Acta Crystallographica Section B
Structural Science
ISSN 0108-7681

Diffuse X-ray scattering from 4,4′-dimethoxybenzil, $C_{16}H_{14}O_4$: analysis *via* automatic refinement of a Monte Carlo model

T. R. Welberry* and A. P. Heerdegen

Research School of Chemistry, Australian National University, Canberra, ACT 0200, Australia

Correspondence e-mail: welberry@rsc.anu.edu.au

Received 16 January 2003
Accepted 1 March 2003

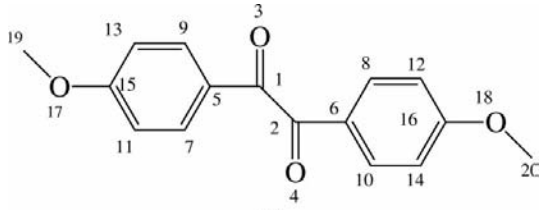
A recently developed method for fitting a Monte Carlo computer-simulation model to observed single-crystal diffuse X-ray scattering has been used to study the diffuse scattering in 4,4′-dimethoxybenzil, $C_{16}H_{14}O_4$. A model involving only nine parameters, consisting of seven intermolecular force constants and two intramolecular torsional force constants, was refined to give an agreement factor, $\omega R = \left[\sum \omega (\Delta I)^2 / \sum \omega I_{\rm obs}^2\right]^{1/2}$, of 18.1% for 118 918 data points in two sections of data. The model was purely thermal in nature. The analysis has shown that the most prominent features of the diffraction patterns, *viz.* diffuse streaks that occur normal to the $[10\bar{1}]$ direction, are due to longitudinal displacement correlations along chains of molecules extending in this direction. These displacements are transmitted from molecule to molecule *via* contacts involving pairs of hydrogen bonds between adjacent methoxy groups. In contrast to an earlier study of benzil itself, it was not found to be possible to determine, with any degree of certainty, the torsional force constants for rotations about the single bonds in the molecule. It is supposed that this result may be due to the limited data available in the present study.

1. Introduction

The present investigation is part of a long-term study of diffuse X-ray scattering from disordered single crystals in which we attempt to understand and explain the origins of the disorder and the resultant diffuse scattering in terms of basic interatomic or intermolecular interactions. In recent years, the advent of cheaper and faster computing has led to a considerable enhancement of our ability to analyse and model such diffuse scattering. In our current methodology, we use a suitably parameterized Monte Carlo model of the disordered system, set up in terms of basic interactions between atoms or molecules, and directly fit this model to the observed X-ray data using least-squares refinement (see Welberry *et al.*, 1998). This methodology is very computer intensive and is only just becoming a viable means of analysis; even with the latest very fast machines, the method cannot be used without some simplification of the problem. An important aspect of the work is to continue to develop the methodology and assess the best strategies that may be employed in given circumstances using the available computing resources, and to gain experience in the quality of fit that might be expected.

In a previous paper (Welberry *et al.*, 2001) we reported a study of the molecular crystal benzil, which had been known for many years to exhibit strong diffuse scattering, but the origins of this scattering had not been explained. That work showed that the particularly distinctive diffuse scattering

patterns of benzil are purely thermal in origin and arise because of a coupling of the internal motions of the molecules with more distant neighbours *via* a network of hydrogen bonds. Furthermore, it was found that the torsional force constants describing the three basic internal degrees of freedom of the molecule are crucial in giving a good fit between the observed and calculated model diffraction patterns. In the present paper we report on a comparable study of 4,4′-dimethoxybenzil, $C_{16}O_4H_{14}$ (I) (hereafter referred to as DMOB). The atom-numbering scheme used in the scheme below is used in later sections. Note that H atoms have been omitted and are not considered throughout the paper.



(I)

Although the central part of the molecule has a very similar molecular geometry to benzil itself, the terminating methoxy groups provide two additional degrees of internal flexibility, and the space group and general molecular packing are quite different in this compound. In addition, it was considered that the presence of additional O atoms in the DMOB structure might lead to further possibilities in terms of hydrogen-bonding network interactions.

The average structure of DMOB was refined by Cannon *et al.* (1989), following an earlier determination by Crowley *et al.* (1983), to a respectable *R* factor of 4.1%. The space group was given as *C*2/*c*, with $a = 21.912$, $b = 4.055$, $c = 15.197$ Å and $\beta = 102.13^\circ$; $Z' = 0.5$, and the molecules lie on sites of *C*2 symmetry. The analysis revealed no unusual thermal displacement parameters and the structure appeared in every respect a normal ordered molecular crystal. Nevertheless, preliminary investigations revealed the presence of strong and quite highly structured diffuse scattering, thus indicating that significant information over and above that contained in the Bragg experiment would be obtainable from an analysis of the diffuse scattering.

2. Experimental

Diffuse scattering data were collected from a needle-shaped crystal, ∼3 mm long, with cross-section dimensions of 0.25 × 0.35 mm. The needle axis corresponded to the crystallographic *b* axis. These data were obtained using the position-sensitive detector (PSD) diffractometer system described by Osborn & Welberry (1990), which uses flat-cone Weissenberg geometry. In the present case, data for DMOB were collected only on reciprocal planes (*h*0*l*) and (*h*1*l*). Because of the small *b* cell dimension, the μ angle required to obtain (*h*2*l*) was beyond the accessible range of the diffractometer. Despite numerous attempts, we were unable to grow crystals of a sufficient size to allow us to shape a sample that would have been suitable for data collection about a different axis. However, since the molecules were clearly resolved in projection along *b*, it was considered that useful information would be obtainable from the two available reciprocal sections.

The PSD was used in a single setting, which was chosen to span the range of approximately $\sin\theta/\lambda = 0.04$–$0.301$ Å^{-1}, using Cu $K\alpha$ radiation. The maximum value of $\sin\theta/\lambda$ was thus a little less than that used in the earlier benzil study, but more of the low-angle data were included. Initially, sections of data consisted of 400 × 400 pixels, at which resolution a single pixel corresponds approximately to the experimental resolution. For use in the analysis, however, the data were rebinned into smaller arrays of 300 × 300 pixels. Previously we have used even smaller arrays for the analysis because of the limitations of available computer time, but since in the present case we use only two sections of data it was feasible to use these somewhat larger arrays. These two sections of data, displayed as false colour images, are shown in Figs. 1(*a*) and 1(*b*). Pixels in these images that are displayed as black (zero value) correspond to unmeasured points, and these were omitted from the analysis. In all, a total of 118 918 measured data points were used.

