

Precursor effects of the orthorhombic to monoclinic phase transition in benzocaine form (II) revealed by X-ray diffuse scattering

E. J. Chan and T. R. Welberry*

Research School of Chemistry, Australian
National University, Canberra, ACT 0200,
AustraliaCorrespondence e-mail:
welberry@rsc.anu.edu.au

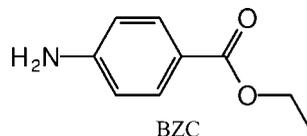
Received 4 December 2009

Accepted 12 January 2010

We described the development of a Monte Carlo computer model for the room-temperature form (II) polymorph of benzocaine that incorporates, on a local scale, structural features derived from the low-temperature form (III) polymorph. The introduction of this extra information convincingly reproduces those observed diffraction features that an earlier harmonic model was unable to achieve. In both form (I) and form (II) the hydrogen-bonded chains of molecules that extend along the respective *a* axes tend to slide backward and forward along their lengths. While in form (I) the motion is well modelled by a harmonic potential in form (II) there is a degree of anharmonicity that means that some intermolecular contact vectors, which are identical in the average structure, are distributed bimodally with either longer or shorter vectors being preferred to the average. Moreover there is a tendency for these deviations from average to be correlated to give short-range ordered domains that are the precursors of the two twinned variants of the long-range ordered low-temperature form (III) structure.

1. Introduction

The work presented in this paper is part of a long-term study in which diffuse X-ray scattering methods are being used to probe in detail the local structure and dynamics of molecular materials that exhibit polymorphism in order to gain an understanding of how and why polymorphism occurs and how it might be predicted. Polymorphism is of particular importance in the pharmaceutical industry, but is also important in the manufacture of other materials such as pigments, dyes and explosives (Bernstein, 2002). Much of the structural information concerning polymorphic materials has come from conventional crystallography (Day *et al.*, 2009), but diffuse scattering can provide valuable additional information (Beasley *et al.*, 2008).



In a previous paper we reported on the analysis of the diffuse scattering in the two room-temperature polymorphs of the drug benzocaine, C₉H₁₁O₂N (BZC; Chan, Welberry *et al.*, 2009). The analysis involved the fitting of Monte Carlo (MC) computer simulation models of the structures to the observed diffuse scattering patterns. The MC models used in that work consisted of sets of harmonic springs to represent *effective* intermolecular interactions and torsional springs to represent restoring forces on the intramolecular rotations around single bonds. The models assumed that the diffuse scattering was

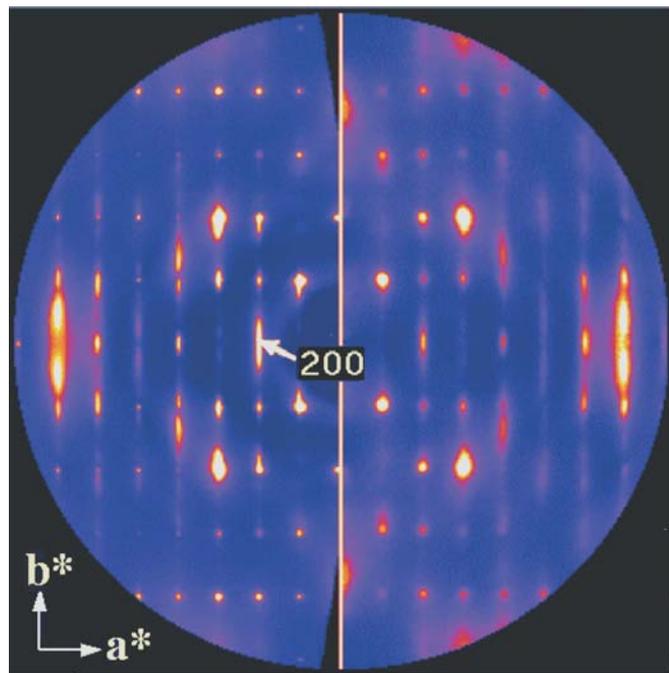
Table 1

Cell data for the three polymorphs, (I), (II) and (III), of benzocaine, $C_9H_{11}O_2N$ (BZC).

Polymorph	BZC (I)	BZC (II)	BZC (III)
Space group	$P2_1/c$	$P2_12_12_1$	$P2_1$
a (Å)	8.2570 (2)	8.2424 (9)	8.1883 (4)
b (Å)	5.5009 (4)	5.3111 (3)	10.6374 (5)
c (Å)	19.9560 (2)	20.9044 (2)	20.4761 (1)
α (°)	90.0	90.0	90.0
β (°)	91.699 (4)	90.0	90.0
γ (°)	90.0	90.0	99.370 (2)
V (Å ³)	906.02 (14)	915.12 (9)	1760.05 (15)
Z	4	4	8
V/Z	226.50	228.78	220.01
Temperature (K)	300	300	150

purely thermal in origin, *i.e.* thermal diffuse scattering (TDS). The analyses revealed the presence of strongly correlated molecular motions which were very similar in the two different forms. One notable type of motion observed was that chains of molecules linked by end-on N–H...O hydrogen bonds tend to slide back and forth along their length, parallel to the a axis.

However, despite the fact that the two models were very similar and the data of comparable quality, the fit that was obtained for form (I) was significantly better than for form (II), suggesting that the model for form (II) was lacking some essential ingredient. The inability of the harmonic model to completely fit the data for form (II) was most obvious in the $hk0$ section where the strong elongated diffuse peak in the vicinity of the 200 reflection, in particular, was very poorly modelled (see Fig. 1). As a result of this, low-temperature

**Figure 1**

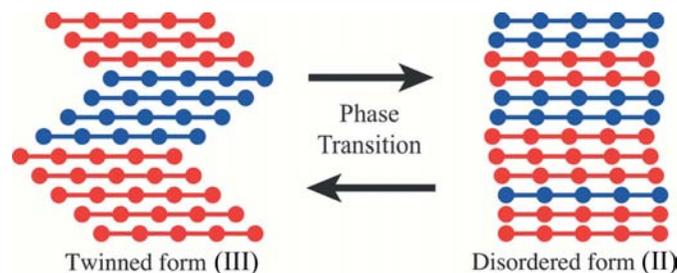
The diffraction pattern of the $hk0$ section of form (II) calculated from the harmonic spring model (right) compared with the observed data (left). Note the poor agreement for the elongated diffuse peak around the 200 position.

experiments were carried out and it was discovered that form (II) undergoes a reversible phase transition to a new form (III). The average structure of form (III) has been reported in Chan, Rae *et al.* (2009). The cell data for the three polymorphs are reproduced here for convenient reference (see Table 1). The phase transition occurs only just below room temperature (at ~ 270 K) so that it seems feasible that the inability of the harmonic spring model for form (II) to fully fit the observed data was due to the presence of precursor effects of the phase transition in the room-temperature structure. It was conjectured that these could be manifest as anharmonic motions or even as static disorder.

Although in principle it should be possible to build a model for a molecular crystal that so closely resembles the real system it would be able to simulate its phase transition behaviour, such a detailed and accurate description of a system is still well beyond current capabilities. Therefore, in order to obtain insight into the nature of the transition in form (II) of benzocaine we attempt in this paper to build a model based on the twinned low-temperature structure and track how this might progress towards a room-temperature structure in which attributes of the low-temperature form are retained on a local scale. The aim is to test whether such a structure can account for the features of the observed X-ray patterns that were not modelled well by the harmonic spring model.

