Acta Crystallographica Section A
Foundations of Crystallography
ISSN 0108-7673

Received 10 July 2000
Accepted 2 October 2000

Diffuse X-ray scattering from benzil, $C_{14}H_{10}O_2$: analysis *via* automatic refinement of a Monte Carlo model

T. R. Welberry,[a]* D. J. Goossens,[a] A. J. Edwards[a] and W. I. F. David[b]

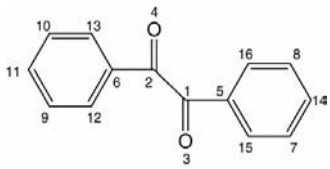
[a]Research School of Chemistry, Australian National University, Canberra, ACT 0200, Australia, and [b]ISIS, Rutherford Appleton Laboratory, Didcot, Oxfordshire, England. Correspondence e-mail: welberry@rsc.anu.edu.au

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 benzil, diphenylethanedione, $C_6H_5-CO-CO-C_6H_5$. A model involving 13 parameters consisting of 11 intermolecular force constants, a single intramolecular torsional force constant and a local Debye–Waller factor was refined to give an agreement factor, $R = [\sum \omega(\Delta I)^2 / \sum \omega I_{obs}^2]^{1/2}$, of 14.5% for 101324 data points. The model was purely thermal in nature. The analysis has shown that the diffuse lines, which feature so prominently in the observed diffraction patterns, are due to strong longitudinal displacement correlations. These are transmitted from molecule to molecule *via* a network of contacts involving hydrogen bonding of an O atom on one molecule and the *para* H atom of the phenyl ring of a neighbouring molecule. The analysis also allowed the determination of a torsional force constant for rotations about the single bonds in the molecule. This is the first diffuse scattering study in which measurement of such internal molecular torsion forces has been attempted.

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 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 workstations, 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 gain experience in the quality of fit that might be expected.

In a recent study (Welberry, 2000), we attempted to apply the method to a complex disordered system in which each molecular site was occupied by one of four basically different molecular orientations. In view of this degree of occupational complexity, it was assumed for simplicity that for the accompanying relaxational displacements the molecules behaved as rigid bodies, even though they possessed some internal degrees of flexibility. Although the model obtained provided a reasonable description of the disordered system, it was clear that further information was present in the experimental data and this provided the impetus for us to attempt to incorporate the possibility of internal degrees of freedom into future models. In the present study, as a first step, we apply the method to a simpler system in which the molecule has internal degrees of freedom and still gives strong diffuse X-ray scattering patterns but has only one basic molecular orientation for each site. This is the organic molecular crystal benzil [(I), diphenylethanedione, $C_6H_5-CO-CO-C_6H_5$].



(I)

Here an unconventional numbering of the atoms was chosen for convenience and is referred to later in the text.

Crystals of benzil have been known for many years to exhibit strong diffuse scattering. For example, a Laue

Table 1
The z-matrix representation of the molecule used in the analysis.

The nine figures in bold lettering were obtained by fitting the prototype z-matrix to the observed structural coordinates. Figures in italics are related to these. The remaining angles were fixed constants. *a* and *b* are dummy atoms used to define the local axial system.

Atom	Distance from *l* (Å)	Angle with *l*–*m* (°)	Dihedral angle with *l*–*m*–*n* (°)	*l*	*m*	*n*	Dihedral angle refine flag
a	–	–	–	–	–	–	
b	**0.7615**	–	–	*a*	–	–	
C1	*0.7615*	90.0	–	*b*	*a*	–	
C2	*1.0769236*	45.0	180.0	*a*	*b*	C1	
O3	**1.215**	**115.8**	**54.4***	C1	*b*	*a*	1
O4	*1.215*	*115.8*	*54.4**	C2	*b*	*a*	1
C5	**1.449**	**123.6**	**177.9**	C1	C3	*b*	
C6	*1.449*	*123.6*	*177.9*	C2	C4	*b*	
C7	**2.410**	150.0	**172.0***	C5	C1	*b*	1
C8	*2.410*	60.0	180.0	C7	C5	C1	
C9	*2.410*	150.0	*172.0**	C6	C2	*b*	1
C10	*2.410*	60.0	180.0	C9	C6	C2	
C11	*1.391*	30.0	180.0	C9	C10	C6	
C12	*1.391*	30.0	180.0	C6	C9	C10	
C13	*1.391*	30.0	180.0	C6	C10	C9	
C14	*1.391*	30.0	180.0	C7	C8	C5	
C15	*1.391*	30.0	180.0	C5	C7	C8	
C16	*1.391*	30.0	180.0	C5	C8	C7	

(stationary-crystal) X-ray photograph showing the scattering appears on the cover of the book by Kathleen Lonsdale (1948). Despite the presence of this diffuse scattering, conventional crystal structure determination has led to a respectable solution for the average structure (More *et al.*, 1987) and only the presence of rather high atomic displacement parameters indicates any sign of anomalous behaviour. They reported the room-temperature structure to be $P3_121$ with $Z = 3$ and $a = 8.402$, $c = 13.655$ Å. Isotropic atomic displacement factors reported ranged from $U_{eq} = 0.086$ to 0.173 Å^2 corresponding to B values ($B = 8\pi^2\overline{u^2}$) of 6.8 to 13.6 Å^2. The molecular twofold axes coincide with the space-group twofold axes. However, though different aspects of the structure have been studied, a completely satisfactory description of the diffuse scattering has not been forthcoming. To date, the most detailed study of the problem was that of Moss *et al.* (1997). These authors measured the diffuse intensity at points on a lattice with half the dimensions of the reciprocal lattice and analysed these data by extending the traditional **TLS** theory of rigid-body displacements to include correlations between different rigid bodies. The C_7H_5 groups of the molecule were taken as the rigid bodies and the correlations between these groups, defined in terms of tensors, were subject to least-squares refinement against the diffuse data.

