Acta Crystallographica Section B
Structural Science
ISSN 0108-7681

Spatial arrangement of molecules in homomolecular $Z' = 2$ structures

Elna Pidcock

Cambridge Crystallographic Data Centre, 12 Union Road, Cambridge CB2 1EZ, England

Correspondence e-mail: pidcock@ccdc.cam.ac.uk

Received 14 September 2005
Accepted 3 January 2006

The Box Model of crystal packing describes unit cells in terms of a limited number of arrangements of molecular building blocks. An analysis of $Z' \leq 1$ structures has shown that cell dimensions are related to molecular dimensions in a systematic way and that the spatial arrangement of molecules in crystal structures is very similar, irrespective of Z or space group. In this paper it is shown that the spatial arrangement of molecules in $Z' = 2$ structures are, within the context of the Box Model, very similar to that found for $Z' \leq 1$ structures. The absence of crystallographic symmetry does not appear to affect correlations between molecular dimensions and cell dimensions, or between the packing patterns and the positions of molecules in the unit cell, established from the analysis of $Z' \leq 1$ structures. The preference shown by $Z' = 2$ structures for low surface-area packing patterns and the observation that strong energetic interactions are most often found between the large faces of the independent molecules reaffirms the importance of molecular shape in crystal packing.

1. Introduction

A search of the CSD (Vol. 5.26, November 2004 and three updates; Allen, 2002) using *ConQuest* (Bruno *et al.*, 2002) shows that 27 362 of the 320 253 structures (8.5%) with three-dimensional coordinates contain more than one molecule in the asymmetric unit. Homomolecular $Z' = 2$ structures (the two molecules of the asymmetric unit are chemically identical) account for 16 067 structures (5.0%). Our understanding of when $Z' > 1$ structures will be observed or why they form is limited. It has been suggested that molecules of an awkward shape or molecules for which there is a conflict between close-packing requirements, and those of intermolecular interactions, tend to form structures where $Z' > 1$ (Steed, 2003, and references therein). A selection of previous work concerning $Z' = 2$ structures covers topics such as space-group distributions with respect to Z' (Steiner, 2000) and the identification of pseudosymmetric or hypersymmetric interactions between molecules of the asymmetric unit (for examples see Zorky, 1996; Britton, 2000; Marsh *et al.*, 1998). Here we extend an analysis of the spatial arrangement of molecules, performed on structures with a maximum of one molecule in the asymmetric unit to $Z' = 2$ homomolecular structures. Inspired by Kitaigorodskii (1961) and his theory of close packing we have looked for, and found, regularity in the spatial arrangement of molecules in the unit cell.

The unit cell is a construct used by the scientific community to describe a crystal structure. It is not uniquely defined and the same structure can be described by unit cells with quite different parameters. However, in practice, unit cells are often presented in a standard way, so whilst not being unique, they

are at least consistent with one another within a particular space group setting. For example, it was found from a search of the Cambridge Structural Database (Allen, 2002, henceforth CSD) that the cell parameters given for structures in the space-group setting $P12_1/c1$ were those of the reduced cell parameters in 90.4% of cases. Similarly, for structures belonging to $P\bar{1}$ it was found that the cell parameters given were those of the reduced cell in 79.1% of cases. More importantly, however, the unit cell is a building block from which the entire crystal structure can be constructed and therefore encapsulates all information required to understand the composition of the crystal structure. Advancing our understanding of the structure of the unit cell, for example, understanding the spatial arrangement of molecules, the relevance of the position of molecules and the interaction between molecules in the unit cell will only help increase our understanding of crystal structures. The Box Model of crystal packing (Pidcock & Motherwell, 2003; Pidcock & Motherwell, 2004*a*) is a new model describing the structure of unit cells. By understanding the basic construction of a unit cell it is hoped a new framework is provided upon which to build an understanding of the complex intermolecular interactions that result in crystal structures.

The Box Model (Pidcock & Motherwell, 2003; Pidcock & Motherwell, 2004*a*) provides a very simple representation of molecular crystal structures. By likening a molecule to a box with three unequal dimensions, models of unit cells can, literally, be built by stacking boxes with faces touching and edges aligned. For example, with two boxes, three close-packed arrangements are possible: the largest faces of the boxes may be placed in contact, the medium faces may be in contact, or the smallest faces may be in contact (Fig. 1). These arrangements, called packing patterns, are models for the possible geometries of unit cells where $Z = 2$. The packing patterns represent idealized unit cells where the 'molecules' are aligned perfectly with the unit-cell axes and where the layers of molecules touch but do not interpenetrate. An extensive survey of molecular crystal structures, where $Z' \leq 1$ and $Z = 2$, 4 or 8 has shown that packing patterns represent a good approximation to the structures of unit cells (Pidcock & Motherwell, 2004*a*).

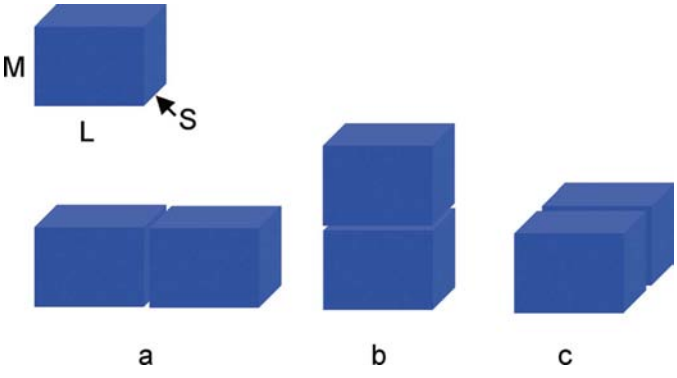

Figure 1
A box of three unequal dimensions, L, M and S where $L > M > S$ represents a molecule. Arrangements of a given number of boxes, with faces touching and edges aligned, represent idealized unit cells. There is a limited number of close-packed arrangements or packing patterns for a given number of boxes. For example, for two boxes the three possible packing patterns are shown: (*a*) 112*L*, (*b*) 112*M*, (*c*) 112*S*.

The packing patterns, for a given number of boxes, have the same volume, but the external surface area depends on which faces of the boxes are in contact within the pattern. Therefore, for two boxes, the packing pattern with the minimum surface area is found when the largest faces of the boxes are in contact. The packing pattern where the smallest faces are in contact has the maximum surface area. Experimental cells have been shown to be predominantly of low surface area where the large faces of the molecules are related by symmetry and not unit-cell translations (Pidcock & Motherwell, 2003; Pidcock & Motherwell, 2004*a*). This preference for low surface-area packing patterns indicates that molecular shape is of primary importance in crystal packing and that there is some driving force to minimize surface area for a given volume. The results obtained from this simplified description of molecular crystal packing are very much in accordance with the work of Kitaigorodskii (1961), in particular his theory of close packing. Indeed, the regularity found in the spatial arrangement of molecules, within the context of the Box Model, has led us to 'parameterize' close packing (Pidcock & Motherwell, 2004*c*).

We have applied the framework provided by the Box Model of crystal packing to $Z' = 2$ structures as a first step to understanding the phenomenon of $Z' = 2$ structures. If the Box Model truly encapsulates the fundamentals of crystal packing, as we suggest, then it should be applicable to structures with more than one molecule in the asymmetric unit. By establishing some empirical observations about these structures, it may be possible to build an understanding of them.

2. Calculation details

For a box of three unequal dimensions, long (L), medium (M) and short (S), the pattern coefficients of the Box Model are the integer multiples which describe the overall dimensions of the packing patterns, in terms of the box dimensions. Thus, the pattern coefficients of the family of packing patterns shown in Fig. 1 are 1, 1 and 2, because the dimensions of the arrays of boxes can be described as 1 box × 1 box × 2 boxes. The packing-pattern names describe the relationships between the box dimensions and the pattern coefficients. Thus, packing pattern 112*S* is so named because it describes the arrangement of boxes given by $1 \times L$, $1 \times M$ and $2 \times S$. The unique pattern coefficient of the packing pattern is given at the end along with the relevant box dimension. So, the packing pattern 112*L* describes the arrangement given by $1 \times M$, $1 \times S$ and $2 \times L$. When there are four boxes, there are six possible close-packed arrangements: three members of the 221 family (221*L*, 221*M*, 221*S*) and three members of the 114 family (114*L*, 114*M*, 114*S*). For eight boxes, the possible packing patterns are 222, the 421 family (six members) and the 118 family (three members). This nomenclature is described in more detail in earlier publications and is used throughout this paper.

