Acta Cryst. (1998). A**54**, 661–674

Analysis of Single-Crystal Diffuse X-ray Scattering *via* Automatic Refinement of a Monte Carlo Model

T. R. Welberry,* Th. Proffen and M. Bown

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

(*Received 26 January 1998; accepted 18 March 1998*)

Abstract

In this paper, the feasibility of automatic least-squares refinement of a Monte Carlo model of a disordered structure by quantitative comparison of its calculated diffraction pattern with observed diffuse scattering data is demonstrated. At each stage of the refinement, the goodness-of fit parameter χ^2 is used as a quantitative measure of the agreement of the model with the observed data. Numerical estimates of the differentials of χ^2 with respect to the model parameters are calculated and these are used to provide automatic updating of the parameters *via* a least-squares algorithm. The method has been applied to the material $Fe_3(CO)_{12}$ which shows strong diffuse scattering caused by the fact that the triangular Fe_3 group can occupy a given molecular site, in either of two different orientations. A model has been refined involving 16 independent parameters that describe both the correlation of neighbouring occupancy variables and local orientational and centre-of-mass translational relaxation displacements. Although the method still presents a formidable computational task even for the fastest of modern computers, the work described represents a significant first step towards the ultimate goal of allowing routine analysis of diffuse scattering data.

1. Introduction

Although diffuse scattering from disordered crystals has been known since the earliest days of X-ray diffraction, its interpretation and analysis is still far from being a routine process and quantitative studies are relatively rare. The most extensive and quantitative methods have been developed in the field of alloys and simple oxides where the number of parameters needed to define a disordered structure is small and the diffracted patterns can be attributed unambiguously to basic interatomic interactions (see *e.g.* Epperson *et al.*, 1994; Borie & Sparks, 1971; Tibballs, 1975; Hayakawa & Cohen, 1975). For structures containing more than a few atoms, however, this type of detailed analysis is simply not feasible, and alternative approaches must be adopted (see *e.g.* the reviews by Welberry & Butler, 1995; Frey, 1995, 1997). For large (macromolecular) structures, interpreting diffuse scattering, when the structure contains so many parameters, would appear to be a quite intractible problem; but even here useful information can be extracted (see *e.g.* Clarage *et al.*, 1992; Doucet *et al.*, 1992; Kolatkar *et al.*, 1994, Clarage & Phillips, 1997). The present work forms part of a long-term study to develop methods that will allow routine analysis of diffuse scattering for a diverse range of materials.

In numerous recent papers, we have reported investigations of disordered crystals in which we have used Monte Carlo (MC) simulation of a computer model to aid in the interpretation of observed diffuse X-ray scattering patterns (Welberry, Butler, Thompson & Withers, 1993; Welberry, Butler & Heerdegen, 1993; Welberry & Glazer, 1994; Welberry & Mayo, 1996; Welberry & Christy, 1997). First a model is set up in the computer in terms of sets of random variables representing the atomic occupancies and positions or molecular orientations. A relatively small number of energy parameters describes the way in which these atoms or molecules interact. Monte Carlo simulation is then carried out for a time sufficient to allow the system to approach equilibrium. The final atomic coordinates of this model crystal realisation are then used to calculate diffraction patterns, which may be compared with the observed X-ray patterns. After assessment of the points of agreement and disagreement from this comparison, the model parameters are adjusted and the whole process is repeated iteratively until a satisfactory agreement between observed and calculated patterns is obtained.

Although convincing results have been obtained by this method for a variety of quite different systems, the crucial step of comparing the observed and calculated patterns has, to date, been performed visually and adjustment of the system parameters has relied heavily on an accumulation of experience, gained over a number of years. In the present paper, we describe our first attempts to perform this iterative MC methodology solely by computer, using quantitative rather than visual comparison of observed and calculated diffraction patterns and automatic updating of model parameters, using a least-squares algorithm. It should be stressed at

ISSN 0108-7673

Table 1. *Cell data and average structure – coordinates used in the present paper (obtained from Cotton & Troup, 1974)*

	Fractional coordinates					
	Orientation (+)			Orientation (−)		
	x	*y*	*z*	*x*	*y*	*z*
Fe(1)	−0.0544	−0.0204	−0.1684	0.0544	0.0204	0.1684
Fe(2)	0.1735	−0.0162	0.0716	−0.1735	0.0162	−0.0716
Fe(3)	−0.1123	0.0476	0.1100	0.1123	−0.0476	−0.1100

the outset that this represents a formidable computational task. At each stage of iteration, complete MC simulations are carried out for the current set of model parameters, diffraction patterns calculated and the goodness-of-fit parameter, χ^2, obtained as a quantitative measure of the agreement with the observed data. In addition, complete MC simulations, together with accompanying calculations of their diffraction patterns, are carried out for sets of parameters in which each of the parameters has, in turn, been changed by a small amount. From these, numerical estimates of the differentials of χ^2 with respect to each of the system variables are obtained, and are used to form the least-squares matrix, **A** (see later). Although computer speeds have increased rapidly over the last few years, it is still not feasible to perform such a calculation using extensive three-dimensional diffuse scattering data or using models that represent the real-space structure in as much detail as would ideally be desired. In order to make any progress, a number of approximations have had to be made and these are discussed in the sections that follow.