In contrast to that work, we make use of diffuse scattering data covering whole sections of reciprocal space on a much finer reciprocal grid. These data were obtained using the position-sensitive detector (PSD) Weissenberg system described by Osborn & Welberry (1990). In the present case, data for benzil were collected on reciprocal planes (*hk*0), (*hk*0.5), (*hk*1), (*hk*1.5), (*hk*2), (*hk*2.5), (*hk*3), (*hk*4). The PSD was used in a single setting, which was chosen to span the range of approximately 26.0–77.2° in 2θ, using Co $K\alpha$ radiation. 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 200 × 200 pixels. Because of the limitations of available computer time, only four of these sections were used in the data analysis. These sections of data, displayed as false colour images, are shown in Fig. 1.

In all calculations carried out in the present work, cell dimensions of $a = 8.409$, $c = 13.672$ Å, newly measured at 295 K, were used.

2. 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). We here give 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 will depend on the details of the local occupational ordering. (In the present study, it should be noted, it is assumed there is no occupational disorder.) Parameters, p_i, specifying these interactions are then adjusted iteratively until best agreement is obtained between the observed diffraction patterns and ones calculated from the model coordinates. 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

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

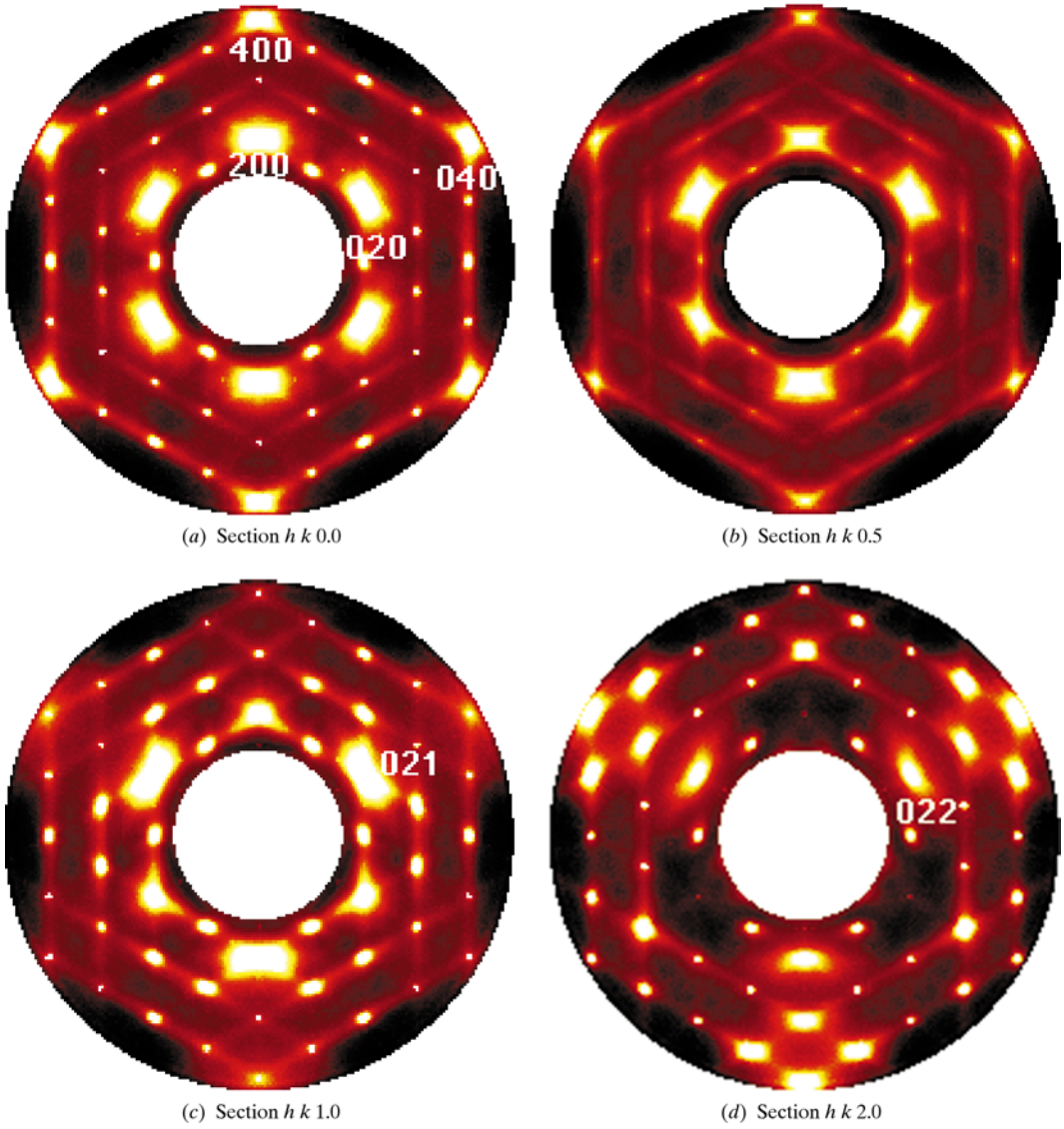
Here the summation is over all non-integral reciprocal points h, k, l corresponding to individual pixels in the m measured sections of data. f_m is a scale and b_m a background correction applied to section m. [Note that b_m and f_m are determined as described by Proffen & Welberry (1997) and are not 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 $\omega_{hklm} = 1/I_{obs}$. Numerical estimates of the differentials of χ^2 with respect to the model parameters are calculated by performing additional 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 iteration.