A search of the CSD (Version 5.26, November 2004 and two updates; Allen, 2002) for structures where $Z' = 2$, with only one chemical species in the structure (three-dimensional coordinates required, *R*-factor less than 0.05, but no further filters applied), was performed. The file of structures was then processed by *RPluto* (Motherwell *et al.*, 1999) using the SHAPE command. This command calculates the three molecular dimensions, *L*, *M* and *S* where $L > M > S$, using the principal axes of inertia, and determines to which cell axis each molecular dimension is most closely aligned. The three pattern coefficients (ratio of cell axis to molecular dimension) are calculated for each molecule. A program was then written to process the output from *RPluto* to determine which packing pattern represented the best fit to the pattern coefficients calculated for each molecule in the cell. The Euclidian distance metric was used to calculate the fit between the calculated pattern coefficients and target pattern coefficients, as described in an earlier publication (Pidcock & Motherwell, 2004*a*). Only structures displaying space groups $P1$, $P\overline{1}$, $P2_1$, $C2$, Pc, Cc, $P2_1/c$, $C2/c$, $P2_12_12$, $P2_12_12_1$, $Pna2_1$, $Pca2_1$ and $Pbca$ were included. When the two molecules of the asymmetric unit were oriented the same way with respect to the cell axes, the two packing patterns assigned to the structure were the same and these molecules form the 'aligned' data set. A total of 320 structures which were found to have more than two residues in the asymmetric unit (either due to disorder or due to occupancy of special positions) were removed from the 'aligned' dataset. When the molecules were oriented differently with respect to the cell axes, then two different packing patterns were returned.

For structures where $Z' = 2$ the two molecules presented by the crystallographer are picked arbitrarily and do not imply structural importance or energetically strong interactions. Thus, the strongest energetic interaction between the two independent molecules, *A* and *B*, was used to identify a new, energetically significant asymmetric unit. To identify the energetically important interaction between *A* and *B*, structures were processed by *RPluto*. The positions of any missing H atoms were calculated using ideal geometry about C and N atoms. The crystal-packing potential energy was calculated by summing molecular interactions, about a central reference molecule, using the atom-pair empirical potentials of Gavezzotti (1994). The elements included in the calculation of intermolecular potentials were C, H, N, O, S, Cl and F; Br was approximated by S and all other elements made no contribution. The Gavezzotti potentials that describe common hydrogen-bonding interactions were used and approximations were made for potentials of hydrogen bonds not included by Gavezzotti. The default maximum of 200 atoms per asymmetric unit was applied to ensure tractable potential energy calculations. Each structure gives a list of intermolecular interactions sorted into order by energy, strongest first, and a flag that identifies whether the interactions are between *A*–*A*, *A*–*B* or *B*–*B*. The strongest energetic interaction between *A* and *B* was used to identify the energetically important asymmetric unit. Also returned by the calculations is a vector describing the geometric relationship between the two interacting molecules. This vector is used to identify the faces of the independent molecules which are interacting. The orientations of the molecules in the cell are known with respect to the *x*, *y* and *z* components of the vector. By dividing each component of the vector by the relevant molecular dimension, the spatial arrangement of the independent molecules is described in terms of multiples of *L*, *M* and *S*, the molecular dimensions. To estimate which faces of the independent molecules are in closest proximity, the vector component with the largest multiple is chosen. For example, if the vector between the centroids of the independent molecules was given by $1L$, $0M$ and $0S$, it is clear that the smallest (*MS*) faces of the independent molecules are in closest proximity. Similarly, if the vector is given by $0L$, $0M$ and $1S$, the largest (*LM*) faces of the molecules are interacting. In general, independent molecules of the unit cell are not aligned perfectly and hence the geometric description of the energetically significant asymmetric unit is an approximation to the real geometry. These calculations were performed only for structures belonging to the 'aligned' data set.

During the calculation of the crystal-packing potential energy, a centre of mass is calculated for each molecule of the asymmetric unit. This information is extracted from the output of *RPluto*. To generate the scatterplots of molecular centres in this paper, the centre of mass of each molecule is positioned in a cell 0–1, 0–1, 0–1 using shifts of 1 and −1.

3. Results and discussion

A search has been made of the CSD (Allen, 2002) using *ConQuest* (Bruno *et al.*, 2002) for structures where $Z' = 2$ and where the two molecules belonging to the asymmetric unit are chemically identical. Structures where molecules were residing on special positions were excluded from the dataset, not because the Box Model was inapplicable but because the coordinates of the molecules were then fixed (at the special positions), influencing the subsequent analysis of the spatial arrangement of the molecules in the cell. The packing patterns (of the Box Model) for each structure were calculated, using both molecules of the asymmetric unit, hence each structure was initially described by two packing patterns. When the molecules were oriented in the same way with respect to the unit-cell axes then the two packing patterns returned for the structure were the same. Of the 6565 structures containing two whole and chemically identical molecules in the asymmetric unit (structures with disorder were excluded) and belonging to space groups $P1$, $P\overline{1}$, $P2_1$, $C2$, Pc, Cc, $P2_1/c$, $C2/c$, $P2_12_12$, $P2_12_12_1$, $Pna2_1$, $Pca2_1$ and $Pbca$, 4934 structures (75%) fell into the category of structures where the molecules are oriented the same with respect to the cell axes. These structures were collated to form the 'aligned' dataset and it is these data which are the focus of this paper. The work of Sona & Gautham (1992) and Gautham (1992) showed that the symmetry-independent molecules belonging to structures where $Z' = 2$ have, in general, similar conformations. It was found that in 91.9% of structures belonging to the 'aligned' dataset, the two molecules of the asymmetric unit were

described by the same molecular dimensions, within a tolerance of 1 Å on each dimension. The remaining 1631 structures (of the original dataset of 6565 structures) return different packing patterns. It was found that the proportion of structures where the molecules were described with the same molecular dimensions decreased marginally in this dataset (81.4% of structures contained molecules with the same molecular dimensions, within a tolerance of 1 Å). Structures belong to this dataset because of the different orientations of the independent molecules with respect to the cell axes and not because the molecules are of different conformations.

3.1. Pattern coefficients

The assignment of structures to a packing pattern requires the calculation of pattern coefficients for each structure (Pidcock & Motherwell, 2003, 2004*a*). The pattern coefficients are calculated by dividing each cell-axis length by the molecular dimension most closely aligned with it (see §2). When $Z' \leq 1$, each structure is described by three pattern coefficients. For $Z' = 2$ structures belonging to the 'aligned' dataset, pattern coefficients are calculated for, and a packing pattern assigned to, each molecule of the asymmetric unit. However, because the independent molecules are aligned the same with respect to the unit-cell axes, only a single packing pattern is needed to describe the structure.

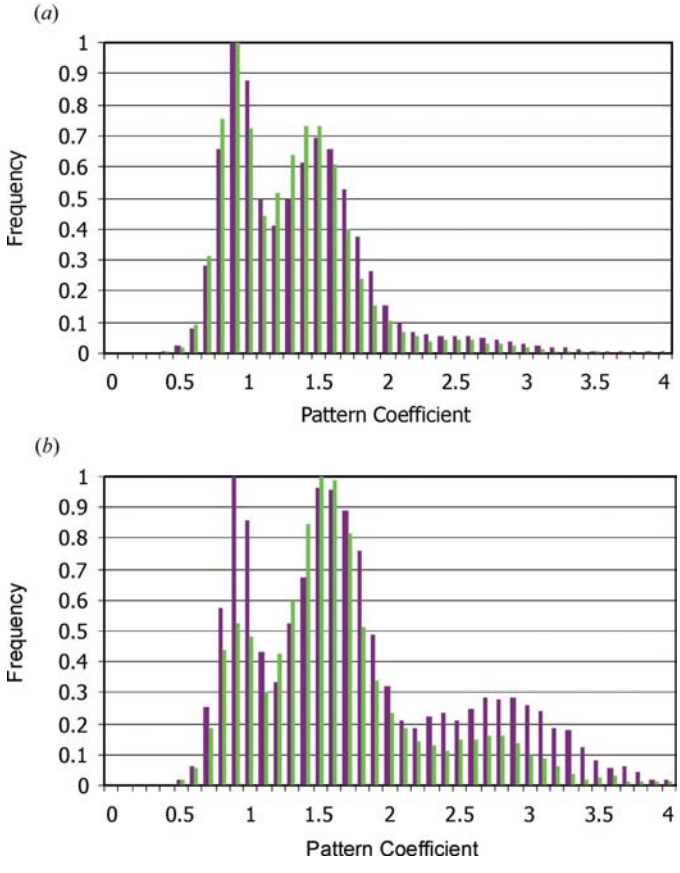

Figure 2
Histograms of calculated pattern coefficients *versus* frequency for (*a*) Z = 4 structures and (*b*) Z = 8 structures. The light-coloured bars represent pattern coefficients calculated for $Z' = 2$ structures (this work) and the dark-coloured bars represent $Z' \leq 1$ structures (Pidcock & Motherwell, 2004*c*). The histograms have been normalized, to aid comparison, such that the most intense peak of each histogram has a value of 1.0.