It should be noted that each section of the data displayed was obtained by combining two data sets in which the sphere of reflection swept through the reciprocal plane in a clockwise or an anticlockwise direction. The purpose of combining these two sets was to overcome the loss of data that occurs when a strong Bragg peak (particularly a low-angle peak) is incident on the wire detector. Since such strong scattering overloads the detector, the collection software steps off the Bragg position until an acceptable count rate is obtained. This process results in an arc segment of reciprocal space (of up to ∼1–2° in ω) being omitted from the collection. By combining the clockwise and anticlockwise data sets, data are lost only where the arcs of missing data on the two sets intersect. These unmeasured pixels can be seen, for example, at the centre of very strong TDS peaks in the (*h*0*l*) section and, more widely, in the (*h*1*l*) section because of the presence of numerous intense low-angle peaks.

3. Monte Carlo refinement method

In this work we utilize the method of automatic refinement of a Monte Carlo (MC) model first described by Welberry *et al.* (1998) and utilized in the previous study of benzil (Welberry *et al.*, 2001). We give here, for convenience, a brief summary of the basic steps. First, the MC model is set up in terms of basic interactions between atoms or molecules, taking into account the average structure as revealed by Bragg analysis and allowing for the fact that intermolecular or interatomic distances might depend on the details of the local occupational ordering. (In the present study, it should be noted, it is assumed that there is no occupational disorder.) Monte Carlo simulation is then carried out on the system at a particular

temperature for a suitably large number of cycles to allow the system to (approximately) reach equilibrium. Diffraction patterns are calculated from the resulting atomic coordinates. Parameters, p_i, specifying the interactions are then adjusted iteratively by performing further MC simulations until the best agreement is obtained between the observed and calculated diffraction patterns. At each stage of the refinement, a goodness-of-fit parameter, χ^2, is used as a quantitative measure of the agreement of the model diffraction pattern with the observed data

$$\chi^2 = \sum_{h,k,l,m} \omega_{hklm} (\Delta I)^2, \tag{1}$$

where

$$\Delta I = I_{\rm obs} - (b_m + f_m I_{\rm calc}). \tag{2}$$

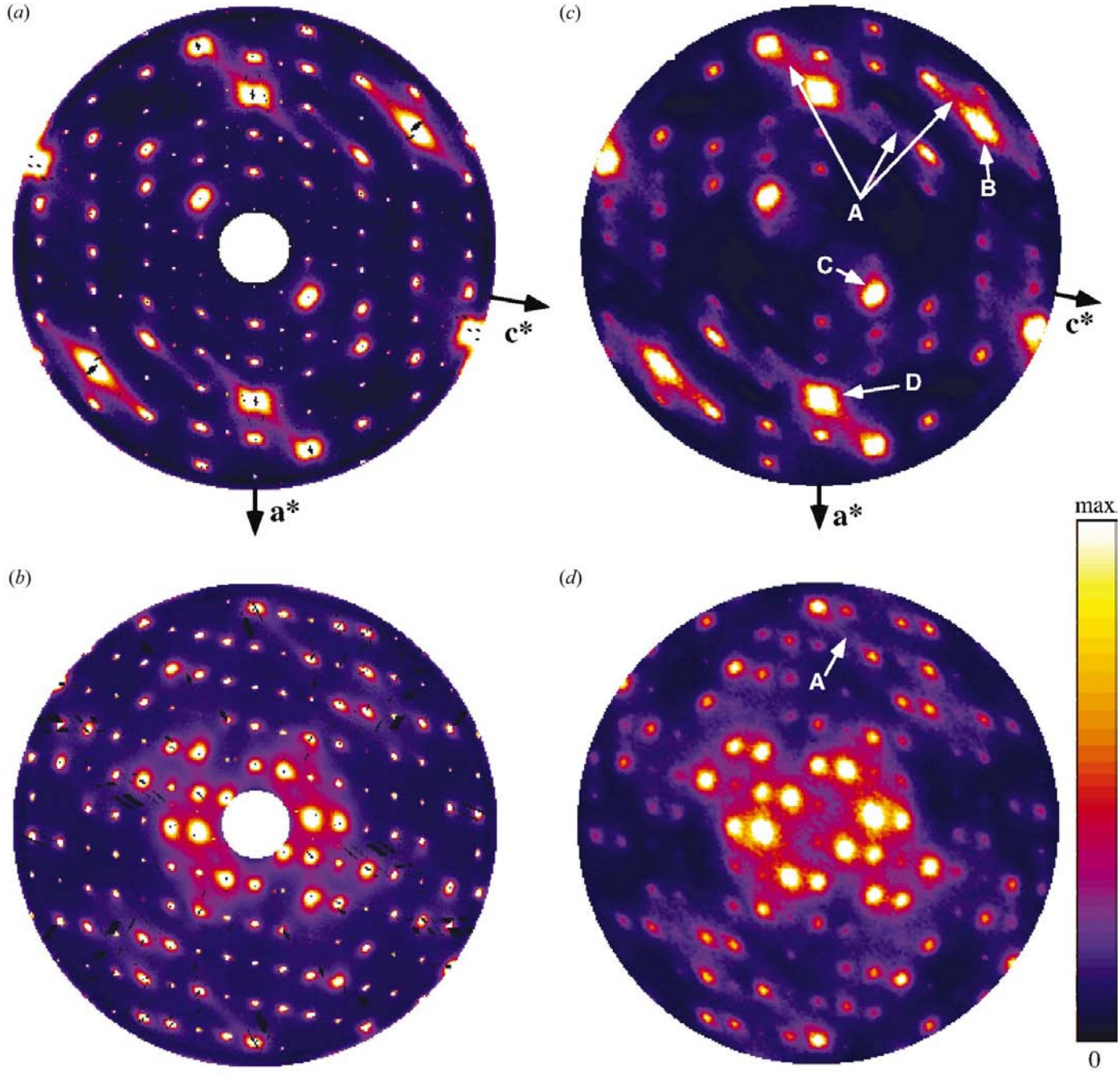


Figure 1
The observed diffraction patterns of 4,4′-dimethoxybenzil and the corresponding diffraction patterns calculated from the final MC model. (*a*) The observed (*h*0*l*) section. (*b*) The observed (*h*1*l*) section. (*c*) The calculated (*h*0*l*) section. (*d*) The calculated (*h*1*l*) section. Note the calculated patterns are of the diffuse scattering only and do not contain Bragg peak information. The outer edge of the diffraction patterns corresponds to the maximum recorded value of $2\theta = \sim 54^\circ$.