3. Molecular representation

In order to allow for the internal flexibility of the benzil molecule, a z-matrix representation was used. This method, 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 fixed with the exception of the dihedral angles corresponding to rotation about the central C1—C2 bond (φ_1) and rotations of the phenyl groups about the C1—C5 and C2—C6 bonds, respectively (φ_2 and φ_3). Parameters defining a prototype benzil *z*-matrix were used together with a quaternion representation of the orientation matrix to generate fractional crystal coordinates. (For a description of quaternions see *e.g.* Goldstein, 1980). Both the parameters of the *z*-matrix and the elements of the quaternion matrix were then adjusted iteratively until the fractional coordinates agreed with the corresponding observed crystal structure coordinates. The final *z*-matrix used in the analysis is given in Table 1. It should be noticed that two dummy atoms are included at the beginning. These are used to define the local orthogonal axial system. The origin of the local axial system is at the centre of the C1—C2 bond, with one axis extending along the bond and another along the C_2 axis of the molecule. Nine *z*-matrix parameters were used to obtain the fit to the observed coordinates. These are the bond lengths and angles printed in bold lettering. Other elements of the *z*-matrix that are related to these nine parameters are printed in italics. The remaining quantities in plain text are fixed constants that constrain the phenyl rings to be regular hexagons. After this initial fitting, only the dihedral angles indicated by an asterisk were allowed to vary during the remainder of the study.

It should be noted that the *z*-matrix in Table 1 defines the atoms starting from the centre of the molecule. The particular ordering of the atoms was chosen so that for much of the MC



(*a*) Section *h k* 0.0 (*b*) Section *h k* 0.5

(*c*) Section *h k* 1.0 (*d*) Section *h k* 2.0

Figure 1
The four sections of observed diffuse X-ray scattering data used in the analysis. The range of data recorded was approximately 26.0–77.2° in 2θ, using Co $K\alpha$ radiation.

work a cut-down version of the *z*-matrix could be used in which atoms C11–C16 were omitted.

Each molecule in the model crystal was represented in the computer by 11 variables. These consisted of three centre-of-mass coordinates, four quaternion parameters (one of these is redundant) and four dihedral angles (one of these is redundant since φ_1 is split into two parts, one either side of the local symmetry axis).

4. 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 workstations. 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 iterations 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 with that of the observed diffraction patterns, we have found in previous studies that a minimum system size of ~$32 \times 32 \times 32$ unit cells is required. In the present case, however, the *c* dimension is relatively long with three molecules in a spiral arrangement comprising the unit-cell repeat (see Fig. 2*a*) and it was found satisfactory to use a simulation system size of only $32 \times 32 \times 16$ unit cells. With 48 non-hydrogen atoms per unit cell, the model crystal thus comprises 786432 individual atoms.

To represent the intermolecular interactions, we use the idea of 'effective' interactions. Each molecule is in contact with 15 neighbouring molecules and if all atom–atom contacts were to be included the number of these would be prohibitively large to be contemplated for the MC simulations. With

Table 2
The intermolecular contact vectors used in the MC model.

Contacts		Cell translation			Symmetry equivalent	Distance	Vector
From atom	To atom	*x*	*y*	*z*	*s*	(Å)	type
O3	C14	0	1	0	1	3.42	11*
	C7	0	1	0	1	4.50	1
	C8	0	1	0	1	4.01	2
	C8	0	0	0	2	3.64	5
	C9	1	1	0	2	3.35	4
	O4	0	0	0	2	4.04	6
O4	C7	1	1	−1	3	3.35	4
	C10	1	0	0	1	4.01	2
	C10	0	0	−1	3	3.64	5
	O3	0	0	−1	3	4.04	6
	C9	1	0	0	1	4.50	1
	C11	1	0	0	1	3.42	11*
C7	C8	0	−1	0	2	3.99	7
	C9	0	0	0	2	4.23	8
	C9	0	0	0	3	4.46	10
	C10	0	0	0	2	3.73	9
	O4	1	0	0	2	3.35	4
	O3	0	−1	0	1	4.50	1
C8	C7	1	0	−1	3	3.99	7
	C9	0	0	0	2	3.73	9
	C10	0	−1	0	1	4.34	3
	C10	1	0	0	1	4.34	3
	O3	0	−1	0	1	4.01	2
	O3	0	0	−1	3	3.64	5
C9	C7	0	0	−1	2	4.46	10
	C7	0	0	−1	3	4.23	8
	C8	0	0	−1	3	3.73	9
	C10	−1	0	−1	3	3.99	7
	O3	0	1	−1	3	3.35	4
	O4	−1	0	0	1	4.50	1
C10	C7	0	0	−1	3	3.73	9
	C8	−1	0	0	1	4.34	3
	C8	0	1	0	1	4.34	3
	C9	0	1	0	2	3.99	7
	O4	−1	0	0	1	4.01	2
	O4	0	0	0	2	3.64	5
C14	O3	0	−1	0	1	3.42	11*
C11	O4	−1	0	0	1	3.42	11*