The system chosen for this study was dodecacarbonyltriiron, $Fe_3(CO)_{12}$. The molecule consists of a triangle of Fe atoms, surrounded by an approximately icosahedral cage of coordinating carbonyl (CO) groups (see Fig. 1). Crystal structure determination (Cotton & Troup, 1974; Braga *et al.*, 1994) revealed that the molecular sites are disordered but can be well modelled by assuming two different orientations of the same molecular geometry, related by a centre of symmetry. Cell data and the average structure coordinates used in the present work are given in Table 1. Although not previously reported, the presence of strong diffuse scattering was established in initial experiments and the rather simple form of these diffraction patterns seemed to indicate that a relatively simple model would be capable of describing them.

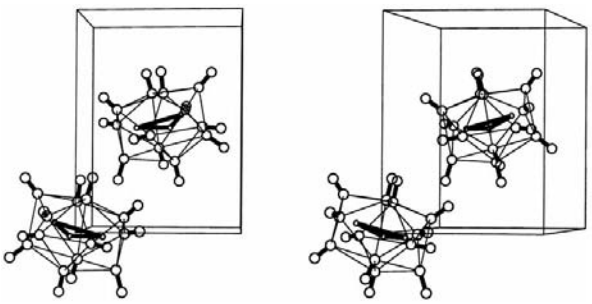

Fig. 1. Stereoscopic view down **c** of the structure of $Fe_3(CO)_{12}$. **b** is vertical and **a** is horizontal. Only one (−) of the two possible molecular orientations is shown for each molecular site. Heavy lines indicate Fe—Fe and C—O bonds. Thin lines indicate intramolecular C···C contacts less than 3.0 Å.

2. Experimental data

The compound dodecacarbonyltriiron was prepared by the published route (McFarlane & Wilkinson, 1966). Large single crystals of the compound suitable for diffuse X-ray scattering studies were grown by placing the compound in the thimble of a soxlet apparatus and extracting with *n*-pentane under an atmosphere of dry argon. Most crystals showed prominent {110}, {011} and {10$\bar{1}$} faces with {010} and {100} appearing occasionally. A number of crystals were selected and each was aligned about a different zone axis using optical reflection from the developed faces listed above. The three zone axes obtained were [100], [101] and [010]. Initial investigations were carried out using crystals in the [100] and [101] settings but subsequently further data were obtained using the [010] setting, when it became apparent that two sections were insufficient to obtain an unambiguous solution. For the crystals used, typical dimensions were ~0.4–0.5 mm for the edge length of the most prominent face. Complete sections of diffuse X-ray scattering data normal to each of these zone axes were recorded using the position-sensitive detector (PSD) system described by Osborn & Welberry (1990). This system uses a flat-cone Weissenberg geometry. The raw recorded data are obtained as a two-dimensional array of points in the Weissenberg space (2θ *versus* ω). Before use, these are re-binned into a square array of points (400 × 400 pixels) in the undistorted reciprocal-lattice plane. The experimental resolution due to beam divergence, crystal size and mosaic spread then corresponds approximately to a single pixel in this plane. Each data point is renormalized by dividing by the number of raw data points contibuting to it, so that no correction due to the Lorenz effect is required. In the present case, corrections due to X-ray polarization and specimen absorption were neglected. Graphite-monochromated Co $K\alpha$ ($\lambda = 1.790$ Å) radiation was used. Though the measured data are obtained in digital form, it is convenient for comparison purposes to display the data as false-colour or grey-scale images. We show in Fig. 2 for future reference the diffuse sections obtained normal to the three zone axes. In the case of the [010] zone axis, the section is the $\frac{1}{2}\mathbf{b}^*$ section since most of the diffuse scattering for this zone axis lies between the Bragg layers.

3. Monte Carlo model

Of crucial importance in this process is the need to have a model system of sufficient size that statistical variations are small and the resulting calculated patterns are relatively noise-free and of a comparable quality to the observed data. Our experience is that a minimum system size of $\sim 32 \times 32 \times 32$ unit cells is required and this was adopted throughout the study (see *e.g.* Welberry & Proffen, 1998). Binary random variables $s_{i,j,k,l}$ $(= \pm 1)$ were used to represent the two molecular orientations that are assumed to occur at each molecular site i, j, k, l. Here, i, j, k specify the unit cell and l one of the two sites within each cell. Hereafter, we will refer to the two different orientations as '+' and '−'. Although the initial work described here relates to the use of this set of binary occupancy variables, the model also incorporated further continuous random variables $X_{i,j,k,l}$, $Y_{i,j,k,l}$, $Z_{i,j,k,l}$, which allowed rigid-body centre-of-mass translations of the whole molecules, and $\varphi_{i,j,k,l}$ defining molecular librations.