2. Development of a disordered model for form (II)

A key result from the analysis using purely harmonic springs was that hydrogen-bonded chains of molecules tended to slide back and forth along the a axis [in both polymorphs (I) and (II)]. It was conjectured that in form (II) some kind of anharmonic effect was present in this sliding motion. That is, the average position is not the energy minimum and chains are more likely to be displaced one way or the other from the average position. As a result of this we envisage that the phase transition in form (II) occurs as a result of the chains locking into one or other of the two kinds of displacement. In form (III), below the transition, the structure is comprised of domains of two different twin components in which the chains have slipped progressively further and further in one or other of the two alternative directions. Above the phase transition

**Figure 2**

Schematic representation of the phase transition from a disordered high-temperature structure to an ordered (but twinned) low-temperature structure.

the chains are envisaged to be displaced in a more random fashion about the average molecular position. This is shown schematically in Fig. 2. However, at room temperature, some 20–25° above the phase transition, the structure shows precursor effects of what is about to happen and the displacements will not be completely random but show some local ordering.

In this section we describe our attempts to build a model to describe these precursor effects and show how they are able to account for the anomalous diffraction features (see Fig. 1) that are not accounted for by a purely harmonic model.

2.1. Outline

As a general outline we follow the steps:

(i) Fractional atomic coordinates are taken from one (monoclinic) twin of the form (III) structure and these are

plotted (Fig. 3c) in an orthorhombic cell [corresponding to a doubled form (II) cell].

(ii) This is repeated using coordinates corresponding to the second twin component of form (II) (Fig. 3d).

(iii) Steps (i) and (ii), as well as deforming the form (III) unit cell, result in slightly deformed molecules. These deformed molecules are substituted as closely as possible by undeformed molecules specified by z -matrices and quaternions.

(iv) Each molecular site i in the crystal is assigned a random variable, $\sigma_i = \pm 1$. $\sigma_i = 1$ is taken to mean that the site is occupied by a molecule derived from twin component 1 and $\sigma_i = -1$ is taken to mean that the site is occupied by a molecule derived from twin component 2. This structure is thus statically disordered.

(v) Monte Carlo simulation is used to introduce correlations between neighbouring σ_i variables so that domains are formed in which locally the structure is

comprised of all twin component 1 coordinates or all twin component 2 coordinates. This structure is still just statically disordered.

(vi) The molecules are then connected by harmonic springs to represent the effective intermolecular interactions analogous to those used in the original harmonic spring model analyses (see Chan, Welberry *et al.*, 2009). With the σ_i variables fixed, a second MC simulation is carried out in order to induce thermal displacements and hence thermal diffuse scattering (TDS).

2.2. Using coordinates from the form (III) structure

Fig. 3(a) shows a plot of a single molecular layer normal to c in form (II) of benzocaine. The superimposed blue rectangles have been drawn to indicate the size of two unit cells of form (II) for comparison with the doubled cell of form (III). Fig. 3(b) shows the same molecular layer in form (III) and here the blue parallelogram shows the size of the actual unit cell. In both of these figures the vertical columns of molecules correspond to the hydrogen-bonded chains (running along a)

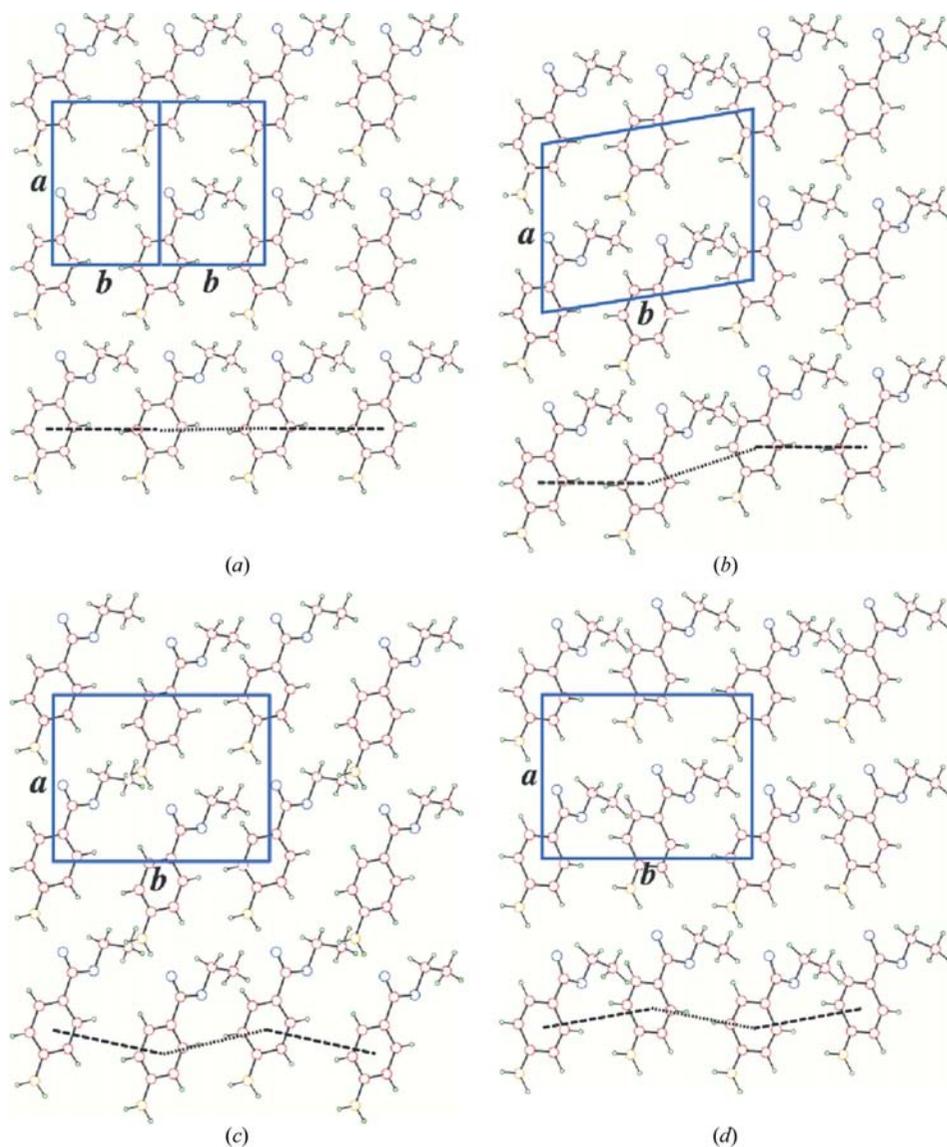


Figure 3
A single layer of molecules of benzocaine viewed down the c axis in the following structures: (a) form (II); (b) form (III) (twin component 1); (c) the same fractional coordinates as in (b) but plotted in an orthorhombic cell; (d) an analogous plot to (c) but using coordinates from twin component 2.

mentioned earlier. It is seen that in form (II) (Fig. 3*a*) these chains are all related by the *b*-axis translational symmetry, so that the dashed and dotted vectors drawn on the bottom row of molecules are equivalent. In form (III) (Fig. 3*b*) the relationship between the first two chains is very similar to that in form (II), but then the next two chains are offset by a translation along *a*. This results in a doubling of the *b* cell dimension as well as a departure from 90° for the γ cell angle. Now the dashed and dotted vectors are no longer equivalent.