The summation in (1) is over all non-integral reciprocal points h, k, l corresponding to individual pixels in the m measured sections of data. f_m is a scale factor and b_m is a background correction applied to section m. [Note that b_m and f_m are determined as described by Proffen & Welberry (1997) and are not included as variables in the least-squares matrix.] ω_{hklm} is the weight for the corresponding data point h, k, l of data plane m. The weights used in the work described here were taken as $\omega_{hklm} = 1/I_{\rm obs}$. Numerical estimates of the differentials of χ^2 with respect to the model parameters are calculated by performing additional MC simulations in which each parameter, p_i, in turn has been incremented by a small amount, δp_i. These differentials are then used to form the least-squares matrix and provide automatic updating of the parameters before the next cycle of the least-squares calculation.

In order to monitor progress in the refinement and to give a comparative assessment with other systems, we make use of the agreement factor

$$\omega R = \left[\sum \omega (\Delta I)^2 / \sum \omega I_{\rm obs}^2\right]^{1/2}. \qquad (3)$$

4. Model crystal

4.1. Molecular representation

In order to allow for the internal flexibility of the DMOB molecule, a z-matrix representation was used. This method, which is commonly used, for example, in *ab initio* molecular orbital calculations (see *e.g.* Hehre *et al.*, 1986), allows the geometry of the molecule to be specified in terms of bond lengths, bond angles and dihedral angles. In the present work we used a model in which all bond lengths, bond angles and dihedral angles were initially set to the values obtained from the average crystal structure determination. During simulation, all of these lengths and angles were kept fixed, with the exception of the five dihedral angles that correspond to rotation about the central C1—C2 bond (φ_1), rotations of the phenyl groups about the C1—C5 and C2—C6 bonds (φ_2 and φ_3), and rotations of the methoxy groups around the C15—O17 and C16—O18 bonds (φ_4 and φ_5). The atoms were numbered as shown in the molecular drawing of (I), so that atom positions were progressively defined proceeding from the centre of the molecule outwards. The z matrix used in the analysis is given in Table 1. The dihedral angles that were allowed to vary during the simulations are given in bold and indicated by an asterisk. It should be noted that two of the four molecular sites in the unit cell contain molecules that are enantiomeric to that given in Table 1. For these molecules, the signs of all dihedral angles (column 4) were reversed.

Table 1
The z matrix representation used in the analysis (see text for details).

Atom	Distance from l (Å)	Angle with lm (°)	Dihedral angle with lmn (°)	l	m	n
C1	–	–	–	–	–	–
C2	1.534	–	–	1	–	–
O3	1.220	116.62	–	1	2	–
O4	1.220	116.62	**122.50***	2	1	3
C5	1.466	120.15	−175.13	1	2	3
C6	1.466	120.15	−175.13	2	1	4
C7	1.382	122.47	**−7.43***	5	1	2
C8	1.382	122.47	**−7.43***	6	2	1
C9	1.392	119.08	179.64	5	1	7
C10	1.392	119.08	179.65	6	2	8
C11	1.379	120.96	178.83	7	5	1
C12	1.379	120.96	178.83	8	6	2
C13	1.366	120.43	−178.90	9	5	1
C14	1.366	120.42	−178.94	10	6	2
C15	1.361	120.21	0.33	11	7	5
C16	1.361	120.21	0.33	12	8	6
O17	1.359	125.32	−179.12	15	11	7
O18	1.359	125.32	−179.12	16	12	8
C19	1.434	117.82	**2.20***	17	15	11
C20	1.434	117.81	**2.21***	18	16	12

Individual molecules were located within the model crystal using a quaternion representation of the orientation matrix and centre-of-mass coordinates. [For a description of quaternions see *e.g.* Goldstein (1980).] Each molecule in the model crystal was thus represented in the computer by 12 variables. These consisted of three centre-of-mass coordinates, four quaternion parameters (one of these is redundant) and five dihedral angles. A stereo view of the average structure of DMOB is shown in Fig. 2.

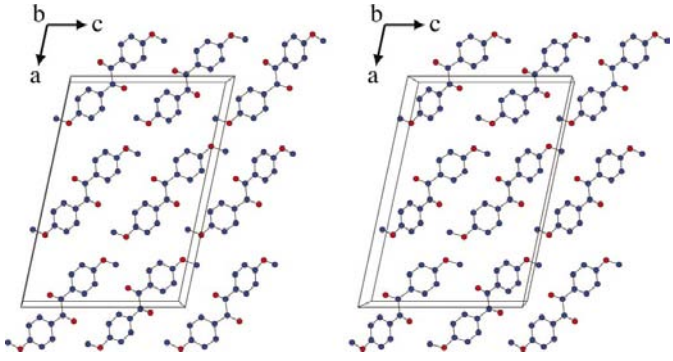


Figure 2
Stereoview of the average structure of DMOB, viewed along b. Only one layer of molecules is shown for clarity.

4.2. Monte Carlo simulations

The automatic MC model refinement methodology outlined in §2 is very computer intensive and is only just becoming a viable means of analysis with the advent of very fast and inexpensive computer processors. Even then it cannot generally be used without some simplification of the problem. The calculations are in two stages: the Monte Carlo simulation and the subsequent calculation of diffraction patterns. The computational resources required depend on a number of factors. For the Monte Carlo simulation, the time for a calculation depends on the system size, the number of interactions per molecule and the number of MC cycles to be carried out. The calculation of the diffraction pattern depends on the quality of the pattern required. In order to be able to produce a calculated diffraction pattern of a quality (high resolution, low noise) comparable to that of the observed

Table 2
The parameters used in the least-squares calculations and their mean refined values.

The s.u. values in column 4 were estimated by observing the fluctuation of the parameter values over 40 cycles of iteration. For the intercolumn vectors, the force-constant index in column 1 corresponds to the vector labelling shown in Fig. 3. For each of the original force constants, column 5 gives the number of the corresponding constant used in the simplified model. The last column gives the refined value of those nine constants.