3.1. *Ordering of the two molecular orientations*

The Monte Carlo energy used to specify local ordering of the molecular orientations was of the Ising-model form:

$$E_1 = \sum_n a_n s_{i,j,k,l} s_{i_n,j_n,k_n,l_n}. \quad (1)$$

Here the summation is over all n types of neighbour of a particular site. In this case, eight neighbouring vectors and their symmetry equivalents were included. Details of these vectors are given in Table 2. Unique values of the parameters a_n are required for these eight vectors which comprise four vectors between nearest neighbours along each type of body diagonal, three vectors between next-nearest neighbours in each of the axial directions and one additional vector between third-nearest neighbours in the [101] direction.

In all simulations, the starting configuration was a random distribution in which each of the $s_{i,j,k,l}$ variables was either $+1$ or -1 with equal probability. Consequently, the average, $\langle s_{i,j,k,l} \rangle$, was zero. The MC iteration was carried out by choosing a pair of sites at random and interchanging the variables in the two sites. In this way, the overall equal concentration of the two molecular orientations was maintained.

For use in the iterative fitting process described later, it would have been possible to use the interaction energies, a_n, as the system variables. However, it has been our usual practice in previous studies to use a feedback mechanism in our MC runs in order to obtain specified values of correlations along given interaction vectors (see Welberry, Butler & Heerdegen, 1993). Thus, we use correlation values $c_n = \langle s_{i,j,k,l} s_{i_n,j_n,k_n,l_n} \rangle$ as system variables. A given MC run therefore has as input the set of target correlations c_n and the corresponding set of energies a_n are determined during the course of the iteration. For all MC runs in the present work involving occupancies, 50 cycles of iteration were used. A cycle is defined as that number of individual MC steps required to visit each site once on average (*i.e.* in the present case $32 \times 32 \times 32 \times 2$).

Table 2. *Definition of the eight intermolecular vectors used in the analysis*

Contact type	Vectors to centre of neighbouring molecules from central molecule at (000)	
1	$\frac{1}{2}[\bar{1}\bar{1}\bar{1}]$	$\frac{1}{2}[\bar{1}1\bar{1}]$
2	$\frac{1}{2}[1\bar{1}1]$	$\frac{1}{2}[111]$
3	$[100]$	$[\bar{1}00]$
4	$\frac{1}{2}[1\bar{1}\bar{1}]$	$\frac{1}{2}[11\bar{1}]$
5	$\frac{1}{2}[\bar{1}\bar{1}1]$	$\frac{1}{2}[\bar{1}11]$
6	$[001]$	$[00\bar{1}]$
7	$[010]$	$[0\bar{1}0]$
8	$[101]$	$[\bar{1}0\bar{1}]$

3.2. *Centre-of-mass displacements and orientational relaxation*

Any local ordering of site occupancies such as that described above is invariably accompanied by local relaxation displacements that depend on the particular type of intermolecular contact, *i.e.* if neighbouring sites are occupied by molecules in orientations $(++)$, $(+-)$, $(-+)$ or $(--)$. In the case of a simple binary alloy, the phenomenon is generally known as the 'atomic size effect' (Warren *et al.*, 1951; see also Welberry, 1986; Welberry & Butler, 1994). Though such 'size-effect' displacements are generally small, they can contribute substantially to the diffuse scattering signal. In the present case, we are dealing not with single atoms but with a complex molecular shape. Even if the molecule is assumed to be rigid, local relaxation may involve not only a shift of the centre of mass but also a rotation away from its average orientation.

Such a complex means of relaxation would require many parameters to specify the problem completely and this is beyond our present computational resources. For the present purposes, we restricted the motion of each molecule to consist simply of a rigid-body translation of the centre of mass together with a single librational degree of freedom. In this way, it was possible to limit the number of parameters required to define the relaxation model in the MC program to manageable proportions.