Fig. 3(*c*) shows a plot using the same fractional coordinates as for Fig. 3(*b*) but these have now been plotted in an orthorhombic cell identical to the doubled form (II) cell (*cf.* Fig. 3*a*). Since the fractional coordinates of this structure are the same as for the structure in Fig. 3(*b*) the Bragg intensities will be the same (to a very good approximation) as for form (III), but the peaks will fall on the orthorhombic [form (II)] reciprocal lattice. Comparing this modified form (III) structure (Fig. 3*c*) with the original average form (II) structure (Fig. 3*a*) the two look very similar except for the relative displacements along *a* of neighbouring chains.

Fig. 3(*d*) shows a similar plot which has been derived from a form (III) crystal having the opposite twin orientation, *i.e.* with the successive molecular chains displaced down rather than up. We can therefore envisage a disordered model for form (II) that on a local scale resembles one or other of these two modified forms. It should be noted that each structure has four molecular layers per cell stacked along *c*. Those shown in Figs. 3(*a*) and (*b*) are only one of the four. However, the same kind of treatment can be applied to each of the other three layers. In each case a given molecular site will contain a molecule in either of two orientations/positions: **A**, corresponding to twin component 1, and **B**, corresponding to twin component 2. (It should be noted that, taken on their own, the single layers in Fig. 3*c* and Fig. 3*d* appear to be different structures. However, when all four layers in the unit cell are considered together two identical but twin-related structures occur when either all molecules are in the **A** position or all in the **B** position.)

The average structure of this disordered model will thus show a molecular site with the **A** and **B** molecules superposed.

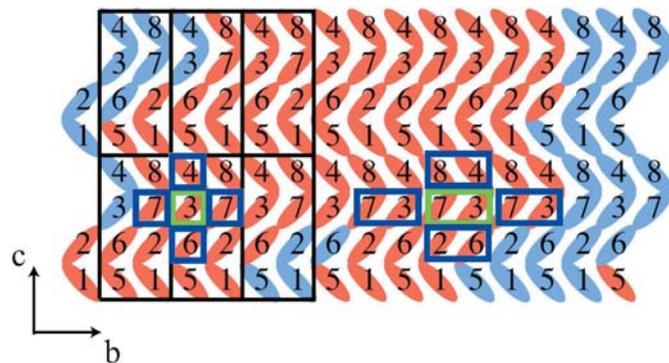


Figure 4
Schematic representation of the model structure viewed down the *a* axis. See text for further details.

In all, because we are using a doubled form (II) cell there are eight molecular sites per cell. Viewed down the *c* axis these are all superposed, but may be clearly seen when viewed down the *a* axis, as shown in Fig. 4. In this figure each elliptic motif represents one of the molecular chains viewed end-on and the colour represents the type of molecule present (**A** or **B**). Black rectangles indicate the size of the doubled orthorhombic unit cell. (Note these rectangles have been shifted compared with the actual unit cells seen in Fig. 5 in order to completely enclose the eight molecular motifs.) The numbers label the eight molecular sites within each cell. It should be noted that in the average form (II) structure sites 1, 2, 3 and 4 are equivalent to sites 5, 6, 7 and 8.

2.3. Ordering the molecular site variables

It is envisaged that at higher temperatures there is no correlation between the occupancy (**A** or **B**) of different sites but, on cooling to room temperature and below, correlations increase and the domains of the two twin orientations of form (III) begin to form. In this subsection we describe the MC models used to induce this ordering. Two different models have been used. Model 1 involves interactions between nearest-neighbour chains only. In Fig. 4 the green square identifies a typical target site for this model and the blue squares identify the neighbouring sites with which the target site interacts. For Model 2 the simulation is carried out using

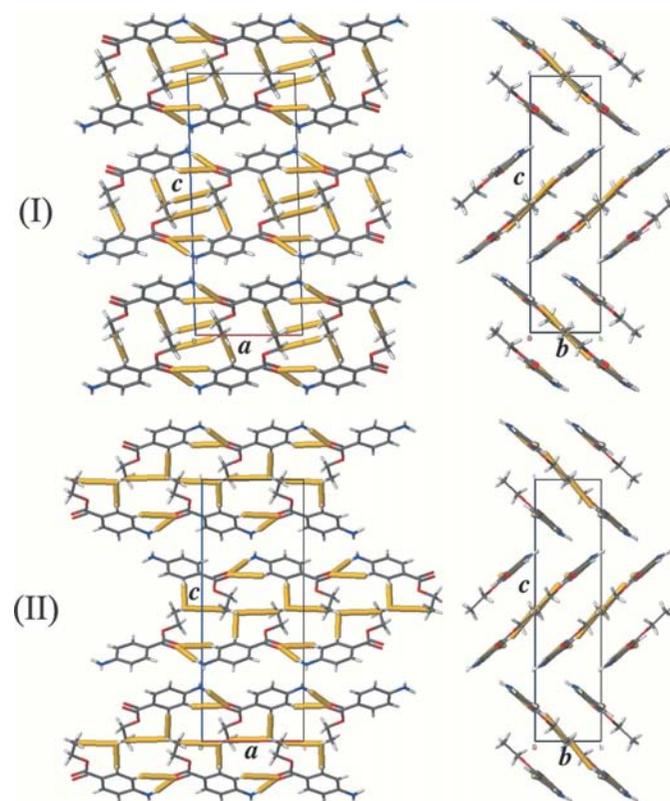


Figure 5
Cell plots viewed down *b* (left) and down *a* (right) for forms (I) and (II) of BZC. Yellow contacts between molecules represent group (II) of the interaction vectors (see Table 2). See text for details.

Table 2

Summary of the simplified groups of interaction vectors used.

Values here are for interactions with a target molecule with symmetry code x, y, z . G = group type; V = vector type; O = origin atom; C = cell translation of destination molecule with respect to symmetry-operated origin molecule; S = symmetry relationship of destination molecule to origin molecule; D = destination atom; L = length of vector; F = spring constant. $x24$ is a dummy atom, placed at the centroid of the phenyl ring, that was used in the z -matrix description of the molecule.

Form (I)								Form (II)							
G	V	O	C	S	D	L	F	G	V	O	C	S	D	L	F
1:	1	H21	0 0 0	i	$x24$	2.824	2.0	1:	1	H21	0 0 0	i	$x24$	3.879	2.0
2:	2	H20	1 1 0	i	C10	3.191	4.0	2:	2	H20		i	C11	3.313	4.0
	3	C11	0 1 0	i	H13	2.889			3	C20	0 1 0	i	C13	2.411	
	4	O8	1 0 0	ii	H14	2.746			4	O8	1 0 0	ii	H14	2.759	
	5	O8	1 0 0	ii	H22	2.105			5	O8	1 0 0	ii	H22	2.051	
3:	6	C6	0 0 0	iii	H15	3.078	1.0	3:	6	C6	0 0 0	iii	H15	3.061	1.0
	7	C2	0 1 0	ii	H15	3.195			7	C2	0 1 0	ii	H15	3.195	
	8	H13	0 1 0	ii	C5	3.216			8	H13	0 1 0	ii	C5	3.218	
4:	9	N12	1 0 0	iii	H23	2.426	8.0	4:	5	N12	1 0 0	iii	H23	2.332	8.0

Symmetry codes for form (I): (i) $1 - x, 1 - y, 1 - z$; (ii) x, y, z ; (iii) $1 - x, \frac{1}{2} + y, \frac{1}{2} - z$; for form (II): (i) $\frac{1}{2} + x, 1 - y, 1 - z$; (ii) x, y, z ; (iii) $1 - x, \frac{1}{2} + y, \frac{1}{2} - z$.

pairs of sites. The blue and green rectangles in Fig. 4 identify the corresponding target and neighbouring sites used for this model.