Force constant	Vector distance, d_{0i} (Å)	Type of contact	Final value (s.u.)	New parameter	New refined value
F_1	3.396	O—Me	8.4 (9)	1	6.0 (2)
F_2	3.474	O—Me	4.1 (5)	1	
F_3	3.426	O—O	4.6 (5)	2	3.8 (3)
F $_4$	***3.474***	***O—Ph***	***4.3 (3)***	***7***	***3.6 (2)***
F_5	3.543	Me—Me	6.4 (4)	3	5.3 (2)
F_6	3.632	O—Me	5.3 (5)	1	
F_7	3.621	O—Ph	4.6 (2)	4	4.8 (2)
F_8	3.595	O—Me	5.1 (5)	1	
F $_9$	***3.693***	***O—Ph***	***4.3 (3)***	***7***	
F $_{10}$	***3.658***	***C—Ph***	***4.7 (6)***	***7***	
F_{11}	3.674	Me—Ph	4.9 (3)	5	4.9 (2)
F $_{12}$	***3.737***	***Ph—Ph***	***4.9 (3)***	***7***	
F $_{13}$	***3.699***	***O—Me***	***5.0 (3)***	***7***	
F $_{14}$	***3.698***	***C—O***	***4.9 (3)***	***7***	
F $_{15}$	***3.830***	***Ph—Ph***	***4.1 (4)***	***7***	
F $_{16}$	***3.834***	***Ph—Ph***	***4.9 (4)***	***7***	
F_{17}	3.961	O—O	4.2 (5)	2	
F_{18}	3.948	Ph—Ph	7.6 (7)	6	6.5 (2)
F $_{19}$	***3.959***	***Ph—Ph***	***4.2 (7)***	***7***	
F $_{20}$	***3.963***	***Ph—Ph***	***4.6 (3)***	***7***	
F $_{21}$	***4.010***	***C—Ph***	***4.1 (3)***	***7***	
G_1	–	Torsion	0.0	8	0.0
G_1–G_3	–	Torsion	0.0	8	
G_4–G_5	–	Torsion	0.26	9	0.13

diffraction patterns, we have found in previous studies that a minimum system size of ~32 × 32 × 32 unit cells is required. In the present case, however, with four molecules per unit cell, it was found satisfactory to use a simulation system size of only 32 × 16 × 32 unit cells. With 80 non-H atoms per unit cell, the model crystal thus contained 1 310 720 individual atoms.

To represent the intermolecular interactions we use the idea of 'effective' interactions. Each molecule is in contact with 14 neighbouring molecules, and if all atom–atom contacts were to be included the number of these would be too large for MC simulations to be feasible. With each molecule in its equilibrium position, all intermolecular atom–atom contact distances less than 4.1 Å were computed. A smaller number of contacts was obtained by pruning this list. In the final list, each intermolecular contact was represented by a small number of atom–atom contacts, which were considered sufficient to define the equilibrium separation and mutual orientation of neighbouring molecular fragments. For each molecular site, a total of 78 intermolecular atom–atom contact vectors were used, and these consisted of 21 distinct symmetry-unrelated vector types. Of the 78 contact vectors, 44 were between a given molecule and those immediately above and below it along the *b* axis. Similarly, of the 21 symmetry-unrelated vector types, 11 were involved in these contacts. The structure clearly consists of columns of these closely packed molecules and the remaining 34 contacts (10 vector types) may be considered as intercolumn interactions. These are shown in Fig. 3.

Interactions between the molecules were included along each of these vectors in the form of Hooke's law (harmonic) springs with force constants, F_i, corresponding to the 21 different vector types. The intermolecular contribution to the total energy of the system was thus

$$E_{\rm inter} = \sum_{\substack{\text{all contact}\\ \text{vectors}}} F_i (d_i - d_{0i})^2. \qquad (4)$$

Here, d_i is the instantaneous length of a particular vector of type i, and d_{0i} is the equilibrium length of that vector. Values for d_{0i} are given in Table 2.

In addition to the intermolecular spring constants, terms were also included in the MC energy to provide intramolecular forces in the form of simple torsional force constants, G_i, which depend on the deviations, $\Delta\varphi_i$, of the angles φ_1, φ_2, φ_3, φ_4 and φ_5 from their equilibrium values. These intramolecular contributions to the energy were

$$E_{\rm intra} = \sum_{\substack{\text{all}\\ \text{molecules}}} \sum_i G_i (\Delta\varphi_i)^2. \qquad (5)$$

It is assumed that, because of symmetry, $G_2 = G_3$ and $G_4 = G_5$. The total energy of the crystal is then expressed in the form

$$E_{\rm tot} = E_{\rm inter} + E_{\rm intra}, \qquad (6)$$

and this expression was used in the application of the normal Monte Carlo algorithm as follows. A site in the crystal was

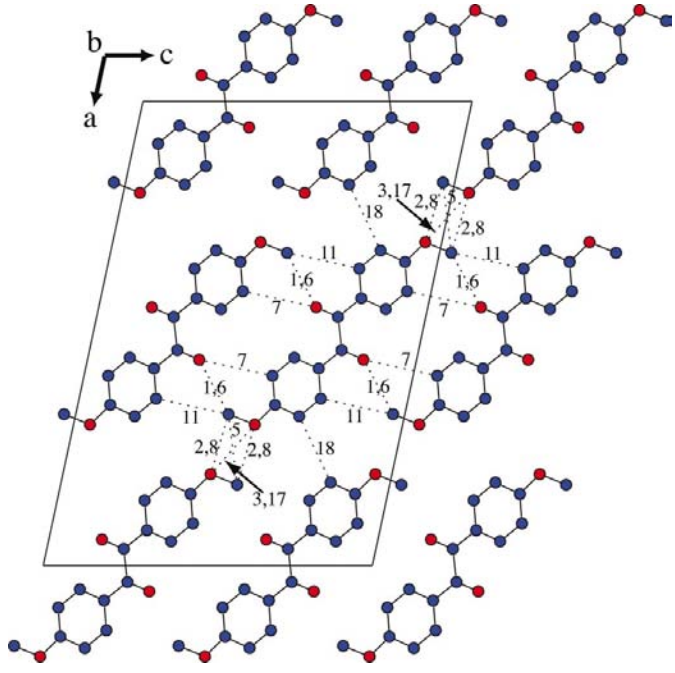


Figure 3
Contact vectors used in the analysis. The numbering corresponds to the force constants given in Table 2. Labels consisting of two integers separated by a comma refer to superimposed vectors. The contact vectors between the central molecule and those immediately above and below it (viewed along *b*) have been omitted for clarity.

chosen at random, and the set of 12 variables defining the conformation, orientation and position of the molecule were subjected to a small random increment. The contribution to the energy of all terms that are dependent on the altered variables was computed before and after the change was made. The energy difference, $\Delta E = E_{\rm new} - E_{\rm old}$, was used to decide whether the new configuration should be kept or the system should be returned to the original configuration. A pseudo-random number, η, chosen uniformly in the range [0,1], was compared with the transition probability, P, where

$$P = \exp(-\Delta E/kT). \quad (7)$$

T is the temperature and k is Boltzmann's constant. For $\eta < P$, the new configuration was accepted, while for $\eta > P$ it was rejected and the system returned to the original state. A Monte Carlo cycle is defined as that number of such individual steps needed to visit each site once on average. In the present work we assumed $kT = 1$ throughout, and consequently the force constants determined were relative to this energy scale.