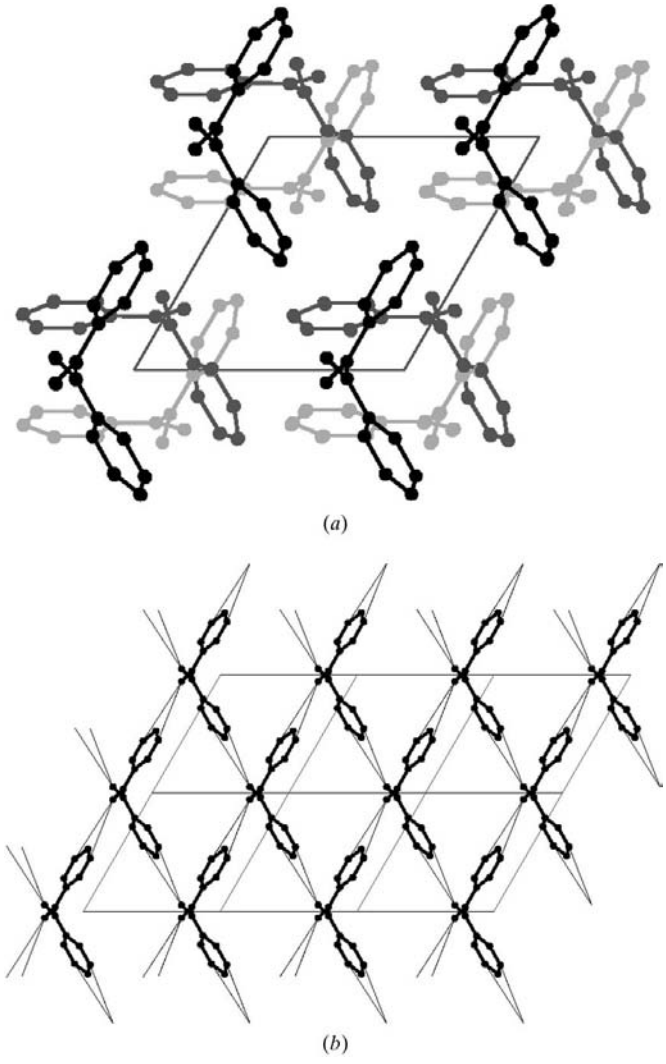


Figure 2
(*a*) Projection of the benzil structure down **c**. The three different shadings of grey indicate the three different symmetry-equivalent molecules occurring in the unit cell. (*b*) The network of contact vectors within a single layer of molecules formed by the interactions 1 and 2.

each molecule in its equilibrium position, all intermolecular atom–atom contact distances less than 4.5 Å were computed. A smaller number of contacts was obtained by pruning this list. Each intermolecular contact was represented by a small number of atom–atom contacts (2 to 3), which were sufficient to define the equilibrium separation and mutual orientation of neighbouring molecular fragments. Since each molecule was made up of rigid fragments with only three internal degrees of freedom, the atom–atom contacts involving only the atoms O3, O4 and C7, C8, C9 and C10 were considered sufficient for this purpose. These contact vectors, numbering 34 in total, are given in Table 2. The 34 vectors comprise ten symmetry-inequivalent vector types. At a later stage, four additional contact vectors (one additional vector type) involving atoms C11 and C14 were included. These are indicated with an asterisk in Table 2. 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 10 (subsequently 11) different vector types. The intermolecular contribution to the total energy of the system was, thus,

$$E_{\rm inter} = \sum_{\rm all\ contact\ vectors} F_i(d_i - d_{0i})^2. \quad (3)$$

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} may be found in Table 2.

The internal flexibility of the molecule was initially assumed to involve no energy penalty but in later stages of the analysis it became necessary to include such intramolecular forces. Two types of term were considered. First, simple torsional forces were included that depended on the deviations, $\Delta\varphi_i$, of the angles $\varphi_1, \varphi_2, \varphi_3$ from their equilibrium values. Secondly, terms to account for the possibility of correlated motion of the different intramolecular fragments were added. These intramolecular contributions to the energy were

$$E_{\rm intra} = \sum_{\rm all\ molecules} F_{12}(\Delta\varphi_1)^2 + F_{13}[(\Delta\varphi_2)^2 + (\Delta\varphi_3)^2] + F_{14}\Delta\varphi_2\Delta\varphi_3 + F_{15}(\Delta\varphi_1\Delta\varphi_2 + \Delta\varphi_1\Delta\varphi_3). \quad (4)$$

The total energy of the crystal is then expressed in the form

$$E_{\rm tot} = E_{\rm inter} + E_{\rm intra} \quad (5)$$

and this was used in the application of the normal Monte Carlo algorithm as follows. A site in the crystal was chosen at random and the set of 11 variables defining the conformation, orientation and position of the molecule was subjected to a small random increment. The contribution to the energy of all terms that are dependent on the altered variables is computed before and after the change is 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 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

$$P = \exp(-\Delta E/kT), \quad (6)$$

T is the temperature and k Boltzmann's constant. If $\eta < P$, the new configuration was accepted while, if $\eta > P$, it was rejected and the system returned to the original state. A Monte Carlo cycle is defined as the 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.