2.3.1. Model 1. The Monte Carlo energy for this first model was in the form of a simple Ising model

$$E = \sum_t \sigma_t \sum_n J_n \sigma_n. \quad (1)$$

Here t corresponds to a target site and n to a neighbouring site. The first summation is over all sites in the crystal and the second over all sites neighbouring the target site. Three different values of J_n were used, J_a, J_b and J_c , corresponding to neighbours in the a, b and c directions. Note that these neighbours do not correspond to molecules related by cell translations, but refer to the rows and columns of elliptic motifs shown in Fig. 4. A suitable choice of these three parameters enables either positive or negative correlation to be induced in each direction.

2.3.2. Model 2. For model 2, four-state random variables ξ_i were used to represent pairs of molecular sites, as shown in Fig. 4. The aim in this case was to allow correlations between the second nearest σ_i variables to be induced. The four states of each variable are used to represent the four possible pair combinations, **AA**, **AB**, **BA** or **BB**. In this case the MC energy that was used was given by

$$E = \sum_t \sum_n \mathcal{F}_n(\xi_t, \xi_n). \quad (2)$$

\mathcal{F}_n is a function whose value depends on the values of ξ_i on both the target and the neighbouring site. This is best expressed in matrix form. For neighbours in the b direction

$$\mathcal{F}_b(\xi_t, \xi_n) = \begin{array}{c|cccc} & \mathbf{AA} & \mathbf{AB} & \mathbf{BA} & \mathbf{BB} \\ \mathbf{AA} & J_1 & J_1 & J_2 & J_2 \\ \mathbf{AB} & J_2 & J_2 & J_1 & J_1 \\ \mathbf{BA} & J_1 & J_1 & J_2 & J_2 \\ \mathbf{BB} & J_2 & J_2 & J_1 & J_1 \end{array}. \quad (3)$$

Similarly, for neighbours in the c -direction

$$\mathcal{F}_c(\xi_t, \xi_n) = \begin{array}{c|cccc} & \mathbf{AA} & \mathbf{AB} & \mathbf{BA} & \mathbf{BB} \\ \mathbf{AA} & J_3 & J_4 & J_4 & J_5 \\ \mathbf{AB} & J_4 & J_3 & J_5 & J_4 \\ \mathbf{BA} & J_4 & J_5 & J_3 & J_4 \\ \mathbf{BB} & J_5 & J_4 & J_4 & J_3 \end{array}. \quad (4)$$

For interactions along the a axis, *i.e.* along the molecular chains, no special treatment was required and an energy term identical to that in (1) was used.

The symmetries of the matrices given in (3) and (4) were required in order to ensure that the average occupancy of each molecular site was 50% **A** and 50% **B**, and that the two halves of the doubled cell were the same on average.

2.4. Modelling thermal displacements

2.4.1. A simplified spring model. A simplified version of the original harmonic model is first described. This is used as a starting or reference point for subsequent models in which the **A** and **B** molecular positions have been inserted instead of the average positions.

In our earlier paper on BZC (Chan, Welberry *et al.*, 2009) the harmonic spring models used to describe both the form (I) and form (II) structures comprised sets of 15 linear springs for the intermolecular interactions and four additional torsional springs for the internal degrees of flexibility. For the present study we have used a simplified model in order to facilitate the modelling process outlined in §2.1. The main feature of this simplified model was that the molecules were considered to be

rigid. This in turn allowed a reduced number of intermolecular springs to be used.

Instead of using 15 spring types it was found through successive trials that only nine spring vector types were actually necessary once the molecules were assumed to be rigid. The model was then further simplified by separating these nine spring vector types into four groups with only a single independent spring constant parameter used for all the members of a given group. Details are given in Table 2. An example of one of the spring groups is shown in Fig. 5. All of the springs shown (as yellow rods) represent intermolecular interactions for the ribbon substructure and the same spring constant parameter is used for all of them [for both form (I) and form (II)]. The strengths of the four spring constants were adjusted by trial and error until satisfactory agreement with the

observed data was obtained. Although obtained in this semi-quantitative manner any differences from the original model fits appeared quite small and it was considered that the computational benefits of having a much reduced number of parameters outweighed any loss of accuracy in the fitting. In addition, with the use of the same spring constants for both models, considerable further confidence was obtained that the discrepancies observed in the $hk0$ section (see Fig. 1) for form (II) were real anharmonic features of the structure and not due to the choice of springs in the modelling.

Fig. 6 shows the observed and calculated patterns of the $hk0$ section for both forms (I) and (II) at room temperature using this simplified harmonic model. Of particular note is that, while for form (I) the feature near 200 is well modelled (see inset), for form (II) the feature is not modelled well at all. As in the earlier paper (Chan, Welberry *et al.*, 2009) this diffuse peak is observed to be much stronger and more elongated than the purely harmonic model is able to reproduce.

2.4.2. Transforming the structure. In order to show how the inclusion of attributes of the form (III) structure in a model of the room-temperature form (II) structure can explain the anomalous diffraction effects that have been observed, the following strategy was adopted. First a disordered model was created as described in §2.2. Then short-range correlations (or short-range order, SRO) were introduced in the distribution of the two types of molecular positions, as described in §2.3. Starting from this initial structure the molecules were then connected by the harmonic springs derived from the simplified harmonic model described in §2.4. Finally MC simulation was carried out on this system in order to induce thermal displacements while maintaining the SRO present in the starting model. If this simulation is continued indefinitely the model eventually reaches the final harmonic model. However, at intermediate stages of the simulation a situation arises where the model still has significant character of the low-temperature twinned structure while also having attributes of the thermal disorder of the harmonic model.