Pattern coefficients calculated for $Z' \leq 1$ structures, when collated and displayed as histograms (see Fig. 2), demonstrated the regularity in the spatial arrangement of molecules associated with close packing (Pidcock & Motherwell, 2004*c*). Each pattern coefficient calculated for a structure represents the length of a cell axis in terms of one of the molecular dimensions. The clear and distinct peaks in the histograms (Fig. 2) indicate that unit-cell dimensions are related to molecular dimensions in a systematic way. The work on $Z' \leq 1$ structures also showed that histograms plotted for structures belonging to different space groups, but of the same Z, were almost identical to one another, and even histograms representing structures with different Z values had features in common. Therefore, the presence of a peak at a pattern coefficient value of ~ 0.9 in histograms calculated for $Z = 2$, $Z = 4$ and $Z = 8$ structures indicated that one cell length is commonly described by $0.9 \times$ a molecular dimension, L, M or S. This first peak represents the '1-direction' of the packing pattern (as in packing pattern family 22*1*, *11*4, 42*1 etc.*) and is the direction of the unit cell along which molecules are related by translation. The second peak of the histograms was found at a pattern coefficient value of approximately 1.5. Thus, a unit-cell axis length is described by $\sim 1.5 \times L$, M or S. This second peak represents a 2-direction of a packing pattern and hence two molecular centres (related by symmetry when $Z' = 1$) are found along the axis before the unit cell repeats. Unit cells are therefore 'constructed' in accordance with a limited number of templates and these templates are defined by the pattern coefficients. For a more detailed discussion of the form of the pattern coefficient histograms see Pidcock & Motherwell (2004*c*). The histograms of the pattern coefficients calculated for both molecules of the structures belonging to the 'aligned' dataset, when $Z = 4$ and $Z = 8$ are shown in Fig. 2 along with data calculated for $Z' \leq 1$ structures. It can be seen that there is very little difference between the histograms of $Z' = 2$ structures and $Z' \leq 1$ structures, particularly in the position of the peaks. Comparison of the $Z = 8$ histograms shows differences in the relative peak heights between the $Z' = 2$ and $Z' \leq 1$ data sets, but these are due to differences in the populations of the 421 and 222 packing patterns only. The similarity of the $Z' = 2$ histograms with the $Z' \leq 1$ structures indicates the fundamental rules governing the spatial arrangement of the molecules in crystal structures remain the same, whether the interactions between molecules are mediated by crystallographic symmetry or not.

3.2. Positions of molecular centres

In previous work, for structures with $Z' \leq 1$, it was found that the positions of the molecular centres in the unit cell were correlated with the packing-pattern assignments (Pidcock & Motherwell, 2004*b*). For a cell axis which corresponds to a 4

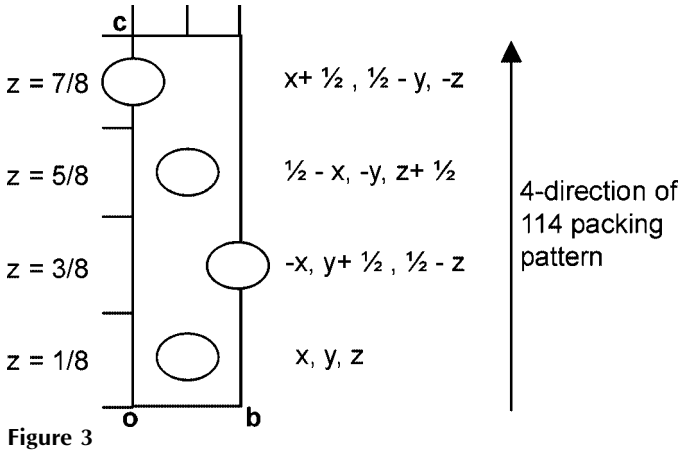

Figure 3
Schematic of a unit cell described by a 114 packing pattern in $P2_12_12_1$. The cell, in terms of the Box Model, is described as a stack of $1 \times 1 \times 4$ molecules. The 4 direction of the packing pattern is aligned with the unit-cell c axis, and the view is down **a**, a 1 direction of the packing pattern. The diagram illustrates the action of the symmetry operators on a molecule residing at $x, \frac{1}{2}, \frac{1}{8}$.

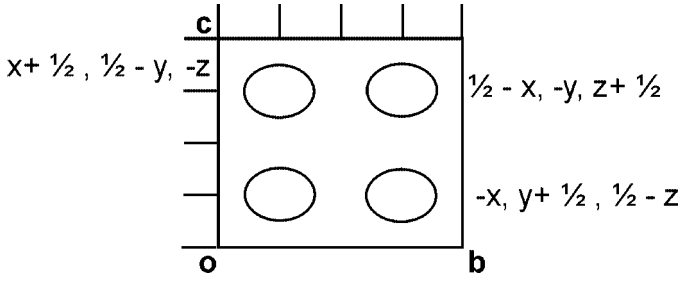

Figure 4
Schematic of a unit cell described by a 221 packing pattern in $P2_12_12_1$. The cell, in terms of the Box Model, is described by an array of $2 \times 2 \times 1$ molecules. The view is down the unit-cell a axis, in the 1 direction of the packing pattern. The diagram illustrates the action of the symmetry operators on a molecule residing at $x, \frac{1}{4}, \frac{1}{4}$.

direction of a 114 packing pattern, for example, the molecular centres were found to cluster at a value of $\frac{1}{8}$ on that axis. Therefore, when a unit cell is constructed of four molecules stacked one on top of each other a molecular centre is commonly found at $\frac{1}{8}$ on the cell axis, which runs parallel to the stack of four molecules (the 4 direction). For a cell axis which corresponds to a 2 direction of a packing pattern, for example in a 112 packing pattern or a 221 packing pattern, the molecular centres were often found at a position of $\frac{1}{4}$. It is not a coincidence that positions of $\frac{1}{8}$ and $\frac{1}{4}$ are favoured. An examination of the action of the symmetry operators along a particular axis demonstrates why clusters of data are observed at these values. The action of the symmetry operators belonging to $P2_12_12_1$ on any of the three equivalent axes, from a starting coordinate p on a particular axis, generates positions at p, $\frac{1}{2} - p$, $p + \frac{1}{2}$, $1 - p$. Therefore, when $p = \frac{1}{8}$, positions are generated at $\frac{1}{8}$, $\frac{3}{8}$, $\frac{5}{8}$ and $\frac{7}{8}$: four positions evenly spaced along the axis, see Fig. 3. When $p = \frac{1}{4}$, positions are generated at $\frac{1}{4}$ and $\frac{3}{4}$ on that axis: 2 positions are evenly spaced along the axis. The accommodation of a 221 packing pattern in $P2_12_12_1$ is illustrated in Fig. 4. The scatterplots of the positions of molecular centres belonging to $Z' \leq 1$ revealed a physical basis for the Box Model; the different packing patterns observed are due, at least in part, to the position of the molecule with respect to the symmetry operators.

For structures with more than one molecule in the asymmetric unit, the relationships between the molecules of the asymmetric unit are not described by space-group symmetry. Therefore, for a $Z' = 2$, $Z = 4$ structure, belonging to the 114 family, the positions of the 4 molecules along an axis are not defined by 4 symmetry operators, rather the action of two symmetry operators on two unique molecules. It is of interest to examine whether the absence of symmetry in $Z' = 2$ structures affects the correlations established between packing patterns and positions of molecular centres. In previous work

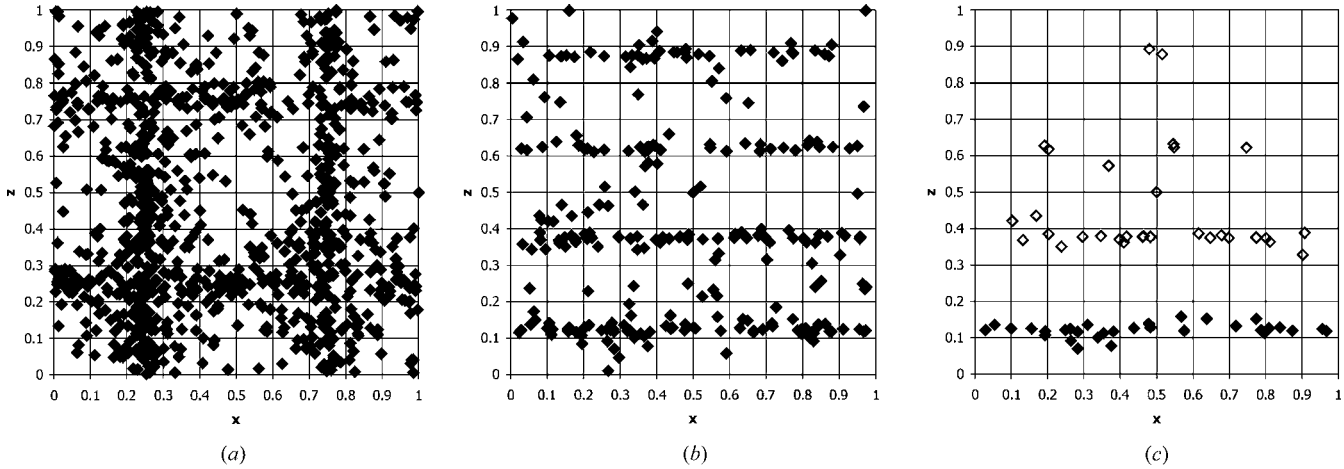

Figure 5
Distribution of molecular centres determined for both molecules of the asymmetric unit in $Z' = 2$, $Z = 4$ structures. (*a*) The positions of molecular centres for structures belonging to the 221 packing pattern family. The '1 direction' is oriented with the unit-cell b axis and the molecular centres are projected onto the xz plane. (*b*) The positions of molecular centres of structures belonging to the 114 packing pattern family where the '4 direction' of the packing pattern is aligned with the unit-cell c axis. Clusters of data at $z = \frac{1}{8}, \frac{3}{8}, \frac{5}{8}$ and $\frac{7}{8}$ can be seen. (*c*) Members of the 114 packing pattern family, where the '4 direction' is aligned with the unit-cell c axis. The position of one molecular centre is constrained to be within 0.12 ± 0.5 (coloured black) and the positions of the second molecule are shown (open diamonds).

for $Z' \leq 1$ structures, the positions of the molecular centres were normalized, using allowed shifts of the space group to place the molecular centre within 0–0.5 on all axes. In this paper, since the datasets contain structures belonging to many different space groups (thus different combinations of shifts of $\frac{1}{2}$ are allowed), the positions of the centres have been placed between 0 and 1 on all axes, but no further manipulations have been performed.