The number of MC cycles required to reach a point sufficiently close to equilibrium is difficult to judge. Broad features of a pattern will be reproduced with relatively few cycles, while features corresponding to relatively long-range effects may take many cycles to settle down. For the benzil study, 400 cycles of iteration were used, and for preliminary stages of the present work, this same number was adopted. However, for final cycles of refinement, this number was increased to 1600. With this number of cycles, a single MC calculation used ~200 min of c.p.u. time on a 700 MHz Pentium III processor.

4.3. Calculation of diffraction patterns

For calculating the diffraction pattern from the output of the MC runs we used the program *DIFFUSE* (Butler & Welberry, 1992). This algorithm obtains the diffraction pattern by taking the average of a large number of diffraction patterns calculated from small regions (or lots), chosen at random, from the simulated real-space array. This process has the effect of removing the high-frequency noise that necessarily occurs if a single calculation from the whole simulated array is performed. It is only necessary that the 'lot size' is large enough to include all significant non-zero correlations. In the case of DMOB, the observed diffraction patterns shown in Fig. 1 display diffuse streaks whose widths correspond to relatively long-range correlation effects. In order to ensure that such effects would be reproduced by the *DIFFUSE* calculations, it was necessary to use a 'lot size' large enough (in the a and c dimensions) to be able to reproduce such features. After some preliminary testing, a lot size of $8 \times 5 \times 6$ unit cells was chosen, and the number of lots sampled in a given calculation was 100. Each of these diffraction pattern calculations took about 150 min of c.p.u. time on a 700 MHz Pentium III processor. Even at this level of calculation the patterns were rather noisy, and further improvement in the signal-to-noise ratio was obtained by averaging the patterns using their inherent twofold symmetry. In order to obtain diffraction patterns from the final model for presentation in this paper, rather better (and proportionately lengthy) calculations were performed using 250 of the same sized lots. These calculations represent something close to the best that can be obtained from the chosen simulation crystal size of $32 \times 16 \times 32$ unit cells; any further improvement would require a larger crystal size.

As was the case in the benzil study (Welberry *et al.*, 2001), it was apparent at an early stage that the calculated intensities at high angles were too strong compared with those at low angles. A similar effect was also found in an earlier study of urea/alkane inclusion compounds (Welberry & Mayo, 1996). In those studies the effect was taken account of by using an overall Debye–Waller factor to modify the standard atomic scattering factors. The problem seems to arise because the true intensity, in sections normal to b, should be obtained by integrating over the whole columns of superposed atoms in the b direction, whereas our calculations only involve very limited integration. Even though the calculations in the present study were from a three-dimensional model, we suspect that the need still arises, partly because of the relatively small lot size (in the b direction) but also because of the limited model crystal size. For the present study, we overcame the problem, as before, by introducing an overall Debye–Waller factor, $\mathbf{B} = 8\pi^2\overline{u^2}$, of 5.0, which was used throughout the analysis.

5. Results

5.1. Progress of refinements

Initially, all force constants, F_i, were set to a value of 5.0, and iterations were carried out using increments, $\delta_i = 1.0$, on these parameters when calculating the differentials for the least-squares matrix. The torsional force constants, G_i, were all initially set to zero, and for these parameters, increments of $\delta_i = 0.05$ were used for calculating the differentials. Numerous least-squares refinement runs were carried out in order to explore the effect of varying such condition specifications as the number of Monte Carlo cycles and the quality of the *DIFFUSE* calculations. It was apparent from these preliminary runs that the initial model gave a good agreement between observed and calculated diffraction patterns, and the least-squares refinement was only able to reduce the agreement factor, ωR, by a relatively small amount (from 0.185 to 0.179 in the best case). After this small initial reduction, subsequent iterations resulted in the system fluctuating within ~0.002 of the minimum. Observation of the behaviour of the force-constant parameters during this period revealed some interesting trends.

First, there was a tendency for the average magnitude of the force constants to increase with time, thereby resulting in a reduction of the mean-squared atomic displacements. At the same time, the value of the scale factor, f_m, in (2) increased. We believe that this behaviour results from the fact that the distribution of the diffuse scattering in the diffraction pattern depends primarily on correlations between atomic displacements. The magnitude of the displacements affects the overall

intensity of the scattering and the overall variation with diffraction angle. If data are available over a wide range of diffraction angles, this latter variation is sufficient to define the magnitude well enough for the scale of the displacements to be fixed. In the present experiments, in which data are recorded over only a relatively small range of 2θ, it appears that there is too strong a dependency between the magnitudes of displacement and the scale factor for the former to be determined with any certainty. Therefore, it was decided in subsequent refinements to renormalize the force constants after each cycle, in order to maintain overall mean-square displacement values comparable to the B values obtained from the Bragg structure determination. This renormalization, which was achieved using a mean value of 5.0 for the F_i parameters, had essentially no effect on the ωR agreement factor.

A second feature of the refinements was that the torsional force constants, G_1–G_3, consistently refused to move away from zero. Any calculation made with a positive value of these parameters resulted in a simulation with a significantly increased ωR agreement factor. In contrast, the torsional constants for the methoxy groups, G_4 and G_5, persistently tended to move to a non-zero value. Over the period of iteration, their values increased gradually to attain the figures given in Table 2. However, since there was no improvement in ωR over the cycles in which these averages were taken, it seems likely that this result merely indicates that the model fit is independent of the value of G_4 and G_5 rather than that G_4 and G_5 should have a significant positive value.