In the present work, we adopted the procedure that at each step only one of the eleven variables (chosen at random) was incremented. In this way, it was possible to monitor the acceptance/rejection ratio for each variable and the random step size for each was adjusted accordingly to maintain an acceptance/rejection ratio close to unity.

The number of MC cycles required 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 benzil, it was apparent that, to reproduce the strong diffuse peaks surrounding the Bragg peaks as well as the rather narrow diffuse lines that are apparent in the sections of data, rather long computation times would be required. Some initial work was carried out using 200 cycles of iteration but final cycles utilized 400 cycles. With this number of cycles, a single MC calculation used ~80 min of CPU time on a 400 MHz Pentium II processor.

5. Calculation of the 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), picked at random, from the simulated real-space array. This has the effect of removing the high-frequency noise, which 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 benzil, the diffraction patterns shown in Fig. 1 display diffuse lines 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' rather larger (in the a and b dimensions) than has been necessary in some previous studies. After some preliminary testing, a lot size of $10 \times 10 \times 2$ unit cells was chosen and the number of lots sampled in a given calculation was 200. Each of these diffraction pattern calculations took about 70 min of CPU time on a 400 MHz Pentium II processor. Even at this level of calculation, the patterns were rather noisy and further improvement in the signal to noise was obtained by averaging the patterns using their inherent $3m$ symmetry.

6. Progress of the refinement

At the start of the refinement, the system energy was considered to be dependent on only the first ten force constants, F_i, acting along the corresponding vectors defined in

Table 2. Values for all F_i were initially chosen to be equal with a value of 25. Differentials of χ^2 with respect to each of these parameters were computed using increments δF_i of ten. Refinement proceeded smoothly within a few cycles to produce a reasonably good fit to the data. The agreement factor, $r = (\chi^2/\sum \omega I_{\rm obs}^2)^{1/2}$, was ~0.19. However, a number of areas of disagreement were evident and steps were taken to modify the model to try to take account of these.

First, it was apparent that the calculated intensities at high angles were relatively too strong (*e.g.* the diffuse peak around 400 relative to that around 200). We found a similar effect previously in a study of urea/alkane inclusion compounds (Welberry & Mayo, 1996), where the effect was taken account of by using an overall Debye–Waller factor to modify the standard atomic scattering factors. In that case, the need for this was attributed to the fact that the calculations had been made on a single layer of the structure. The problem seems to arise because the true intensity, in sections normal to **c**, should be obtained by integrating over the whole columns of superposed atoms in the *c* 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 same limitation is occurring, partly because of the small lot size (in the *c* direction) but also because of the limited model crystal size. Further work is in progress to investigate this effect but for the present study we overcame the problem, as before, by introducing as a refineable parameter an overall Debye–Waller factor. An initial test indicated that employing a value of of 5.0 reduced the *R* value by about 2%.

A second region of discrepancy was that the strong and very broad diffuse peaks that appear in the *hk*1 and *hk*2 sections around the 021 and 022 Bragg positions were calculated to be considerably broader than the observation. It was considered