A scatterplot of molecular centres, for $Z' = 2$ structures belonging to the 221 packing pattern family, where the '1 direction' of the packing pattern is aligned with the unit cell b axis, is shown in Fig. 5(a). These structures belong, predominantly, to space groups $P\overline{1}$ and $P2_1$. The molecular centres are projected onto the xz plane, and the view is down y. It is clear from the scatterplot that the molecular centres of $Z' = 2$ structures are not randomly distributed. The clustering of the molecular centres at positions of $\frac{1}{4}$ is consistent with observations made for $Z' \leq 1$ structures. As mentioned above, a position of $\frac{1}{4}$ is often observed for a molecule when there is a '2 direction' of the packing pattern, along the corresponding axis. For $Z' = 2$ structures, despite the fact that a '2 direction' of the unit cell is constructed by the two independent molecules, positions of $\frac{1}{4}$ remain popular. Thus, the absence of a symmetry relationship does not alter the close-packing 'constraints'. The scatterplot shown in Fig. 5(b) shows the distribution of molecular centres for $Z' = 2$ structures belonging to the 114 packing pattern family. The 4 direction of the packing pattern is aligned with the unit-cell c axis and clusters of data are clearly seen at positions of $\frac{1}{8}$, $\frac{3}{8}$, $\frac{5}{8}$ and $\frac{7}{8}$ on the z axis. However, in the case of $Z' = 2$ structures belonging to space groups such as $P2_1$ and $P\overline{1}$, the symmetry operators do not generate 4 positions along the axis, only 2 positions. Thus, the molecular centre of one molecule is located at $\frac{1}{8}$ and the centre of the second independent molecule is found at $\frac{3}{8}$ or $\frac{5}{8}$ (Fig. 5c). Therefore, the two independent molecules adopt positions which 'mimic' those dictated by symmetry in $Z' = 1$ structures. In the absence of symmetry, $Z' = 2$ structures belonging to 114 packing patterns resemble very closely $Z' = 1$, 114 structures.

Similar observations were made for $Z = 8$ structures. For structures assigned to the 222 packing pattern, clusters of data were found at values of $\frac{1}{4}$ (Fig. 6a), as expected. For structures assigned to the 421 packing pattern family, when the 4 direction of the packing pattern was aligned with the unit cell c axis, for example, clusters of data were found at values of $z = \frac{1}{8}, \frac{3}{8}, \frac{5}{8}$ and $\frac{7}{8}$, as expected (Fig. 6b). In addition, clusters of data are observed at values of $z = 0$ and $z = \frac{1}{4}$; this observation will be discussed in more detail below. However, molecular centres for $Z' = 2$, $Z = 8$ structures are found in positions expected from similar studies performed for $Z' \leq 1$ structures. To reiterate the point made earlier, the absence of formal symmetry relationships between independent molecules of an asymmetric unit does not appear to affect the spatial arrangements adopted by the molecules in the unit cell.

3.2.1. *A–B–B–A* structures and *A–B–A–B* structures. The previous section has shown that the molecular centres of $Z' = 2$ structures are not randomly distributed through the unit cell, but follow patterns, consistent with the packing-pattern assignment, which were observed for $Z' = 1$ structures. However, the existence of two independent molecules in the asymmetric unit introduces some flexibility into how unit cells belonging to a packing pattern are constructed. For example, a closer look at $Z' = 2$ structures that belong to the 114 packing pattern reveals two types of structures, an *A–B–B–A* structure where the independent molecules *A* and *B*, and their symmetry generated counterparts, form layers (Fig. 7) and a structure where the independent molecules alternate in the stack of four molecules, *A–B–A–B*. The reason for the existence of these two types of structures is intimately related to the symmetry operators of the space groups. All 114 structures belonging to $P\overline{1}$ exhibit the *A–B–B–A* structure (Fig. 7a). Molecule *A* resides at the position $\frac{1}{8}$ (on any of the equivalent axes) and the action of the symmetry operator ($p \rightarrow 1 - p$) dictates that the symmetry-generated molecule resides at $\frac{7}{8}$. Molecule *B* of the asymmetric unit is then found at the position $\frac{3}{8}$ (see Fig. 5c) and the symmetry-generated molecule resides at $\frac{5}{8}$. Hence, the *A–B–B–A* structure. This type of layering of molecules is only possible on axes where the symmetry operators transform coordinate p, to $1 - p$, and where $p \simeq \frac{1}{8}$. Therefore, *A–B–B–A* structures are possible in

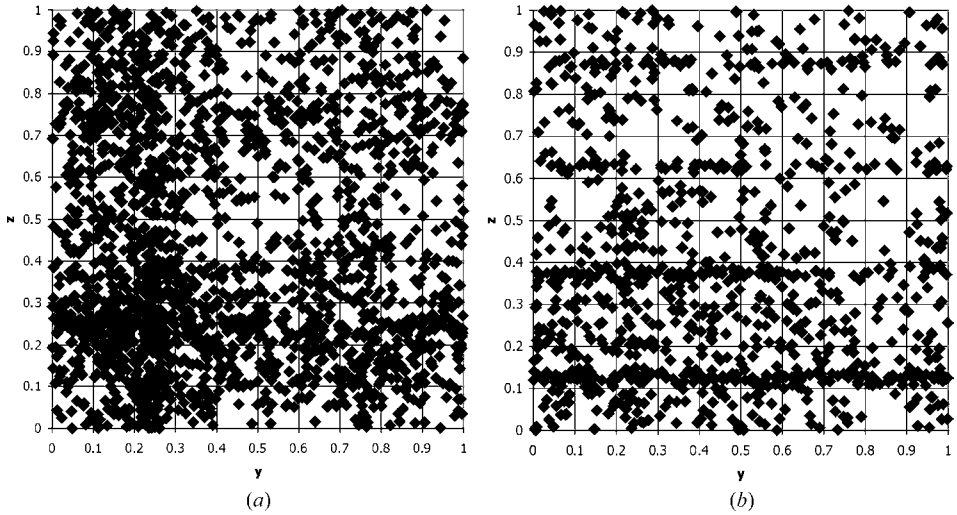

Figure 6
Distribution of molecular centres, determined for both molecules of the asymmetric unit in $Z' = 2$, $Z = 8$ structures. (a) The positions of molecular centres for structures belonging to the 222 packing pattern projected onto the yz plane. (b) The positions of molecular centres for structures belonging to the 421 packing pattern family, where the '4 direction' is aligned with the unit-cell c axis. Clusters of data at values of $z = \frac{1}{8}$ (and equivalent positions) can be seen.

space group $P2_1$ when the '4 direction' of the packing pattern is aligned with the a (or c) axis (Fig. 7b). However, in $P2_1$, for structures where the '4 direction' of the packing pattern is aligned with the unit-cell b axis, the screw axis transforms the y coordinate of molecule A, to $y + \frac{1}{2}$, and so a starting position of $\frac{1}{8}$ generates a symmetry-equivalent molecule at $\frac{5}{8}$. The second independent molecule, B, from a starting position of $\frac{3}{8}$, generates a symmetry-equivalent molecule at $\frac{7}{8}$: an A–B–A–B type structure (Fig. 7c).