The evolution of the structure during this process can be monitored by plotting the deviations (along **a**) of the molecules from their average sites. At the outset the distribution is purely bimodal with shifts either approximately $+0.5$ or -0.5 Å, corresponding to the plotted structures of Figs. 3(c) and (d). As the MC proceeds the two sides of the bimodal distribution first acquire a spread about their local means but then gradually the two halves of the distribution merge and become a single broad distribution, eventually becoming Gaussian. Fig. 7 shows the distribution measured from the simulation coordinates at two different stages of this process. Fig. 7(a) was obtained after 10 cycles of MC iteration, while Fig. 7(b) was obtained after 100 cycles. Fig. 7(b) shows that, even after only 100 MC cycles, the molecular displacements are considerably reduced from the ± 0.5 Å; implied by the initial twin structure coordinates. After 1000 cycles (not shown) the distribution has a similar width but is now a single Gaussian. The fact that the displacements started out from the two quite separate populations and these have now merged to a single distribution does not mean that all memory of the

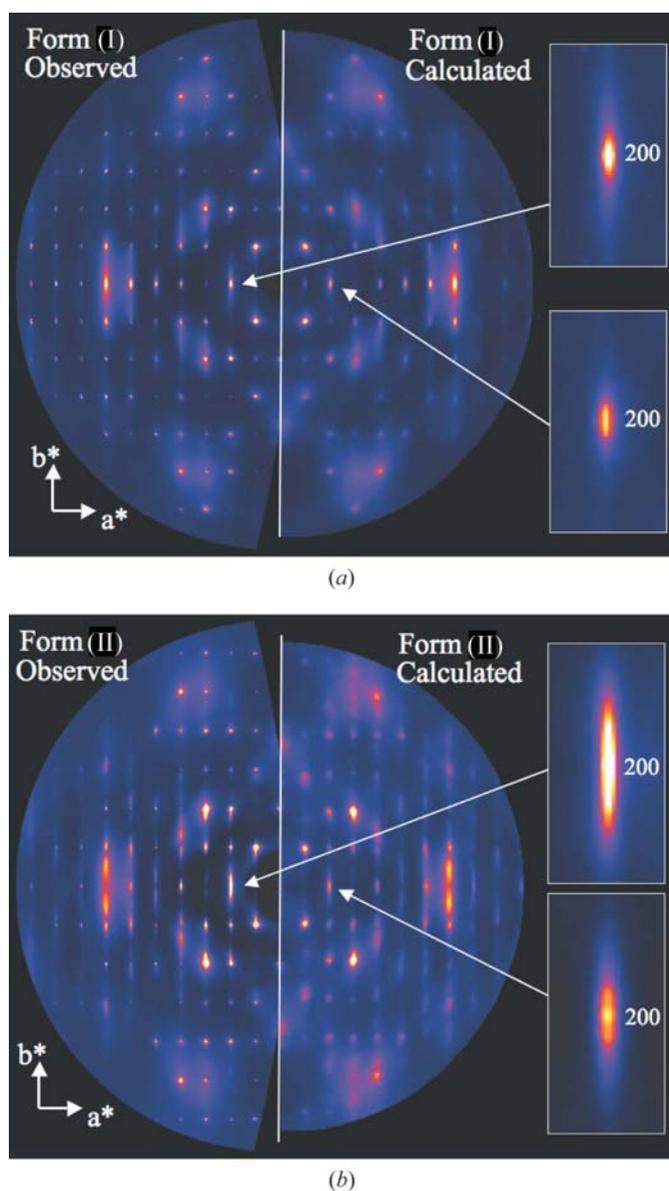


Figure 6
Observed and calculated diffraction patterns for the $hk0$ sections of (a) form (I) and (b) form (II). The simplified harmonic spring model with the same spring constants was used for both calculations.

initial SRO domain structure has been lost. In applying the thermal motion to the SRO domain model, the distribution of molecular displacements has merged to a single Gaussian about the average form (II) position, but the correlation between the displacements of neighbouring molecules has, to a significant extent, been retained, *i.e.* memory of the initial SRO twin structure has not been lost. Clearly, if MC iteration is continued then eventually all memory of the twin structure will be lost. This evolution might be considered, in a rather simplistic way, to parallel the changes that take place as the crystal transforms on heating, the number of MC cycles representing a sort of reaction coordinate.

It was found after extensive tests that stopping the evolution after 1000 cycles gave optimum results for describing the room-temperature observations and so this value was adopted for all the investigations described below. In order to investigate the possibility of obtaining a resultant distribution that gives diffraction patterns in agreement with the observed patterns this process of evolution from a purely bimodal starting distribution was repeated for different values of the ordering energies given by (1) (model 1) or (2) (model 2).

2.5. Simulation details

All of the simulations described in this paper were carried out using a model crystal comprising $48 \times 48 \times 48$ unit cells. The first-stage MC simulations, in which the occupancies of molecular sites were ordered using energies given by (1) (model 1) or (2) (model 2), were carried out for 5000 MC cycles where one cycle corresponds to each molecular site being visited once on average. The method was adopted in which two sites were selected at random and the respective values of the σ_i or ξ_i variables were interchanged. The move was accepted or rejected using the normal Metropolis algorithm (Metropolis *et al.*, 1953) with Boltzmann partition $P = \exp(-\Delta E/k_B T)$, where ΔE is the difference between the new and old system energy. The simulation temperature was maintained at a constant value of $T = 1/k_B$, where k_B is Boltzmann's constant. The J_i values used in the simulation (see Table 2) are measured relative to $k_B T$.

In the second stage of MC simulation, in which thermal displacements were introduced, essentially the same approach as described in Chan, Welberry *et al.* (2009) was used, although simplified because of the adoption of rigid molecules. For each MC step a molecule was chosen at random and the variables specifying its orientation (four quaternion components) and position (x, y, z coordinates) were subjected to random increments. The energy, E_{therm} , was calculated before and after the shift using (5) and the move was accepted or rejected, again using the normal Metropolis algorithm (Metropolis *et al.* 1953) with Boltzmann partition $P = \exp(-\Delta E_{\text{therm}}/k_B T)$. The simulation temperature was again maintained at a constant value of $T = 1/k_B$. The force constants (see Table 2) are measured relative to $k_B T$.

$$E_{\text{therm}} = \sum_{\text{all linear springs}} K_i(d - d_0)^2. \quad (5)$$

With this simplified spring model 1000 MC cycles were found sufficient to achieve equilibrium and this number of cycles was used for the patterns shown in Fig. 6. As explained in §2.4.2, a lesser number of cycles was used to explore the transition.

All calculated diffraction patterns were made using the program *DIFFUSE* (Butler & Welberry, 1992). It should be noted that because the average structure is subtracted within *DIFFUSE*, all calculated patterns are of the diffuse scattering alone and contain no Bragg peaks.

3. Results

3.1. Model 1

Fig. 8 shows a series of three snapshots of the domain structure viewed down **a**, produced using Model 1 [see (1)]. The aim was to produce domains of the **A** and **B** twin components that were relatively large in the *a* and *c* directions and investigate the effect of different amounts of ordering along *b*. For the three examples shown $J_a = J_c = -1.0$ and $J_b = -0.0, -0.05$ and -0.10 for Figs. 8(a)–(c). Note that the images of the domain structure correspond to a view down **a**, *i.e.* a projection along the length of the molecular chain, but the diffuse 200 peak shown alongside each image is viewed down **c**.