A further feature of the behaviour of the force-constant parameters was that, while some particular force constants consistently refined to be larger than the average and others consistently smaller, most remained within one or two standard deviations of the mean value of 5.0. Even after many cycles of iteration, the ratio of the highest to lowest F_i was only $\sim$2. This behaviour was taken to indicate that the system of springs making up the interaction environment of the molecules was over-parameterized; that is, if one spring reduced in strength, others were able to take up the slack without substantially altering the overall pattern of molecular motion giving rise to the diffuse scattering.

We therefore decided to simplify the model by reducing the number of refinable force constants. Although this simplification might have been achieved in a number of different ways, it was decided that a reasonable approach was to use force constants that depended on the type of interaction. Thus, for example, the original force constants F_1, F_2, F_6 and F_8, which all corresponded to interactions between an O atom in one molecule and a methyl group in a neighbouring molecule, were assumed to be equal and to correspond to the new variable 1 (see the fifth column in Table 2). Similarly the new variable 2 corresponds to the original F_3 and F_{17} interactions between pairs of O atoms. Since the observed data contained very little information in directions parallel to b, all force constants involved in interactions between a molecule and that immediately above or below it in the b direction were assumed to be equal and to correspond to the new variable 7 (note that these cases are given in italics in Table 2). Two torsional variables were retained in this simpler model, even though the torsional constants used in the initial model did not appear to be contributing significantly. One of these involved the original force constants G_1–G_3 and the second the constants G_4 and G_5.

Comparable refinements were then carried out with this much simplified model, and parameter values were monitored over the course of the iterations in order to estimate parameter variances. Starting, as before, with all spring constants set to 5.0 and torsion constants to 0.0, the refinement proceeded in a very similar fashion to the full model, with some initial relatively small improvement in ωR followed by a long plateau with the system fluctuating within $\sim$0.002 of the minimum. The lowest value of ωR achieved for this model was 0.181, only 1 s.u. greater than that for the full model. The values of the force constants given in the last column of Table 2 were obtained by averaging over 60 cycles of least-squares iteration. The behaviour of the torsional constants was as before, with the G_1–G_3 parameters consistently remaining zero and the G_4 and G_5 parameters tending upwards. The value given for this parameter in Table 2 is that from the final cycle.

5.2. Diffraction patterns for the final model

Figs. 1(c) and 1(d) show diffraction patterns calculated for the final model, and these can be compared with the observed patterns shown in Figs. 1(a) and 1(b). It should be noted that these calculated patterns are of the diffuse scattering alone and contain no Bragg peak component. It has been our practice in the past to superpose the Bragg peaks onto the calculated diffuse patterns in order to aid in the identification of diffraction features. In the present case, the rather distinct TDS peaks that occur around the Bragg positions obviate the need for this superposition. These patterns have been calculated to the best quality that can be achieved with the chosen crystal size of $32 \times 16 \times 32$ unit cells and are rather better than the patterns actually used in the fitting. Nevertheless, it is clear that Figs. 1(c) and 1(d) are still rather noisy compared with the observed patterns. Further improvement can only be obtained by increasing the system size with corresponding increases in all aspects of the computations.

Although the value of ωR testifies to the good quantitative agreement that has been achieved, it is useful also to point out a number of qualitative points of agreement. First, there are elongated streaks (indicated by A in Figs. 1c and 1d) that extend along rows of Bragg peaks. There are several of these streaks in the ($h0l$) section and one is also visible in the ($h1l$) section. These streaks are somewhat more distinct in the observed data than in the more noisy calculations. The strong diffuse peak B is very elongated, with a length much greater than its width. Diffuse peak C, on the other hand, is nicely elliptical, with its longer axis normal to a vector between the peak and the origin. Diffuse peak D has a more complex shape something like a rhombus. Here there is also some slight discrepancy with the observation, for although the basic

rhombus shape is in agreement, the relative widths of the peak in the vertical and horizontal directions do not quite match.

5.3. Exploration of the simulated model

Although the final model is very simple, it nevertheless produces quite complex diffraction patterns, and it is instructive to look at the results of the Monte Carlo simulation in order to get some insight into exactly what the molecules are doing. In Fig. 4(*a*) we show a plot of the superposition of all of the unit cells in the simulation onto a single cell. This plot clearly shows the mean-square atomic displacement patterns, which are fairly isotropic towards the centre of the molecule but progressively anisotropic towards the methoxy extremities. Fig. 4(*b*) shows a similar plot but with the coordinates in each unit cell plotted relative to the centre of mass of the central molecule. For the central molecule itself, the local shift of origin means that the translational part of the motion has been removed, leaving the displacements due to libration. In this plot it might also be supposed that if there were very strong intermolecular correlations, the plots for the other (neighbouring) molecules in the cell would also be significantly modified. This does not appear to be the case from a simple visual inspection of Fig. 4(*b*).