The same type of observations are made of the $Z = 8$ structures. A–B–B–A and A–B–A–B structures are found and relationships between the type of structure and the symmetry operators and position of the molecule in the cell are observed. For structures assigned to the 421 packing pattern family (a structure assigned to a 421 packing pattern has a unit cell described by an array of $4 \times 2 \times 1$ molecules) belonging to $P2_1/c$, there are many permutations of the position of the molecular centre with the cell axis which are possible, and these affect the type of structure formed. When the '4 direction' of the packing pattern is aligned with the unit-cell a axis, since p transforms to $1 - p$ only, A–B–B–A columns are generated. The action of the symmetry operators on a molecule residing at $\frac{1}{8}, y, \frac{1}{4}$ is shown in Fig. 8. In addition, as only two positions are generated in x for each molecule, the two A–B–B–A columns within the unit cell must be aligned and a layered structure such as that shown in Fig. 11(a) results. The relationships are more complicated when the '4 direction' of the packing pattern is aligned with the b or c axes. Only A–B–A–B columns are possible, irrespective of whether the coordinate for molecule A is located at 0 or $\frac{1}{8}$ (see Figs. 9 and 10). However, the coordinate of molecule A (along the cell axis corresponding to the '4 direction') does affect how the A–B–A–B columns of the unit cell are related to one another. Thus, if molecule A is found at a position of 0 on the relevant axis, then the A–B–A–B columns are aligned in the unit cell and a layered structure like that shown in Fig. 11(b) results. The schematic of Fig. 9 shows the action of the symmetry operators on the symmetry-independent molecules. If, however, the molecular centre of A is at $\frac{1}{8}$ then, since four positions are generated for each independent molecule along the axis (Fig. 10), a structure which looks like a chequerboard is generated (Fig. 11c). The same relationships are true for $P2_12_12_1$. All axes in $P2_12_12_1$ are equivalent and all axes can accommodate either two or four molecules. Therefore, structures belonging to the 421 packing patterns with aligned A–B–A–B layers are found when molecule A resides at 0 (on the axis corresponding to the 4 direction; Fig. 11d) and chequerboard

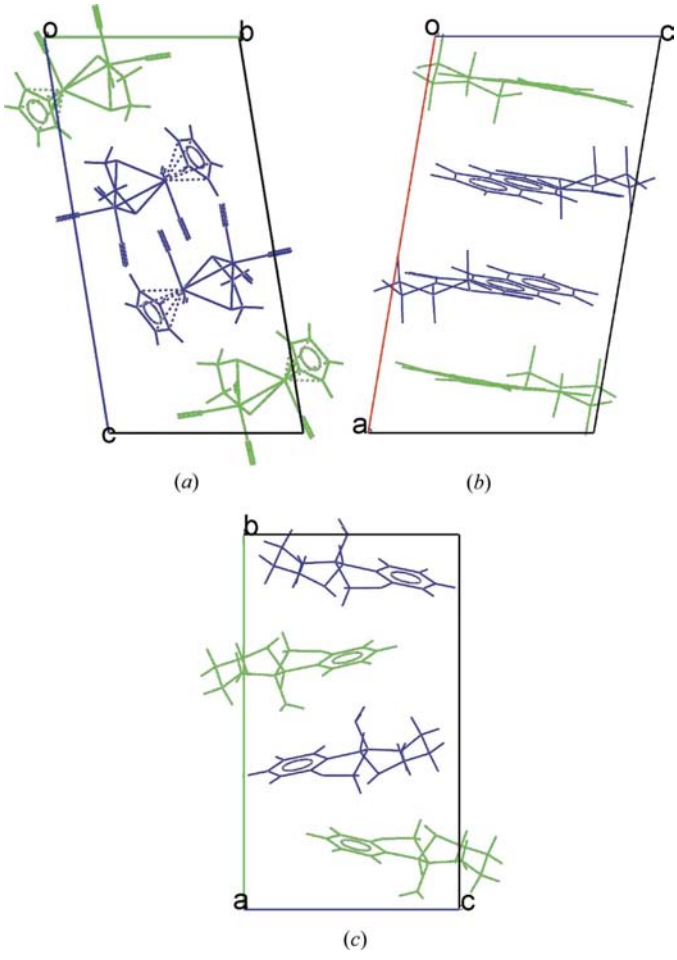

Figure 7
Examples of A–B–B–A and A–B–A–B structures found in $Z' = 2$, $Z = 4$ cells. (a) An A–B–B–A structure in $P\overline{1}$ (BAQXAT; Adams *et al.*, 2003). (b) An A–B–B–A structure in $P2_1$ (CIZXIS; Beurskens *et al.*, 1984). The '4 direction' of the packing pattern is aligned with the unit-cell a axis. (c) An A–B–A–B structure in $P2_1$ (AHUFUE; Pavé *et al.*, 2003), the '4 direction' of the packing pattern is aligned with the unit-cell b axis.

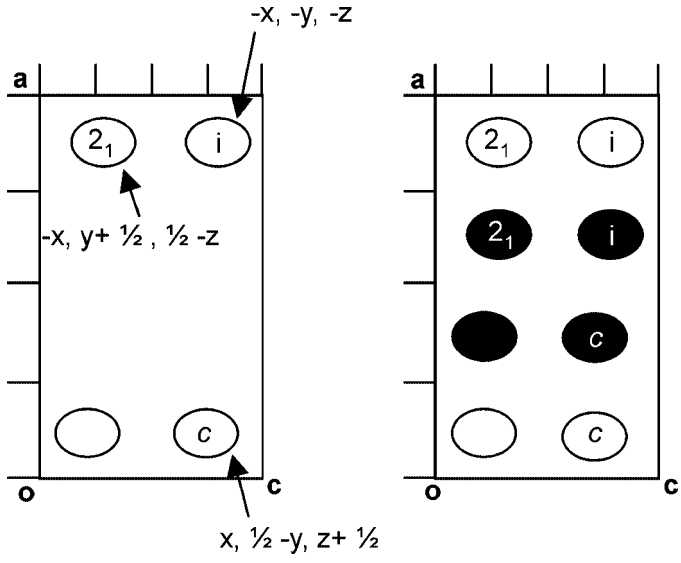

Figure 8
Schematic of a unit cell described by a 421 packing pattern in $P2_1/c$. The 4 direction of the packing pattern is aligned with the unit-cell a axis and the view is down **b**, the 1 direction of the packing pattern. Left: From a starting position of $\frac{1}{8}, y, \frac{1}{4}$, the actions of the symmetry operators generate the distribution of molecular centres shown for molecule A (white ellipses). The molecule generated by the inversion operator is labelled i, the molecule generated by the action of the screw axis operator is labelled 2_1 and the molecule generated by the action of the glide plane is labelled c. Right: From a starting position of $\frac{3}{8}, y, \frac{1}{4}$, molecular centres, generated by the symmetry operators for molecule B (black ellipses) are shown. The resultant, aligned, A–B–B–A columns (along **a**) can be seen.

Table 1
Distribution of $Z' = 2$ structures belonging to the 'aligned' dataset over the packing patterns of the Box Model.

Packing pattern	Number of structures	% of same Z dataset
$Z = 2$		
112*L*	28	9.4
112*M*	70	23.4
112*S*	201	67.2
$Z = 4$		
221*L*	1162	50.8
221*M*	536	23.4
221*S*	299	13.1
114*L*	24	1.0
114*M*	52	2.3
114*S*	215	9.4
$Z = 8$		
222	1132	51.8
1*L*2*M*4*S*	299	13.7
1*L*4*M*2*S*	201	9.2
2*L*1*M*4*S*	239	10.9
2*L*4*M*1*S*	120	5.5
4*L*1*M*2*S*	90	4.1
4*L*2*M*1*S*	72	3.3
118*L*	2	0.1
118*M*	8	0.4
118*S*	24	1.1
$Z = 16$		
224*S*	50	31.3
224*M*	41	25.6
224*S*	24	15.0
441*L*	20	12.5
441*M*	8	5.0
441*S*	6	3.8
148 family	11	6.9

structures with alternating *A*–*B*–*A*–*B* columns are found when molecule *A* resides at $\frac{1}{8}$ (Fig. 11*e*).

3.3. Distribution of structures over packing patterns

In the Box Model formalism, the volume for a given number of close-packed boxes is constant, but the surface area depends on the packing pattern. For two boxes, the packing pattern with the largest external surface area has the smallest faces of the boxes in contact (112*L*) and the packing pattern with the lowest external surface area has the largest faces of the boxes in contact (112*S*), see Fig. 1. It has been shown in previous work that structures for which $Z' \leq 1$ are found more frequently in the low surface-area packing patterns (Pidcock & Motherwell, 2004*a*). The distributions of structures belonging to the 'aligned' dataset over the packing patterns of the Box Model are given in Table 1. The low surface-area patterns, 112*S*, 221*L*, 222, for example, are well populated and represent 52.2% of structures where $Z' = 2$ and $Z \leq 8$. A comparison of the occupancy of packing patterns for $Z' \leq 1$ and $Z' = 2$ structures shows that a greater percentage of structures are found in low surface-area packing patterns when $Z' = 2$. Hence, comparing the occupancy of $Z = 4$, with the 221*L* pattern, we find 36.5% when $Z' \leq 1$ and 52.0% when $Z' = 2$. Similarly, for $Z = 2$, the occupancy of 112*S* is 52.0% when $Z' \leq 1$ and 68.0% when $Z' = 2$. For $Z = 8$, the occupancy of the lowest surface-area packing pattern, 222, is 27.8% when $Z' \leq 1$ and 51.8% when $Z' = 2$. In contrast to the preference for low surface-area packing patterns shown by $Z' = 2$ structures, it is only in the $Z' = 2$ dataset that structures belonging