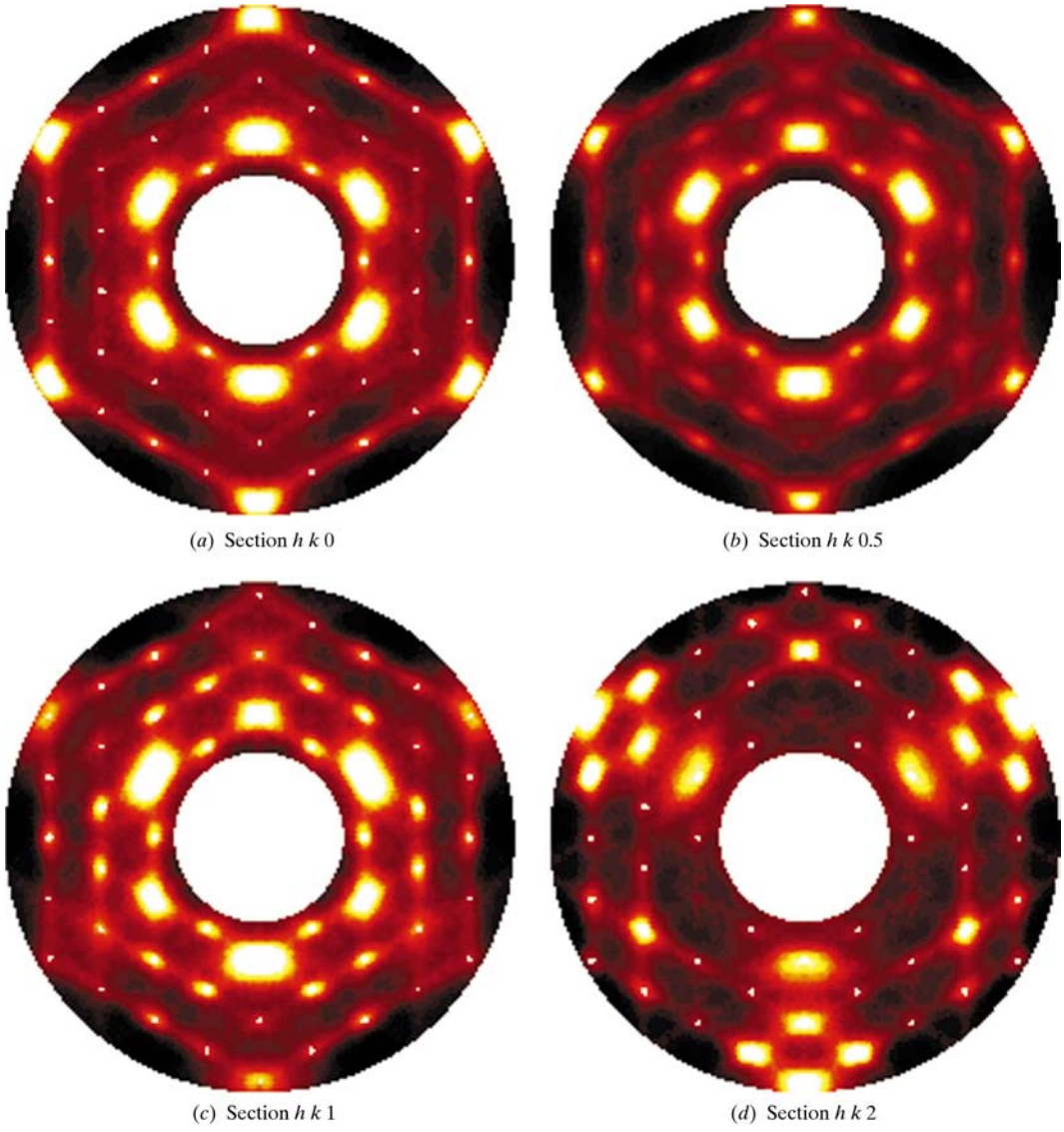


Figure 3
The four sections of diffuse X-ray scattering calculated from the final MC model.

Table 3
The final parameters used in the MC simulations.

Errors were estimated from the variation in the parameter values once a minimum in χ^2 had been reached. Model 1 used the four force constants F_{12}–F_{15} to specify the intramolecular forces while model 2 used a simplifed form in which $F_{13} = F_{12}$ and $F_{14} = F_{15} = 0$.

	Model 1		Model 2			Model 1		Model 2	
Parameter	Value	E.s.d.	Value	E.s.d.	Parameter	Value	E.s.d.	Value	E.s.d.
F_1	131.6	1.8	120.0	3.3	F_9	78.8	10.4	120.4	6.3
F_2	32.0	3.3	48.6	4.7	F_{10}	67.9	5.1	80.3	3.1
F_3	19.0	5.0	15.7	3.6	F_{11}	25.4	2.5	37.8	7.4
F_4	85.2	2.6	85.7	2.8	F_{12}	2.01	0.2	0.159	0.014
F_5	39.5	3.6	34.8	4.1	F_{13}	1.89	0.05	$= F_{12}$	
F_6	15.0	2.7	16.5	2.0	F_{14}	1.69	0.06	–	
F_7	116.3	4.1	145.6	8.8	F_{15}	1.86	0.07	–	
F_8	58.1	3.0	76.8	9.7	$B_{\rm overall}$ (Å^2)	5.17	0.13	6.17	0.09

that this effect could be caused by the fact that the model at this stage had no intramolecular constraints and so the molecule was too flexible. Consequently, the additional intramolecular energy terms given by (4) and involving four additional refineable parameters were incorporated into the model. It was found that including these intramolecular terms reduced the R value by about 3%.

It was apparent from the early cycles of refinement that the narrow diffuse lines that are seen in all the sections of diffuse data originate from a coupled motion of neigbouring molecular fragments along the [100], [010] and [110] directions. The coupling occurs by transmission between the O_3 or O_4 atoms of one molecule and the phenyl group of the neighbour. This corresponds to the contact vector types 1 and 2 in Table 2, and the value of F_1 in particular refined to a large value. However, it was clear from the geometry calculations that, while the F_1 and F_2 links were able to model this effect reasonably well, a direct link to the terminal carbon atom of the phenyl group might be more appropriate since it involved a much shorter distance (3.41 Å compared with 4.50 and 4.01 Å for vectors 1 and 2). Consequently, an additional force constant F_{11} was added along this vector. Although F_{11} refined, within about ten least-squares cycles, to a value comparable with the value for F_2, any tendency for it to take over the rôle previously played by F_1 and F_2 was not very apparent during the number of cycles of refinement that it was feasible to carry out. Moreover, it did not appear to result in significant improvement of the fit.