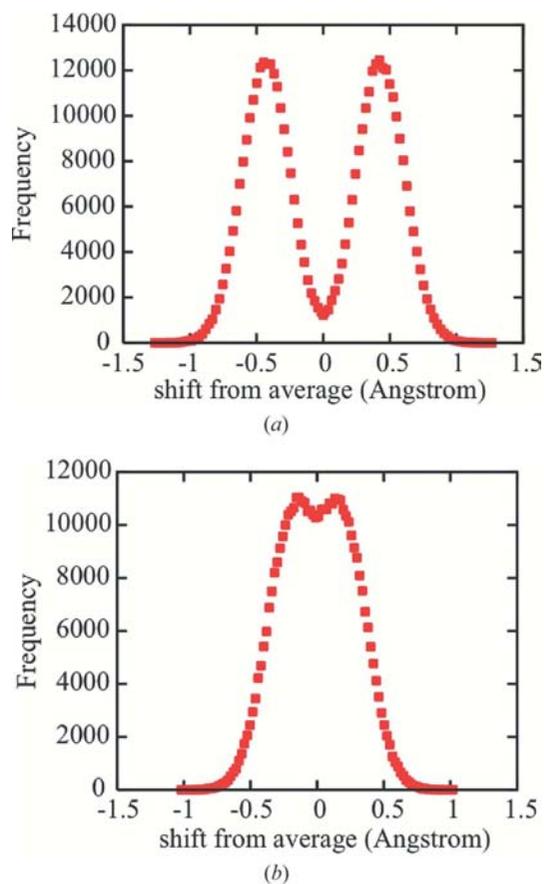


Figure 7
The distribution of the displacements (along **a**) of individual molecules away from their average site positions. (a) After 10 cycles of MC simulation; (b) after 100 cycles.

Table 3

Interaction energies used in the model 2 simulations (BZC).

	J_1	J_2	J_3	J_4	J_5	P_{AA}	C_1
Fig. 9(a)	0.0	0.00	-1.0	0.0	1.0	0.25	0.00
Fig. 9(b)	0.0	0.20	-1.0	0.0	1.0	0.25	0.50
Fig. 9(c)	0.0	0.02	-1.0	0.0	1.0	0.25	0.36
Fig. 9(d)	0.0	0.02	-1.0	0.0	1.0	0.26	0.38

The nearest-neighbour correlation coefficient is defined as

$$C_1 = \frac{(P_{AA} - m_A^2)}{m_A m_B}. \quad (6)$$

Here P_{AA} is the probability that neighbouring sites are both occupied by a molecule in orientation **A** and m_A is the overall fraction of sites occupied by **A**. The nearest-neighbour correlation along *b* actually achieved by this ordering process was 0.0 for Fig. 8(a), 0.34 for Fig. 8(b) and 0.55 for Fig. 8(c).

Despite the large difference in the nearest-neighbour correlation, which is reflected in the domain structure plots, the shape of the diffuse 200 peak does not show the same marked variation. Compared with the purely harmonic model the intensity of the peak is larger but there is no tendency towards the markedly elongated (along **b**) character of the observed pattern. It should be noted that for a statically disordered crystal the intensity of diffuse scattering is proportional to the difference of the molecular scattering factors, $|F_A - F_B|^2$, whereas for TDS the intensity close to Bragg peaks tends to be proportional (like the Bragg peaks themselves) to the average scattering factor, $|m_A F_A + m_B F_B|^2$ (Welberry, 1985, 2004). It is clear from this that the static disorder component has a relatively minor effect on the form of the scattering around 200, apart from the increase in intensity resulting from a change in the average form factor. The majority of the scattering comes from the thermal displacements for which the average structure factor applies.

3.2. Model 2

Fig. 3.3 of Welberry (2004) showed example diffraction patterns calculated from models involving second-nearest neighbour interactions. Amongst these are examples in which the profile of the diffuse peak that occurs for a nearest-neighbour model can either become much narrower or much broader as a result of introducing the second-nearest neighbour. In particular one example in that text (Fig. 3.3c) shows a peak with a broad flat-topped profile that is reminiscent of the profile of the BZC 200 peak under discussion here. Consequently a second model (Model 2) in which second-nearest neighbour interactions (along **b**) were included, was investigated.

A second reason for investigating a model involving second-nearest neighbours was that the form (III) structure shows a sequence of vertical chains (columns) of molecules in which the first two chains appear similar to those in form (II) and the first substantive difference between the two structures does not appear until the third molecular column, *i.e.* the second-neighbour of the first column (see Fig. 3b). The model

implemented (as described in §2.3.2) is actually a nearest-neighbour model for the four-state variable ξ_i , but this clearly involves second- (and third-) nearest neighbours when the basic σ_i variables representing single molecules are considered.

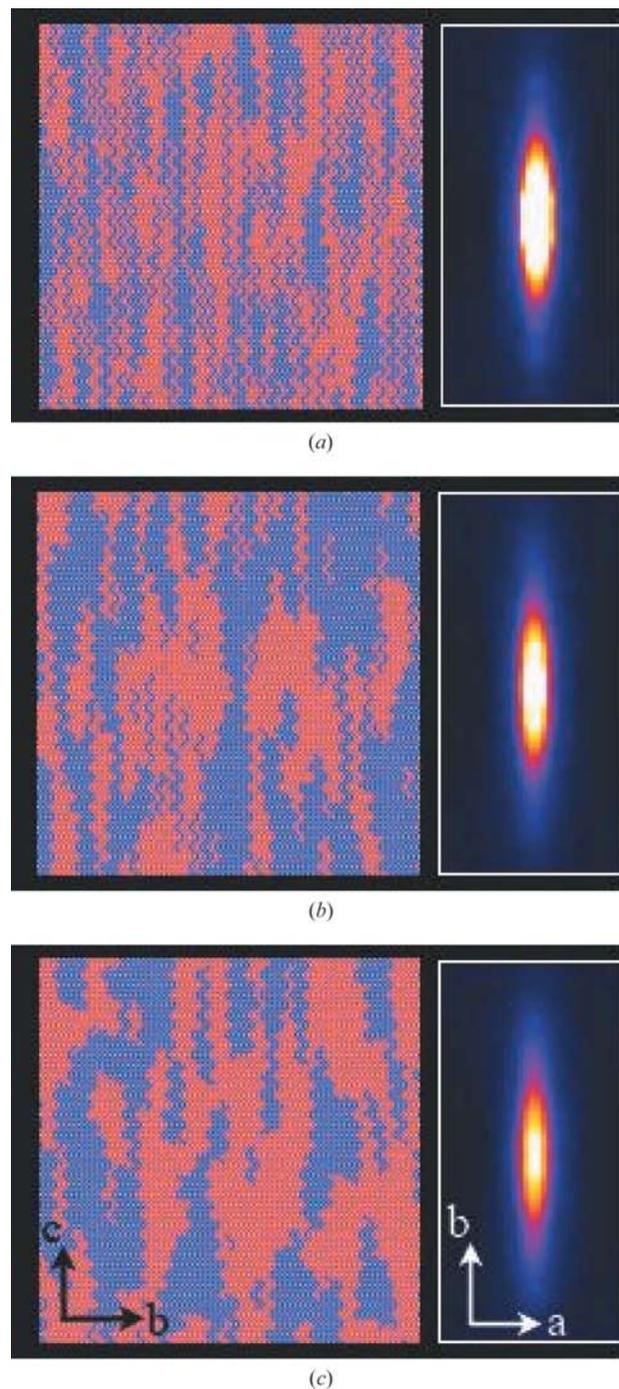


Figure 8 Results from nearest-neighbour Ising model (Model 1) showing the effect on the domain structure viewed down **a** of systematically decreasing the J_b value. Twin component **A** is shown in red, twin component **B** is shown in blue. This ordering in the *bc* plane has relatively little effect on the 200 diffuse diffraction feature viewed down **c**. See text for details.

Table 4

Frequencies of occurrence of different sequences along **b** of four successive molecules of type **A** or **B** for the simulations shown in Fig. 9.