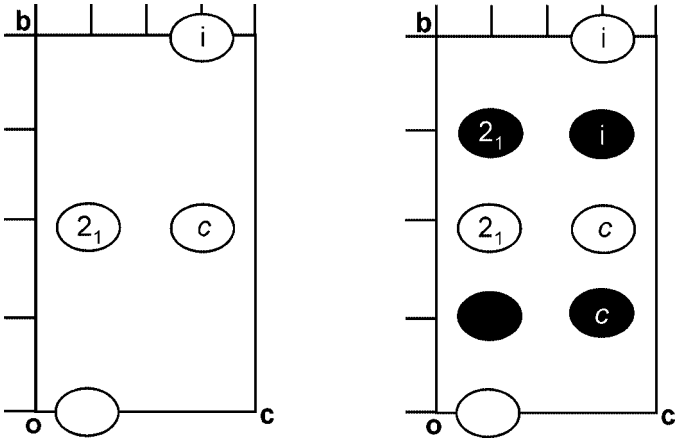

Figure 9
Schematic of a unit cell described by a 421 packing pattern in $P2_1/c$. The 4 direction of the packing pattern is aligned with the unit-cell *b* axis and the view is down **a**, the 1 direction of the packing pattern. Left: From a starting position of $x, 0, \frac{1}{4}$, the actions of the symmetry operators generate the distribution of molecular centres shown for molecule *A* (white ellipses). The molecule generated by the inversion operator is labelled i, the molecule generated by the action of the screw axis operator is labelled 2_1 and the molecule generated by the action of the glide plane is labelled *c*. Right: The action of the symmetry operators on the second molecule of the asymmetric unit, molecule *B* (black ellipses), from a position of $x, \frac{1}{4}, \frac{1}{4}$, results in the distribution of molecular centres shown. The resultant, aligned, *A*–*B*–*A*–*B* columns (along **b**) can be seen.

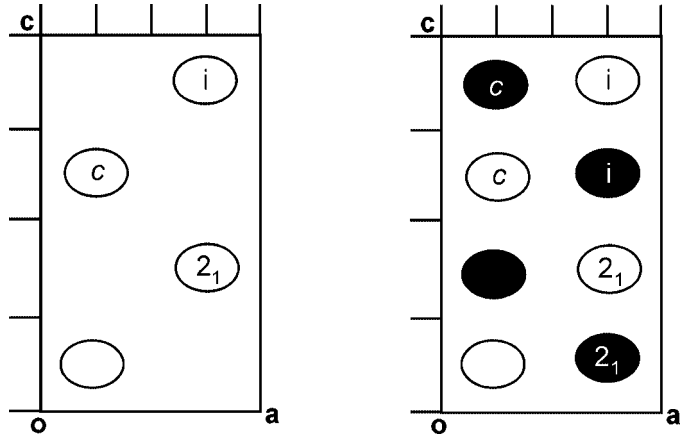

Figure 10
Schematic of a unit cell described by a 421 packing pattern in $P2_1/c$. The 4 direction of the packing pattern is aligned with the unit-cell *b* axis and the view is down **a**, the 1 direction of the packing pattern, as in Fig. 9. Left: From a starting position of $x, \frac{1}{8}, \frac{1}{4}$, the actions of the symmetry operators generate the distribution of molecular centres shown for molecule *A* (white ellipses). The molecule generated by the inversion operator is labelled i, the molecule generated by the action of the screw axis operator is labelled 2_1 and the molecule generated by the action of the glide plane is labelled *c*. Right: The action of the symmetry operators on the second molecule of the asymmetric unit, molecule *B* (black ellipses), from a position of $x, \frac{3}{8}, \frac{1}{4}$, results in the distribution of molecular centres shown. The resultant *A*–*B*–*A*–*B* columns (along **b**) are not aligned with each other, but rather form a 'chequerboard'-type structure, as shown.

to the 118 packing pattern are observed. An example is shown in Fig. 12. This high surface-area packing pattern is by no means popular, representing less than 2% of the $Z = 8$ dataset, but its manifestation in experimental crystal structures serves as a further validation of the Box Model.

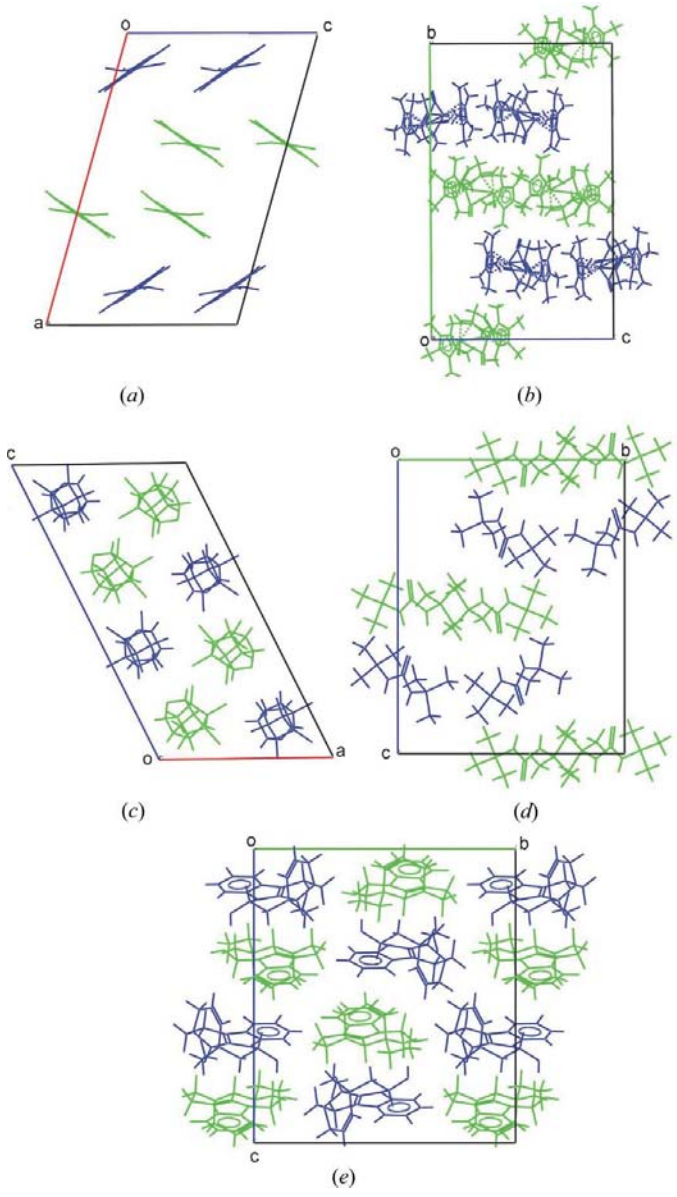

Figure 11
Examples of *A*–*B*–*B*–*A* and *A*–*B*–*A*–*B* structures for $Z = 8$ cells belonging to $P2_1/c$ and $P2_12_12_1$, where *A* and *B* are the independent molecules. The interaction of the molecular centres with the symmetry operators generates these two structural types. (*a*) An *A*–*B*–*B*–*A* type structure in $P2_1/c$ (EHOWIH02; Aakeröy *et al.*, 2003). The '4 direction' of the packing pattern is aligned with the *a* axis. (*b*) An *A*–*B*–*A*–*B* containing structure in $P2_1/c$ (SAYXIZ; Lewandos *et al.*, 1988) where the '4 direction' of the packing pattern is aligned with the unit-cell *b* axis. The molecular centres are found at 0 and $\frac{1}{4}$ on the *b* axis and the *A*–*B*–*A*–*B* columns are aligned leading to a layered structure. (*c*) A 421 structure (BEXGUG; ApSimon *et al.*, 1982) in $P2_1/c$ where the '4 direction' of the packing pattern is aligned with the unit-cell *c* axis. The *A*–*B*–*A*–*B* columns are not aligned leading to a chequerboard pattern. (*d*) An *A*–*B*–*A*–*B* containing structure in $P2_12_12_1$ (EYOMAG; Gel'mbol'dt *et al.*, 2003). The molecular centres of the independent molecules are found at 0 and $\frac{1}{4}$ on the *c* axis, leading to the layered structure observed. (*e*) An *A*–*B*–*A*–*B* containing structure in $P2_12_12_1$ (BOJVIF; Rodier *et al.*, 1982). The centres of the independent molecules are found at $\frac{1}{8}$ on the *c* axis, leading to a chequerboard pattern.

In general, however, it appears from the distribution of $Z' = 2$ structures over the packing patterns that $Z' = 2$ structures have more freedom to adopt low surface-area packing patterns. Alternatively, when the restraints 'imposed' by crystallographic symmetry are lifted, the driving force to minimize surface area for a given volume appears to be more effective.