Final cycles of refinement were carried out with two different models. Model 1 had 16 refineable parameters: 11 intermolecular constants, 1 overall Debye–Waller factor and 4 intramolecular force constants as defined in (4). In model 2, only a single intramolecular force constant, F_{12}, was used, with F_{13} set equal to F_{12} and F_{14} and F_{15} set to zero. The final R value obtained for model 1 was 0.142 for 101 324 data points and, for model 2, the final R value was only a little higher at 0.145. Final values for the parameters are given in Table 3. Estimated errors included in Table 2 were obtained by noting the variation of the parameter values over ten cycles of iteration once a minimum in χ^2 had been reached.

7. Results and discussion

Calculated diffraction patterns obtained from the final MC parameters are shown in Fig. 3. The agreement with the observed data (Fig. 1) is generally very good, as would be expected with an overall R value of ~0.14. The agreement obtained in this study is considerably better than that obtained in any of our previous studies in which the automatic refinement algorithm was used. We attribute this to a number of factors. First, the disorder in benzil was describable purely in terms of thermal disorder without the need to consider occupational disorder and the numerous consequent cross-correlation terms. Secondly, the force model used, though still involving the idea of 'effective' interactions between molecules, used more individual atom–atom interactions than in previous studies, thereby providing a more effective model of the real intermolecular forces. Finally, the calculated diffraction patterns that were used in the least-squares process were of a better quality than those used in previous studies. This was partly due to the longer computation times that it was possible to use but also because of the symmetry averaging that was employed for the first time in the present study.

The analysis has shown quite clearly that the diffuse lines that feature so prominently in the observed diffraction patterns are due to strong longitudinal displacement correlations transmitted from molecule to molecule *via* the network of contacts of types 1, 2 (and 11) depicted in Fig. 2(*b*). Note here that by *longitudinal correlation* we mean that the direction of displacement and the direction of correlation are the same. It seems likely that the molecular basis for this is hydrogen bonding between the O_3 or O_4 atoms on one molecule and the *para* H atom (*i.e.* attached to C_{11} or C_{14}) of the phenyl ring of a neighbouring molecule.

The values of $B = 8\pi^2\overline{u^2}$ found in the final MC simulations ranged from 1.0 to 2.3 Å^2. In view of the fact that the refined models used an additional overall Debye–Waller factor, it is

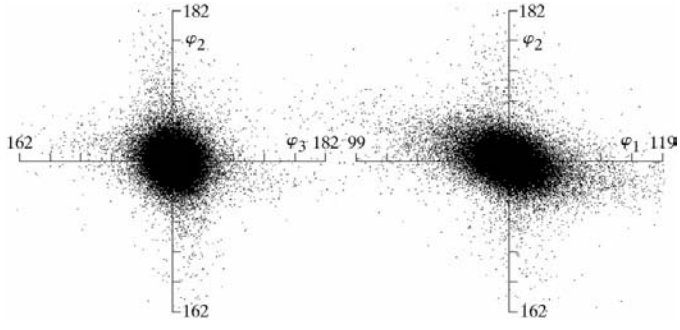


Figure 4
Plots of the distribution of the angles φ_1, φ_2, φ_3 in the final simulation of model 2. A non-circular distribution indicates correlation between the angle values imposed by the external intermolecular forces.

not unexpected that these values are considerably lower than those observed from the Bragg analysis. When combined with the refined $B_{\rm overall}$ (see Table 3), values comparable to those in the Bragg experiment are obtained.

The inclusion of the force constants F_{12}–F_{15} in model 1 resulted in an improvement in the agreement factor R of about 3%. In fact, if the values of F_{12}–F_{15} in this model are set to zero, the R value increases by about 4% and readjustment of the other parameters is only able to recover about 1% of this change. It is clear from this that these intramolecular parameters are important and their effect on the diffuse scattering highly significant. We believe that this study is the first diffuse-scattering study in which measurement of such internal molecular torsion forces has been attempted. At the present time, we are somewhat unsure of how much credence can be attached to the actual values derived. In particular, we were not convinced that the cross-correlation terms, F_{14}, F_{15}, could be as prominent as those determined for model 1. Consequently, we performed the refinement using model 2, in which F_{14} and F_{15} were set to zero. To simplify the model even further, F_{13} was also set equal to F_{12}. This was justified on the grounds that similar force constants in molecular-mechanics programs [such as *TRIPOS* (Clark *et al.*, 1989)] differ by only a few percent for single-bond rotations of the types present in benzil. Refinement of this model involving only 13 parameters yielded an R value almost as low as model 1.

The single value of F_{12} obtained for model 2 at first seems quite out of proportion to the four values, F_{12}–F_{15}, obtained for model 1 since it is more than a factor of ten lower. However, the values of F_{12}–F_{15} make up the elements of a force-constant tensor,

$$\mathbf{F} = \begin{pmatrix} F_{12} & F_{15} & F_{15} \\ F_{15} & F_{13} & F_{14} \\ F_{15} & F_{14} & F_{13} \end{pmatrix}, \tag{7}$$

and further investigation revealed that this tensor has two eigenvalues that are quite small (≤ 0.8) and a third that is very large (≥ 5.0). Moreover, the eigenvector corresponding to the smallest eigenvalue corresponds approximately to a rotation of the angle φ_1 with smaller counter-rotations of φ_2 and φ_3, and the second smallest corresponds approximately to a mode in which φ_2 and φ_3 rotate in opposite directions with φ_1 not involved at all. Consequently, it appears that the model 1 solution may not be very different from model 2 in its essentials. The third eigenvalue (the highest) in model 1 represents a direction in the φ_1, φ_2, φ_3 phase space (all three angles change in phase by approximately equal amounts) in which the molecule has become very stiff. Since this does not appear to affect the level of agreement very much, it is clear that the diffuse-scattering data must be relatively insensitive to this particular kind of stiffness and it should not be refined.