		AA	AB	BA	BB
Fig. 9(a)	AA	24 930	32 139	25 965	27 336
	AB	25 726	27 870	27 702	29 531
	BA	29 923	26 807	29 337	24 465
	BB	29 791	24 013	27 528	29 305
Fig. 9(b)	AA	51 144	58 924	171	131
	AB	174	301	47 471	62 883
	BA	58 811	51 402	166	153
	BB	241	202	62 724	47 470
Fig. 9(c)	AA	46 525	46 463	7830	9552
	AB	8158	6299	48 465	47 907
	BA	47 760	49 346	6867	6559
	BB	7927	8721	47 370	46 619
Fig. 9(d)	AA	55 665	44 941	8938	5267
	AB	8371	9633	40 782	47 602
	BA	44 532	45 022	8606	7906
	BB	6243	6792	47 740	54 328

Fig. 9 shows a series of four snapshots of the domain structure viewed down **a**, produced using Model 2 [see (2)]. The values of the interaction variables J_i used for these

simulations are given in Table 3. This set of examples has been chosen to show the range of effects that can result from introducing the second- and third-nearest neighbours interactions with the use of (3). In Table 4 the frequencies with which different combinations of neighbouring molecular pairs occur in the model realisations are given. It is seen that these generally follow the values of J_1 and J_2 used to induce them.

For the first three examples in Fig. 9 and Table 3 the frequencies with which the four different types of pair (**AA**, **AB**, **BA** and **BB**) occur within each of the double sites in the unit cell (51, 62, 37 or 48, see Fig.4) were all initially set to be equal. Since, in each individual step of the MC simulation, a pair of molecules from one double site is swapped with a pair from a different double site, these initial frequencies are maintained throughout. For the final example, Fig. 9(d), a slight variation of these initial conditions has been used. Here $P_{AA} = P_{BB} = 0.26$ and $P_{AB} = P_{BA} = 0.24$. The values for the nearest-neighbour correlations, C_1 , for these examples are also given in Table 3.

Fig. 9(a), for which both J_1 and J_2 are zero, is very similar to the patterns produced by Model 1. Table 4 shows that the frequencies of all 16 different pair combinations are still of a similar magnitude, although there has been some influence

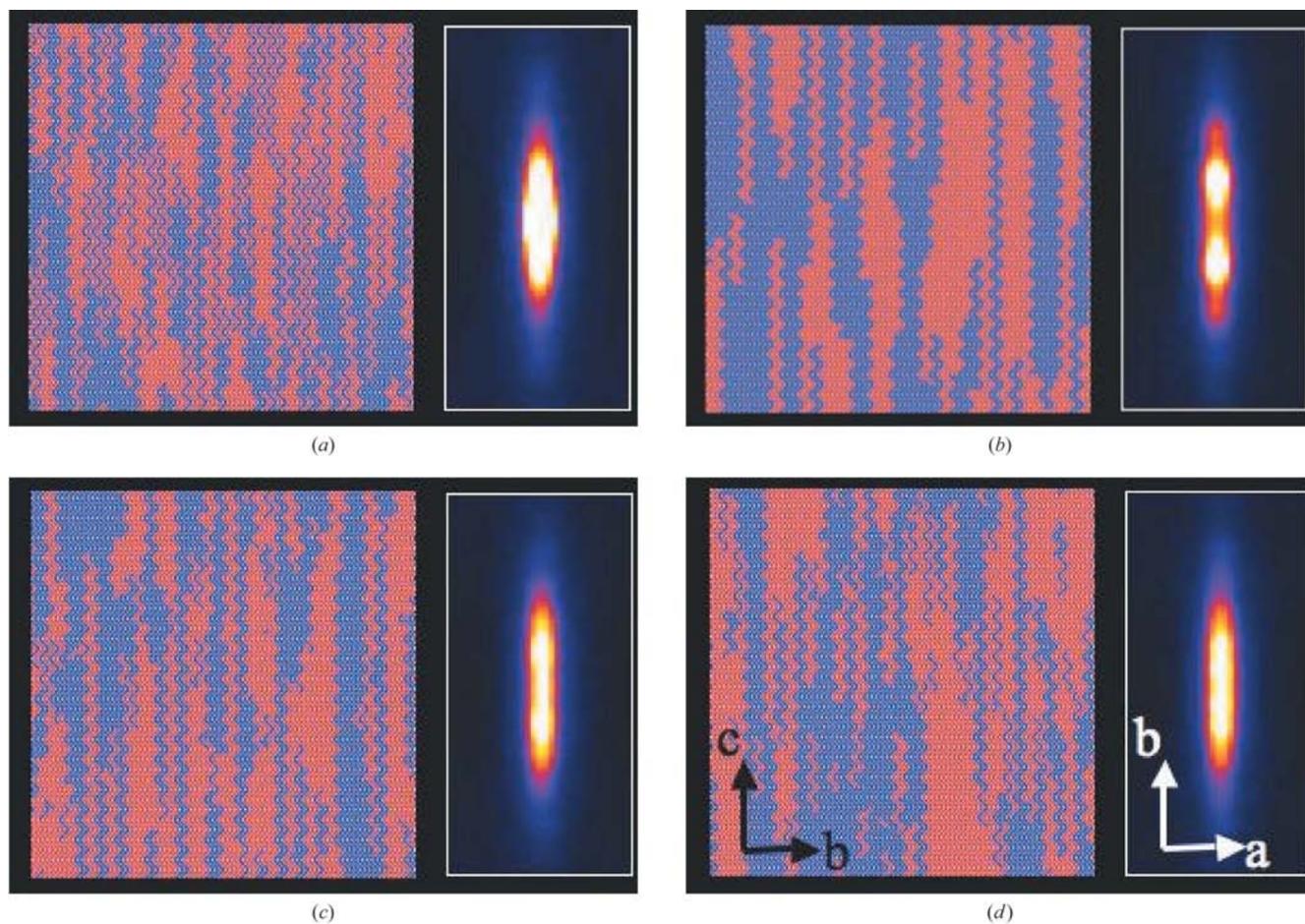


Figure 9

Results from the model involving second- and third-nearest neighbours (Model 2) showing the effect on the domain structure viewed down **a** of different choices of the interaction parameter J_2 (see Table 4). Twin component **A** is shown in red, twin component **B** is shown in blue. This type of ordering in the *bc* plane has a substantial effect on the 200 diffuse diffraction feature viewed down **c**. See text for details.

from the ordering of pairs in the **c** direction. There is thus virtually no occupancy correlation along **b** either within the unit cells or between them. The fall-off of intensity of the diffuse 200 peak along the **b** direction must therefore be due to thermally induced displacement correlations and to changes in the molecular scattering factors away from the 200 Bragg position.

For Fig. 9(b), for which $J_1 < J_2$, a large nearest-neighbour correlation, C_1 , has been induced. Reference to Table 4 shows that the pair combinations that occur most frequently are **|AA|AA|**, **|AA|AB|**, **|AB|BA|** and **|AB|BB|** together with their symmetry equivalents obtained by interchanging **A** and **B**. The lattice realisations are consequently made up of three different types of domains: **AA AA AA AA ...**, **AB BA AB BA ...** and **BB BB BB BB ...**. Since the second of these domain types has a repeat of $4b$ this is clearly the source of the pair of diffuse peaks that are visible at $\pm 1/4b^*$ from the 200 position. As a result of this type of ordering along **b** together with the positive correlation induced along **c** the lattice images clearly show the alternating (blue and red) zigzag rows which represent *ac*-planar molecular bilayers. These molecular bilayers, which slide either into (blue) or out of (red) the plane of the paper, appear to be a fundamental sub-unit of the structure and their presence in the room-temperature form (II) structure is clearly an indication of what eventually occurs at the phase transition.