3.4. Geometry of the energetically significant asymmetric unit

In order to understand the preference for low surface-area packing patterns it was decided to examine how the independent molecules in the cell were related to one another. In previous work, energetic interactions were calculated for a set of $Z' \leq 1$ structures belonging to $P2_1/c$ (Pidcock & Motherwell, 2005). It was found, for these structures, that the interaction between the largest faces of the molecules was frequently ranked highly in terms of energy. Thus, the preference for low surface-area packing patterns was supported by the observation of energetically strong interactions between the large faces of the molecule. In the case of $Z' = 2$ structures, do the strongest energetic interactions between the independent molecules involve the largest faces of the molecules? Using *RPluto* (Motherwell *et al.*, 1999) the crystal packing potential for all structures in the 'aligned' dataset was calculated. Interactions between a central molecule and all the surrounding molecules were calculated and it was straightforward to extract the strongest energetic interaction between the two independent molecules of the unit cell. Thus, a new asymmetric unit was defined, one where the strongest energetic interaction between the two independent molecules was found. A vector between the centroids of the interacting molecules was also returned during the calculation, defining the positions of the molecules with respect to one another. Consequently, it was possible, since the orientations of the molecules in the unit cell are known, to determine which faces of the molecules were involved in the important energetic interaction (see §2). This asymmetric unit can be described in terms of the faces of the two independent molecules which are interacting, the *LM* (largest faces), the *LS* (medium faces) or the *MS* (smallest faces). An alternate way to describe the energetically significant asymmetric unit is to use a packing pattern of the Box Model for two molecules. Hence, all structures can be described in terms of three patterns, 112*S* (largest faces interact), 112*M* (medium faces interact) or 112*L* (smallest faces interact).

In a previous study of the packing interactions of crystallographically independent molecules in organic crystals, Karthe *et al.* (1993) found that the strongest energetic interaction was between the independent molecules of the unit cell in 10 out of the small sample of 18 structures studied (55.6%). In this study it was found that the strongest energetic interaction in the crystal structure was between the independent

molecules in the unit cell in 2729/4934 (55.3%) of structures. An interaction between the independent molecules of the unit cell was present in the top three strongest interactions in 77.4% of cases. These findings support the proposal made by Karthe *et al.* (1993) that whilst the occurrence of some $Z' = 2$ structures may be due to the formation of a dimer or oligomer in solution which crystallizes as such, evidence suggests this is not true in all cases.

The description of the energetically significant asymmetric unit of all structures belonging to the 'aligned' dataset is given in Table 2. The calculations that identify the pair of molecules which compose the energetically significant pair were performed using only the coordinates of the molecules in the cell and no information regarding the packing pattern of the structure was required. Of the 4475 structures included in Table 2, 239 (5.3%) were assigned to a packing pattern which is inconsistent with the contents determined for the energetically significant asymmetric unit. For example, 71 structures were assigned to the 221*L* packing pattern for which the energetically significant asymmetric unit was described with the box pattern 112*L*. Thus, a strong energetic interaction was found between the two independent molecules in a direction where there was only a single molecule present, according to the packing pattern assignment. It is likely that these inconsistencies arise due to the choices that are made when determining the alignment of the molecules in the cell. However, these anomalies only represent approximately 5% of the dataset and therefore do not compromise the integrity of the Box Model or the results discussed here.

Examination of which molecular faces are involved in the energetically strong interaction revealed the importance of molecular shape. Of the 4934 structures of the 'aligned' dataset, 2490 (50.5%) were built from an asymmetric unit where the largest faces of the molecules interact (packing pattern 112*S*), 1492 (30.2%) were built from a unit where the medium faces of the molecules interact (112*M*) and only 952 (19.3%) of the structures were built from a unit where an interaction between the smallest faces of the molecules is responsible for defining the energetically significant asymmetric unit (112*L*). If there was no relationship between the energetics of the intermolecular interaction and the size of the molecular faces involved, it would be expected that the strongest energetic interaction would involve each of the molecular faces in $\sim 33\%$ of cases. Clearly this is not so. As observed in the case of $Z' = 1$ structures in $P2_1/c$, strong energetic interactions often result from interactions of the large surfaces of molecules, thus it is not a phenomenon only associated with $Z' = 2$ structures. Rather, the observation of strong energetic interactions between arrangements of molecules with a minimum surface area for a given volume is evidence for close packing, the efficient use of space by molecules. Its prevalence in structures, irrespective of Z', is perhaps an indication of how fundamental close packing of molecules is to crystal nucleation and growth.

Table 2
The distribution of the energetically significant asymmetric units (whose geometry is described with $Z = 2$ packing patterns), over packing patterns for $Z = 4$ and $Z = 8$ structures.

Only the packing patterns for which there is a choice in the geometry of the asymmetric unit are included.

Packing pattern	112*L*	%	112*M*	%	112*S*	%
$Z = 4$						
221*L*	71	6.1	471	40.5	620	53.4
221*M*	178	33.2	47	8.8	311	58.0
221*S*	108	36.1	167	55.9	24	8.0
$Z = 8$						
222	296	26.1	319	28.2	517	45.7
1*L*2*M*4*S*	20	6.7	46	15.4	233	77.9
1*L*4*M*2*S*	9	4.5	103	51.2	89	44.3
2*L*1*M*4*S*	25	10.5	15	6.3	199	83.3
2*L*4*M*1*S*	29	24.2	82	68.3	9	7.5
4*L*1*M*2*S*	51	56.7	4	4.4	35	38.9
4*L*2*M*1*S*	50	69.4	17	23.6	5	6.9

For structures containing more than two molecules in the cell, there is often more than one possible geometry for the energetically significant asymmetric unit. A schematic diagram is given to illustrate, Fig. 13. From Fig. 13 it can be seen that there are two possible asymmetric units for the unit cell: the two molecules enclosed by the dark-coloured rectangle or the two molecules enclosed by the light-coloured rectangle. The calculations described above allow the identification of the asymmetric unit which is energetically more significant. Examples given in Fig. 14 show two members of the 221*S* packing pattern, one where the energetically significant asymmetric unit is described by 112*M* and the *LS* faces of the molecules interact and for the other, the asymmetric unit is described by 112*L* and the *MS* faces of the molecule interact. For a unit cell belonging to the packing pattern 221*L*, the possible geometries of the asymmetric unit are given by 112*S* or 112*M*. It can be seen from Table 2 that when there is a choice of geometry for the asymmetric unit within a packing pattern, the lower surface-area geometry is more common. For example, for $Z = 4$ structures belonging to the packing pattern 221*M*, the possible geometries of the asymmetric unit are given by 112*L* or 112*S*. The lower surface-area asymmetric unit 112*S* represents 58.0% of structures and 112*L* represents 33.2% of structures. The only exceptions to this preference for low surface area are found in the 421 packing pattern family. For these structures it appears that it is more common for the asymmetric unit which represents half of the '4 direction' to represent the bulk of the structures. For

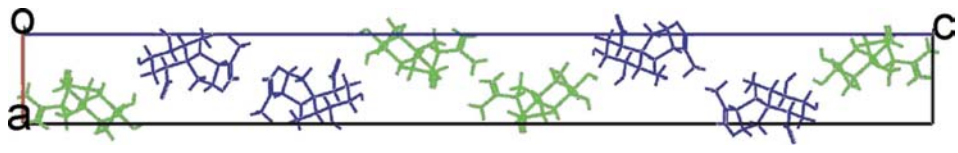

Figure 12
A $Z' = 2$ structure (HXUBIM10; Birnbaum, 1977) which is an example of the rare 118 packing pattern.

example, for structures belonging to the packing pattern 4*L*2*M*1*S*, the most common geometry for the asymmetric unit is 112*L*. The same is true for structures belonging to 4*L*1*M*2*S* even though the alternative 112*S* asymmetric unit has the lower surface area. This seems to suggest that a '4 direction' is a natural extension of a strong energetic interaction between the smallest faces of the molecules. However, in only 18 of the 50 4*L*2*M*1*S* structures for which 112*L* describes the geometry of the asymmetric unit is the interaction between the molecules of the asymmetric unit the strongest energetic interaction of the structure. This result serves as a reminder that although we like to think of crystal structures in terms of 'structure-determining' hydrogen bonds or low surface-area intermolecular interactions, crystal structures represent a compromise between many competing interactions.

4. Summary

Structures where there is more than one molecule in the asymmetric unit represent approximately 9% of the Cambridge Structural Database. A data set of $Z' = 2$ structures where two molecules in the asymmetric unit were chemically identical were extracted from the CSD. For the majority of these structures, the independent molecules were aligned in the same way with respect to the unit cell axes. Thus, 75% of these structures can be described by a single packing pattern of the Box Model. The histograms of pattern coefficients and the scatterplots of the positions of the molecular centres for these $Z' = 2$ structures are very similar to results obtained from similar analyses performed on $Z' \leq 1$ structures. Therefore, despite the lack of crystallographic symmetry relating the molecules of the asymmetric unit, the positions of the molecules in the cell and the spatial arrangement of the molecules are almost indistinguishable from $Z' \leq 1$ structures. The consistency of results returned from the Box Model for $Z' \leq 1$ and $Z' = 2$ structures serves as further proof that the Box Model is based on, and reflects, some very basic principles of crystal packing.