In Fig. 4, we show plots of the distribution of φ_2 *versus* φ_3 and φ_1 *versus* φ_2 for all of the molecules present in the final model 2 simulation. Since model 2 contains no direct forces correlating the angles, any departure from a circularly symmetric distribution is indicative of the fact that the external intermolecular forces act on the molecules to produce correlation between the angles. Fig. 4 shows that there certainly is correlation between φ_1 and φ_2 (and also φ_1 and φ_3) but very little between φ_2 and φ_3. This latter result is in agreement with the findings of Moss *et al.* (1997).

In order to check the reasonableness of the value of F_{12} derived for model 2, we performed *ab initio* molecular-orbital calculations to establish estimates of the lowest-energy modes of vibration of the free molecule. Calculations were carried out using *Gaussian98*. Fully optimized geometries and harmonic vibrational frequencies were obtained at the HF/6-31G(d) and B3-Lyp/6-31G(d) levels of theory. (For a description of the meaning of these theory levels, see Hehre *et al.*, 1986.) These calculations revealed that the lowest mode of vibration at about 25 $\rm cm^{-1}$ corresponded approximately to a twisting around the central bond, *i.e.* involving a rotation of φ_1. Somewhat higher at about 38 and 40 $\rm cm^{-1}$ were modes corresponding to out-of-phase and in-phase rotations of the two phenyl rings about their respective torsion angles φ_2 and φ_3.

To compare with these values, we performed a simple calculation of the oscillation frequency of a phenyl-ring libration by treating it as a simple harmonic torsional pendulum (see *e.g.* Halliday & Resnick, 1966, p. 358). The frequency of a torsional pendulum is

$$\nu = (1/2\pi)(\kappa/I)^{1/2}, \tag{8}$$

where κ is the torsional constant and I is the moment of inertia. For the phenyl ring, the moment of inertia is calculated from the contributions of four C atoms at a distance of 1.2 Å from the rotation axis and four H atoms at 2.08 Å. Using SI units, the moment of inertia is $I = 1.47 \times 10^{-45}$ kg m^2. We relate the torsional force constant κ to our force constant, F, by the relation

$$\kappa = 2F(180/\pi)k_BT. \tag{9}$$

Here, k_B is Boltzmann's constant and the temperature T is taken to be 300 K. Inserting I and κ into the frequency equation, we get

$$\nu = F^{1/2}2.856 \times 10^{12}\ \text{Hz}, \tag{10}$$

which, for $F = 0.159$, corresponds to a wave number of 38.0 $\rm cm^{-1}$, which is in excellent agreement with the *ab initio* derived value. Such good agreement may be somewhat fortuitous as there are numerous factors that might contribute to its inaccuracy. For example, if the MC simulations do not completely reach equilibrium in the finite number of cycles employed, the force constants might be forced to take on too large a value to compensate. It is also uncertain how the rather coarse parameter step size, needed to be employed to refine such parameters at all, will affect the final value. Similarly, the need to incorporate the overall Debye–Waller factor in the present work could also influence the actual value derived.

Despite these reservations, though, it is clear that, as the computational resources available increase and the accuracy of the whole calculation improves, obtaining such results should become much more accurate and reliable.

One of us (WIFD) gratefully acknowledges the award of a Visiting Fellowship from the Australian National University during the tenure of which part of this work was carried out. We are also endebted to Dr Tony Scott for carrying out, on our behalf, the *ab initio* molecular-orbital calculations mentioned in the discussion.

References

Butler, B. D. & Welberry, T. R. (1992). *J. Appl. Cryst.* **25**, 391–399.
Clark, M., Cramer, R. D. III & van Opdenhosch, N. (1989). *J. Comput. Chem.* **10**, 982–1012.
Goldstein, H. (1980). *Classical Mechanics*, 2nd ed. Reading, MA: Addison-Wesley.
Halliday, D. & Resnick, R. (1966). *Physics Parts I and II*, 2nd ed. New York: John Wiley.
Hehre, W. J., Radom, L., Schleyer, P. R. & Pople, J. A. (1986). *Ab Initio Molecular Orbital Theory*. New York: John Wiley.
Lonsdale, K. (1948). *Crystals and X-rays.* London: G. Bell.
More, M., Odou, G. & Lefebvre, J. (1987). *Acta Cryst.* B**43**, 398–405.
Moss, D. S., Wostrack, A. & Tickle, I. J. (1997). *Proc. Workshop on Inelastic Quasielastic Neutron Scattering in Biology*, edited by S. Cusack. New York: Adenine Press.
Osborn, J. C. & Welberry, T. R. (1990). *J. Appl. Cryst.* **23**, 476–484.
Proffen, T. & Welberry, T. R. (1997). *Acta Cryst.* A**53**, 202–216.
Welberry, T. R. (2000). *Acta Cryst.* A**56**, 348–358.
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.