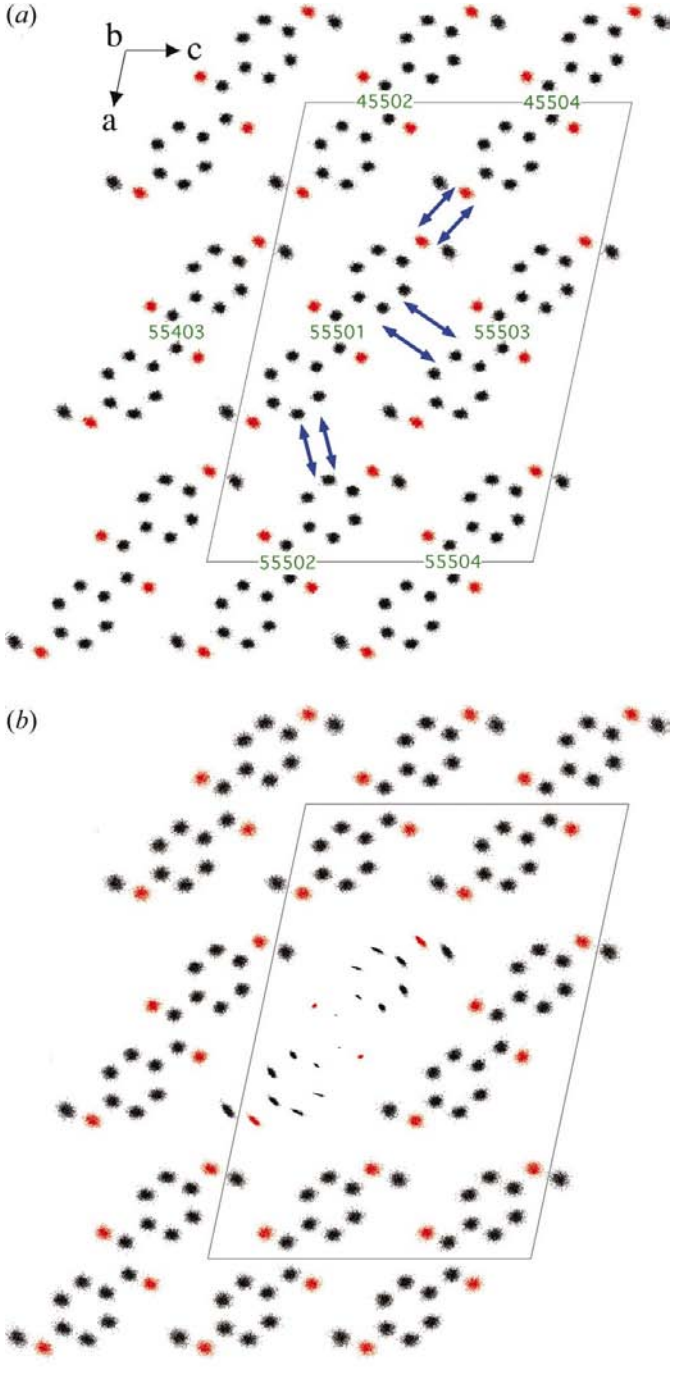


Figure 4
Superposition plots of the atoms in the final simulation. In (*a*), atoms from a single layer (normal to *b*) are plotted. In (*b*), all atoms are plotted relative to the centre of mass of the central molecule. The double arrows in (*a*) marking different intermolecular contacts are referred to in Fig. 5. The labelling of molecular sites in (*a*) follows the *ORTEP*II method (Johnson, 1976).

In order to perform a more sensitive test of the way neighbouring molecules influence each other, we calculated correlation coefficients between the displacements of the centres of mass of various neighbours. More particularly, we measured the correlation between the displacement in a given direction of a central molecule (55501) with the displacement in the same direction of a neighbouring molecule (*e.g.* 45504). [Note that this labelling of molecular sites follows the *ORTEP*II method (Johnson, 1976).] Some example plots are shown in Fig. 5. The orientation of these figures is the same as that of the molecular plots of Fig. 4, with the horizontal axis parallel to the crystallographic *c* axis.

The correlation plots corresponding to the two superposed neighbours 44504 and 45504 have their major axis parallel to the length of the molecule in the general direction [10$\bar{1}$]. It is

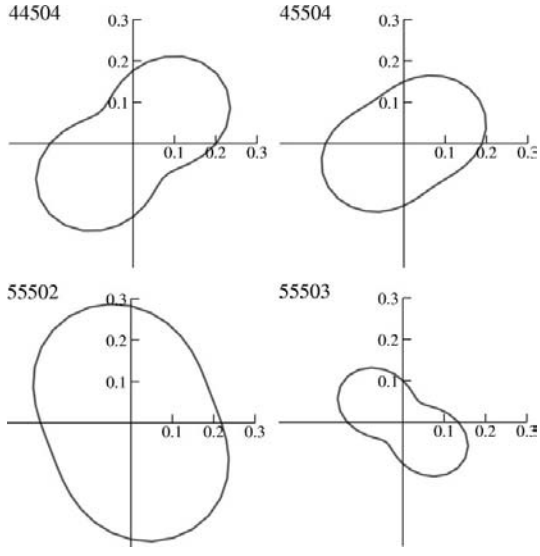


Figure 5
Correlations of the centre-of-mass displacement of a central molecule (55501) with those in neighbouring molecules. The plots show the correlation coefficient as a function of displacement direction. Note that the directions of maximum correlation broadly correspond to the directions of the intermolecular contacts shown by double arrows in Fig. 4(*a*).

these correlations, though only of moderate magnitude, that are responsible for the diffuse streaks *A* in Fig. 1. It is noteworthy that these interactions involve hydrogen bonding between the O atom of one methoxy group and the methyl of the neighbouring methoxy group. The fact that the correlation in the orthogonal direction for these neighbours is very weak results in the peak, for example *B*, being very broad in the direction of the streak. Similarly, the plot for the 55503 neighbour shows very weak correlation in the $[10\bar{1}]$ direction, thus reflecting how neighbouring molecules can easily slide over each other. The 55502 correlation is interesting since the major axis of this figure corresponds to the single highest correlation between molecules in the *ac* plane. Fig. 3 shows, however, that there is only a single direct spring linkage to this neighbour (F_{18}). That this contact is nevertheless important is shown by the fact that the corresponding spring force constant is one of the largest.

In Fig. 6 we examine the mutual dependence of the internal torsion angles. Each point on one of the plots represents the deviations of two of the torsion angles in a given molecule away from their mean values. Each plot contains points for all molecules in the simulation that are in one of the four molecular sites in the unit cell. It can be seen that φ_1 is negatively correlated with φ_3 (and φ_2), but there is also a slight negative correlation between φ_2 and φ_3. The variance of φ_1 is a little less than that of φ_3. The actual mean and s.u. values of the φ_i parameters and the correlations between them are given in Table 3. Since the torsional force constants G_1–G_3 are all zero in the simulation, the restoring force limiting the extent of these angular variations must be provided by the cocoon of intermolecular F_i springs. φ_5 is not correlated with φ_3 but has a much smaller variance. This smaller variance clearly arises from the non-zero value of the G_5 torsional force constant.

Table 3
The mean and s.u. values of the φ_i torsion angles and the correlations between them.