The preference for low surface-area packing patterns had been noted in previous studies performed for $Z' \leq 1$ structures. This preference appears somewhat amplified in the 'aligned' dataset of $Z' = 2$ structures, with a higher proportion being found in low surface-area packing patterns than found for $Z' \leq 1$ structures. Also, it has been found that the strongest energetic interaction between the independent molecules relates the largest faces of the molecules in 50.5% of cases and the smallest faces of the molecules in only 19.3% of cases. Thus, lower surface-area asymmetric units and low surface-area unit cells are preferred. One conclusion which could be drawn form these observations is that when the 'restraints' imposed by crystallographic symmetry are relaxed, the driving

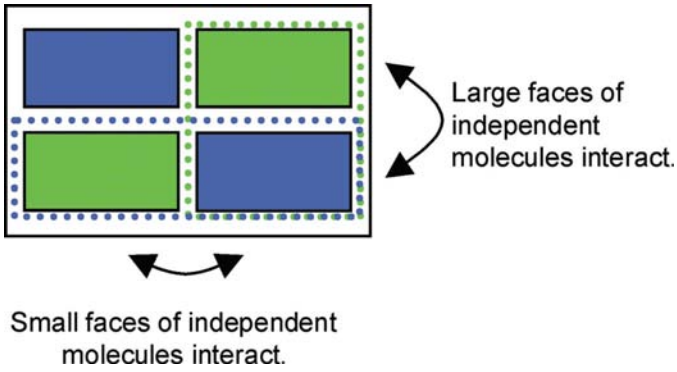

Figure 13
A schematic diagram illustrating a $Z = 4$, $Z' = 2$ unit cell. The symmetry-equivalent molecules are the same colour and the independent molecules a different colour. The two choices for the asymmetric unit are indicated by a dark-coloured dotted rectangle or a light-coloured dotted rectangle. In the light-coloured asymmetric unit the large faces of the molecules interact and this arrangement has the lower surface area. In the dark-coloured unit cell, the small faces of the molecules interact and the surface area of the arrangement is larger.

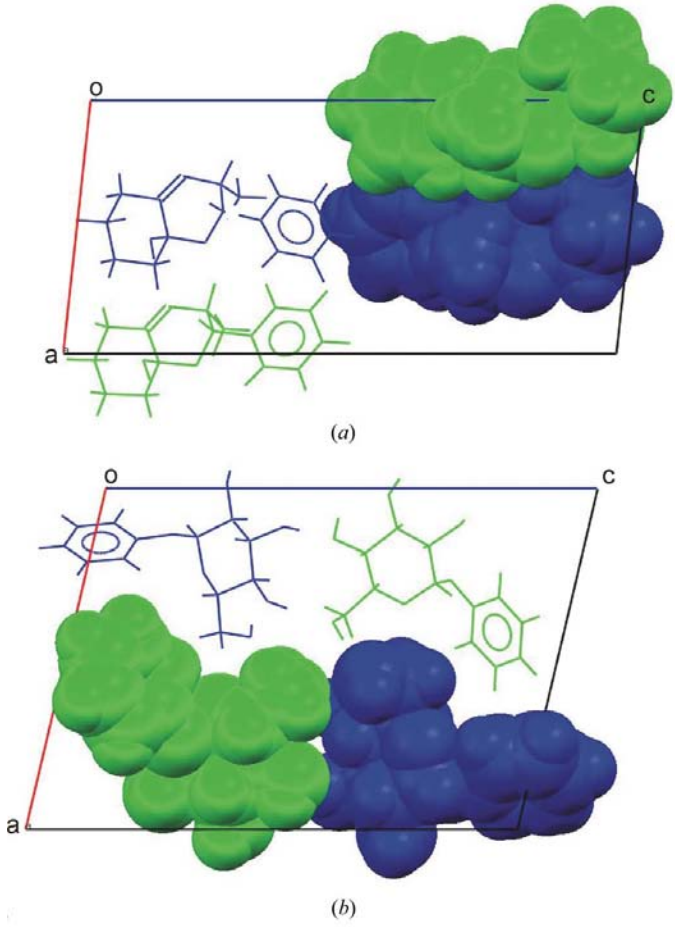

Figure 14
Two crystal structures belonging to the packing pattern 221*S*. The *S* (short) molecular dimension is aligned with the unit-cell *b* axis, the *L* (long) molecular dimension is aligned with the unit-cell *c* axis in both structures. The energetically significant asymmetric unit is indicated by the molecules presented in 'spacefill' style. Left: ZORZIP (Ortiz *et al.*, 1995), where the strongest energetic interaction between the independent molecules is along the *a* axis and relates the *LS* (medium) faces of the molecules. The asymmetric unit, using the packing pattern nomenclature, is 112*M*. Right: BOPXEJ (Jones *et al.*, 1982), where the strongest energetic interaction between the independent molecules is aligned with the *c* axis and relates the *MS* (smallest) faces of the molecules. The asymmetric unit is described by the 112*L* packing pattern.

force to minimize surface area for a given volume exerts more influence.

Finally, the analysis of $Z' = 2$ structures within the context of the model has led to an appreciation of the subtle geometric factors which determine whether a structure is, for example, built from *A*–*B*–*B*–*A* columns or an *A*–*B*–*A*–*B* columns, where *A* and *B* are the independent molecules of the asymmetric unit. It is unlikely that this analysis could have been performed without the framework provided by the Box Model. It is hoped this classification of structures is one of many useful observations that are brought to light through the analysis of $Z' = 2$ structures within the context of the Box Model.

Many thanks to W. D. S Motherwell for a critical reading of this manuscript.

References

Aakeröy, C. B., Beatty, A. M., Helfrich, B. A. & Nieuwenhuyzen, M. (2003). *Cryst. Growth. Des.* **3**, 159–165.

Adams, R. D., Captain, B., Kwan, O.-S. & Miao, S. (2003). *Inorg. Chem.* **42**, 3356–3365.

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

ApSimon, J. W., Yamasaki, K., Fruchier, A., Chau, A. S. & Huber, C. P. (1982). *Can. J. Chem.* **60**, 501–506.

Beurskens, P. T., Beurskens, G., Apreda, M. C., Foces-Foces, C., Cano, F. H. & Garcia-Blanco, S. (1984). *Acta Cryst.* C**40**, 1718–1721.

Birnbaum, G. I. (1977). *Can. J. Chem.* **55**, 1619–1623.

Britton, D. (2000). *Acta Cryst.* B**56**, 828–832.

Bruno, I. J., Cole, J. C., Edgington, P. R., Kessler, M., Macrae, C. M., McCabe, P., Pearson, P. & Taylor, R. (2002). *Acta Cryst.* B**58**, 389–397.

Gautham, N. (1992). *Acta Cryst.* B**48**, 337–338.

Gavezzotti, A. (1994). *Acc. Chem. Res.* **27**, 309–331.

Gel'mbol'dt, V. O., Filinchuk, Ya. E. & Koroeva, L. V. (2003). *Russ. J. Coord. Chem.* **29**, 818.

Jones, P. G., Sheldrick, G. M., Kirby, A. J. & Glenn, R. (1982). *Z. Kristallogr. Kristallgeom. Kristallphys. Kristallchem.* **160**, 259–267.

Karthe, P., Sadasivan, C. & Gautham, N. (1993). *Acta Cryst.* B**49**, 1069–1071.

Kitaigorodskii, A. I. (1961). *Organic Chemical Crystallography*. New York: Consultants Bureau.

Lewandos, G. S., Doherty, N. M., Knox, S. A. R., MacPherson, K. A. & Orpen, A. G. (1988). *Polyhedron*, **7**, 837–845.

Marsh, R. E., Schomaker, V. & Herbstein, F. H. (1998). *Acta Cryst.* B**54**, 921–924.

Motherwell, W. D. S., Shields, G. P. & Allen, F. H. (1999). *Acta Cryst.* B**55**, 1044–1056.

Ortiz, A., Farfan, N., Santillan, R., Rosales, M. de J., Garcia-Baez, E., Daran, J. C. & Halut, S. (1995). *Tetrahedron: Asymm.* **6**, 2715.

Pavé, G., Léger, J.-M., Jarry, C., Viaud-Massuard, M.-C. & Guillaumet, G. (2003). *Tetrahedron Lett.* **44**, 4219–4222.

Pidcock, E. & Motherwell, W. D. S. (2003). *Chem. Commun.* pp. 3028–3029.

Pidcock, E. & Motherwell, W. D. S. (2004*a*). *Cryst. Growth Des.* **5**, 611–620.

Pidcock, E. & Motherwell, W. D. S. (2004*b*). *Acta Cryst.* B**60**, 539–546.

Pidcock, E. & Motherwell, W. D. S. (2004*c*). *Acta Cryst.* B**60**, 725–733.

Pidcock, E. & Motherwell, W. D. S. (2005). *Cryst. Growth Des.* **5**, 2322–2330.

Rodier, N., Baassou, S., Mehri, H. & Plat, M. (1982). *Acta Cryst.* B**38**, 863–867.

Sona, V. & Gautham, N. (1992). *Acta Cryst.* B**48**, 111–113.

Steed, J. (2003). *CrystEngComm*, **5**, 169–179.

Steiner, T. (2000). *Acta Cryst.* B**56**, 673–676.

Zorky, P. M. (1996). *J. Mol. Struct.* **374**, 9–28.