg, P. D. (1975). *Patent Law Fundamentals*. New York: Clark Boardman Corporation.
- Rutherford, J. S. (1997). *Acta Chim. Hung.* **134**, 395–405.
- West, A. R. (1984). *Solid State Chemistry and its Applications*. New York: John Wiley and Sons.
- Wolff, J. J. (1996). *Angew. Chem.* **108**, 2339–2341.
- Young, R. A. (1995). *The Rietveld Method*. Oxford University Press.
- Zambounis, J. S., Hao, Z. & Iqbal, A. (1997). *Nature*, **338**, 131–132.