			Correlation with				
Torsion angle	Mean	s.u.	φ_1	φ_2	φ_3	φ_4	φ_5
φ_1	122.8	3.5	1.0	−0.21	−0.22	−0.04	−0.04
φ_2	−7.9	5.3		1.0	−0.20	−0.01	−0.01
φ_3	−7.9	5.3			1.0	−0.01	−0.01
φ_4	5.0	1.4				1.0	0.01
φ_5	5.0	1.4					1.0

6. Conclusion

In this paper we have described the development and refinement of an MC model that quantitatively fits the observed diffuse diffraction patterns of 4,4′-dimethoxybenzil (DMOB). Although the diffraction patterns are quite complex, with a number of interesting features, the model developed to fit the data is very simple and involves few adjustable parameters. It was initially of concern that the refinement process resulted in only a small improvement of the *R* factor from the starting position in which all intermolecular force constants were set to be equal. Moreover, after an initial model, in which we used 21 spring force-constant parameters and three internal torsional force constants, was simplified to a model using only seven force constants and two torsional constants, virtually no reduction in the quality of the fit was obtained. It seems likely that a further reduction in the numbers of refinable parameters could be made without significantly worsening the fit. We conclude that the reason for this behaviour is that the rather complex observed diffraction patterns occur as a natural consequence of the flexibility of the molecule and this very simple packing model, and there is very little additional information contained in the diffuse scattering.

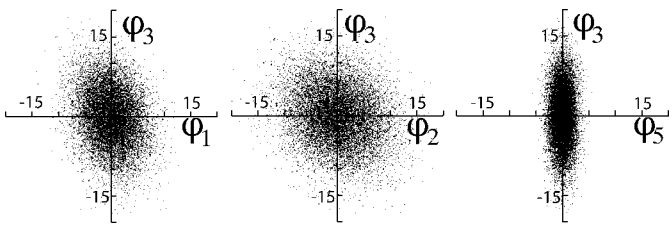


Figure 6
Plots showing the mutual dependence of the internal torsion angles. Angles are shown in degrees as deviations from the average.

In a sense, therefore, the outcome of the present study is a little disappointing, since it had been hoped that the increase in complexity of the molecule would result in the need for a more complex model. On the other hand, it might be considered an encouraging result that the basic intermolecular interactions can be modelled with few parameters. This result should mean that, for systems that contain additional types of disorder (*e.g.* occupational disorder), further variables can be added to describe this disorder while keeping the overall analysis within tractable bounds.

As found in the earlier study of benzil, it is clear that hydrogen bonding plays an important role in the structure of DMOB and the form of the diffuse scattering. The most striking features of the diffraction patterns, *i.e.* the diffuse streaks that occur normal to $[10\bar{1}]$, are due to correlations that are transmitted along rows of molecules in that direction *via* interactions between the methoxy groups of neighbours, each interaction involving two hydrogen-bonding contacts.

In contrast to our earlier study of benzil, the current analysis did not enable realistic measurements of internal torsional force constants to be made. For the angles φ_1–φ_3, all attempts to include non-zero force constants led to significantly worse agreement factors. It appeared that the cocoon of intermolecular springs surrounding the molecule provided sufficient restoring force for perturbations of these angles, so that no additional torsional constants were required. It seems likely that this result stems from the rather limited data that we were able to collect for this material. Because of the needle-shaped habit of the crystals and the inability to grow large samples that could be subsequently reshaped, we were unable to collect data in sections other than those normal to the needle axis. Even here, because of the short *b* axis, the (*h*2*l*) section was beyond the instrumentally accessible region

and the whole analysis was carried out with just the ($h0l$) and ($h1l$) sections. It is possible, if more extensive data could be measured, that sufficient signal would be present to allow external and internal forces to be distinguished. For φ_4 and φ_5, it was found that a torsional force constant comparable to those found in benzil could be included in the model with no adverse effect on the level of agreement, but equally the model was not significantly worse without it. Again, it seems likely that a more extensive data set could resolve this question. The present study clearly points to the need for future studies to include more extensive data sets.

One feature of the present study that was less than satisfactory was the need to fix the average value of the force constants. Although this requirement might be obviated by the inclusion of more (particularly higher-angle) data, alternative strategies may be worth pursuing. One possibility would be to perform an average structure refinement (*i.e.* with Bragg intensities) using a model in which the atomic displacements were formulated in terms of comparable centre-of-mass and internal-angle librations. The derived distributions for the displacements and angles could then be used as constraints in the diffuse scattering analysis. The model to be used for the diffuse scattering could then be set up initially with randomly chosen variables that reproduce these distributions. Then in carrying out the MC simulation, instead of allowing random shifts of the variables for a randomly chosen molecule, a method could be adopted in which the variables from two different molecules are interchanged. In this way, the average structure would be maintained while the system refined to fit the diffuse scattering. Such a method has been described for simple systems in the context of the reverse Monte Carlo (RMC) technique (see Proffen & Welberry, 1998) but has so far not been implemented for more complex systems.

The present study is part of a long-term program in which we aim to investigate a range of flexible molecules in order to gain experience in applying the MC fitting procedure used here for analysing their diffuse X-ray scattering patterns. It is expected that future studies will include cases in which there is static (substitutional) disorder as well as the purely thermal disorder modelled here. In this context, the present results and the experience gained in the study will provide a useful basis for future studies of more complex systems.

Support of this research by the Australian Research Council is gratefully acknowledged. Support by the National Facility of the Australian Partnership for Advanced Computing is also gratefully acknowledged.

References

Butler, B. D. & Welberry, T. R. (1992). *J. Appl. Cryst.* **25**, 391–399.
Cannon, J. R., Patrick, V. A. & White, A. H. (1989). *Aust. J. Chem.* **42**, 1631–1645.
Crowley, J. I., Balanson, R. D. & Mayerle, J. J. (1983). *J. Am. Chem. Soc.* **105**, 6416–6422.
Goldstein, H. (1980). *Classical Mechanics*, 2nd ed. Reading, Massachusetts: Addison-Wesley Publishing Company.
Hehre, W. J., Radom, L., Schleyer, P. v. R. & Pople, J. A. (1986). *Ab Initio Molecular Orbital Theory*. New York: John Wiley and Sons.
Johnson, C. K. (1976). *ORTEP*II. Report ORNL-5138. Oak Ridge National Laboratory, Tennessee, USA.
Osborn, J. C. & Welberry, T. R. (1990). *J. Appl. Cryst.* **23**, 476–484.
Proffen, Th. & Welberry, T. R. (1997). *Acta Cryst.* A**53**, 202–216.
Proffen, Th. & Welberry, T. R. (1998). *Phase Transit.* **67**, 373–397.
Welberry, T. R., Goossens, D. J., Edwards, A. J. & David, W. I. F. (2001). *Acta Cryst.* A**57**, 101–109.
Welberry, T. R. & Mayo, S. C. (1996). *J. Appl. Cryst.* **29**, 353–364.
Welberry, T. R., Proffen, Th. & Bown, M. (1998). *Acta Cryst.* A**54**, 661–674.