For Fig. 9(c) the difference $J_2 - J_1$ is much smaller relative to $k_B T$ and hence the realisation is more disordered. The same most frequent pair combinations are still present, but now there is a significant fraction of other combinations. For this structure the two peaks have broadened to produce a single elongated 200 diffuse peak of the type that was targeted.

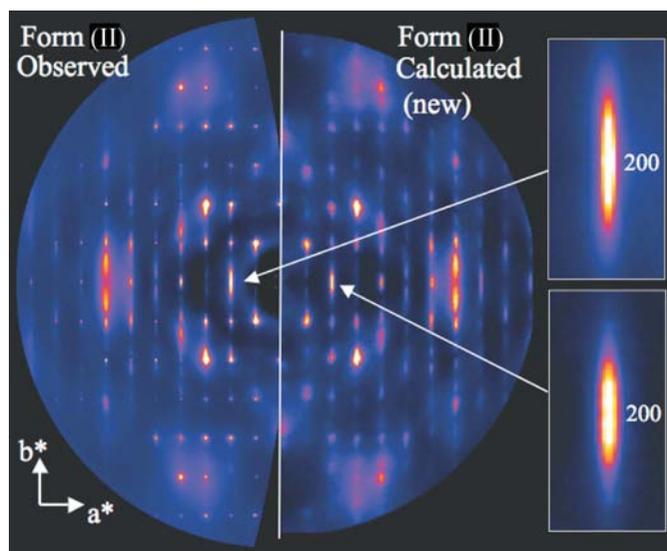


Figure 10

The final calculated *hk0* section showing the best qualitative fit to the observed form (II) data. The model used corresponds to the example in Fig. 9(d). The improvements made to the model through use of a MC simulation involving second- and third-neighbour correlations of molecular types can be seen by comparison of this calculated section to that in Fig. 6(b).

However, there is still a small residual dip in the centre of this peak. The final example, Fig. 9(d), shows that by increasing the relative fraction of the **AA AA AA AA ...** and **BB BB BB BB ...** domains at the expense of the **AB BA AB BA ...** domains this residual dip has disappeared. This required only a small change in the overall pair probability, P_{AA} .

3.3. Diffraction patterns of final room-temperature model

Fig. 10 shows observed and calculated patterns for the complete *hk0* section of the model corresponding to Fig. 9(d). Inset are enlargements of the elongated 200 diffuse peak. The overall agreement is qualitatively very good and the improvement over the original harmonic model in reproducing the 200 diffuse peak is substantial.

This fit was obtained by trial and error using the basic parameterization of Model 2, as described in §2.3.2. No attempt has been made at this stage to refine the model to produce a more quantitative fit since the model that has been created is rather artificial, consisting as it does of a combination of occupancy ordering followed by subsequent imposition of thermal displacements. Nevertheless, the model as described has allowed considerable insight to be obtained on what is happening in the form (II) structure that presages the events that are about to occur at the phase transition.

4. Conclusions

In this paper we have described the development of a model for the room-temperature form (II) polymorph of benzocaine that incorporates, on a local scale, structural features derived from the low-temperature form (III) polymorph. The introduction of this extra information convincingly reproduces those observed diffraction features that an earlier harmonic model was unable to achieve. In both form (I) and form (II) the hydrogen-bonded chains of molecules that extend along the respective *a* axes tend to slide backward and forward along their lengths. While in form (I) the motion is well modelled by a harmonic potential in form (II) there is a degree of anharmonicity that means that some intermolecular contact vectors, that are identical in the average structure, are distributed bimodally with either longer or shorter vectors being preferred to the average. Moreover, there is a tendency for these deviations from average to be correlated to give short-range ordered domains that are the precursors of the two twinned variants of the long-range ordered low-temperature form (III) structure.

The most important outcome of the present work is that it has shown that the MC modelling techniques for analysing diffuse X-ray scattering have now reached a level of precision where subtle details can be modelled and then used to obtain real insight into what is happening in structures at the molecular level. These methods have already benefitted greatly from the increase of available computer processor speeds since the automatic MC refinement method was first introduced (Welberry *et al.*, 1998). With the increase in model crystal sizes now possible together with the increase in the

total number of MC cycles that can be used, calculated diffraction patterns that are of a quality comparable to the observed data can be obtained for an increasing number of two-dimensional sections. With a continued further increase in computer power it is expected that fully three-dimensional analyses should soon become feasible and the level of detail that can be revealed will be correspondingly enhanced.

The study as a whole has added further weight to the premise that diffuse scattering contains a wealth of information over and above that obtainable from Bragg scattering and this can provide meaningful insights into the behaviour of molecules in polymorphic systems. The difficulties inherent in structure prediction of flexible molecular systems (Day *et al.*, 2005, 2009) show that the study of such systems is far from simple, and it is hoped that the insights that diffuse scattering provides can inform the development of more reliable methods of prediction for these important classes of molecular materials, including molecules of industrial and medical significance.

The support of the Australian Research Council, the Australian Partnership for Advanced Computing and the Australian Synchrotron Research Program are gratefully acknowledged. The authors would like to thank Dr Darren Goossens for helpful discussions, Dr Peter Chupas and Dr

Karena Chapman of the Advanced Photon Source for assistance with the collection of the diffuse scattering data, and Dr Aidan Heerdegen for help with data reduction and computing support. Use of the Advanced Photon Source was supported by the US Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract No. DE-AC02-06CH11357.

References

- Beasley, A. G., Welberry, T. R., Goossens, D. J. & Heerdegen, A. P. (2008). *Acta Cryst.* **B64**, 633–643.
- Bernstein, J. (2002). *Polymorphism in Molecular Crystals*. IUCr Monographs on Crystallography. Oxford University Press.
- Butler, B. D. & Welberry, T. R. (1992). *J. Appl. Cryst.* **25**, 391–399.
- Chan, E. J., Rae, A. D. & Welberry, T. R. (2009). *Acta Cryst.* **B65**, 509–515.
- Chan, E. J., Welberry, T. R., Goossens, D. J., Heerdegen, A. P., Beasley, A. G. & Chupas, P. J. (2009). *Acta Cryst.* **B65**, 382–392.
- Day, G. M. *et al.* (2005). *Acta Cryst.* **B61**, 511–527.
- Day, G. M. *et al.* (2009). *Acta Cryst.* **B65**, 107–125.
- Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H. & Teller, E. (1953). *J. Chem. Phys.* **21**, 1087–1092.
- Welberry, T. R. (1985). *Rep. Prog. Phys.* **48**, 1543–1593.
- Welberry, T. R. (2004). *Diffuse X-ray Scattering and Models of Disorder*. IUCr Monographs on Crystallography. Oxford University Press.
- Welberry, T. R., Proffen, Th. & Bown, M. (1998). *Acta Cryst.* **A54**, 661–674.