For the centre-of-mass shift, an energy of the form

$$E_2 = \sum_{n,m} [d_{n,m} - d_{\text{ave}}(1 + \varepsilon_{n,m})]^2 \quad (2)$$

was used. Here, as before, the summation is over intermolecular vectors from a molecule at site n to its neighbour at site m. In this case, only the four shortest

vectors along which molecules are physically in contact were considered (vectors 1, 2, 4 and 5 in Table 2). $d_{n,m}$ is the instantaneous value of the distance between the centres of mass and $d_{\rm ave}$ is the corresponding equilibrium distance in the average structure. $\varepsilon_{n,m}$ is a 'size-effect' parameter which takes different values according to whether the two sites joined by the vector are $(++)$, $(-+)$, $(+-)$ or $(--)$. For each vector, $\varepsilon_{n,m}$ is specified by two variable parameters g_m and u_m,

$$\begin{aligned}\varepsilon_{n,m}(++) &= g_m + u_m \\ \varepsilon_{n,m}(+-) &= \varepsilon_{n,m}(-+) = -g_m \\ \varepsilon_{n,m}(--) &= g_m - u_m.\end{aligned} \tag{3}$$

Although four different combinations of neighbour exist, only two variables are needed since we assume that $(+-)$ and $(-+)$ are equivalent and the overall sum of shifts is zero. The form of (3) allows for both symmetric (g_m) and antisymmetric (u_m) distortions.

The analysis of the thermal ellipsoid tensors carried out by Braga *et al.* (1994) indicated that the Fe_3 triangles appeared to be undergoing librations about an axis parallel to the pseudo-twofold axis of the triangle passing through Fe(1). [In fact, since the Bragg experiment can only detect single-body motions, this model cannot be distinguished from an alternative in which atoms Fe(2) and Fe(3) are in fact moving together in the same direction, corresponding to a libration about an axis normal to that assumed.] Nevertheless, as a starting point, we assumed that a rotation about the pseudo-twofold axis of the triangle was a reasonable one. Such a model could be tested against alternative models.

If the random variable $\varphi_{i,j,k,l}$ is used to represent the angular variation about the pseudo-twofold axis, we can express the way in which the rotations of neighbouring molecules are linked by a size-effect-like expression analogous to that for the centre-of-mass shifts given by (2) above.

$$E_3 = \sum_{n,m} (\Delta\varphi_{n,m} - \xi_{n,m})^2. \tag{4}$$

Here, $\Delta\varphi_{n,m}$ is the difference in the value of the orientational variable at site n and that at a neighbouring site m. $\xi_{n,m}$ is a 'size-effect'-like parameter which takes different values according to whether the two sites joined by the vector are $(++)$, $(-+)$, $(+-)$ or $(--)$. As for the centre-of-mass 'size effect' [see (3)], we allow $\xi_{n,m}$ to have symmetric and antisymmetric components:

$$\begin{aligned}\xi_{n,m}(++) &= \gamma_m + \upsilon_m \\ \xi_{n,m}(+-) &= \xi_{n,m}(-+) = -\gamma_m \\ \xi_{n,m}(--) &= \gamma_m - \upsilon_m.\end{aligned} \tag{5}$$

This formulation thus requires two independent parameters, γ_m and υ_m, in addition to the centre-of-mass size-effect parameters, g_m and u_m, for each intermolecular vector. In the final model reported here, such orientational parameters were used on only the intermolecular vectors 1 and 2 given in Table 2. Although it was considered that further improvement of the fit of the model to the observed data could most likely be obtained by including similar parameters for vectors 4 and 5, this would have involved a further 50% increase in computation time for each cycle of refinement and this has not been pursued.

The relaxation of the molecular centres of mass and orientations, as described above, was applied subsequently to the occupational ordering and used a second stage of MC simulation. For this part of the simulation, a further 25 cycles of iteration were carried out in which the values of the continuous random variables $X_{i,j,k,l}$, $Y_{i,j,k,l}$, $Z_{i,j,k,l}$ and $\varphi_{i,j,k,l}$ were adjusted using the system energy $E = E_2 + E_3$ while the occupancy variables, $s_{i,j,k,l}$, remained constant.

4. Calculated patterns

The calculation time for obtaining diffraction patterns from the MC models is proportional both to the number of atoms in real space as well as to the number of points in reciprocal space at which calculations are performed. In the present case, it was not feasible to use either the full complement of atoms or to utilize the full set of recorded data, and some quite drastic approximations had to be made.

For the real-space part of the problem, to use all 27 atoms (3 Fe, 12 O and 12 C) in each molecule represents a nine times more expensive calculation than to use just the 3 Fe atoms, and throughout the present work only the Fe atoms were used. At first sight, this may appear a rather drastic approximation. It may, however, be justified by the following considerations. Firstly, since the atomic scattering factors for Fe are much larger than those for C and O, the scattering from the Fe atoms will tend to dominate. Secondly, since the lighter atoms are distributed throughout the unit cell rather than being concentrated in specific positions, their contribution would tend not to produce the rather simple and distinctive diffuse scattering patterns that are observed.

Two distinct situations were considered possible:

(i) The whole molecule (Fe's and ligands) behaved as a rigid molecule. In this case, leaving out the C's and O's would affect only the overall molecular structure-factor contribution to the scattering and not the detail of intensity variations within a reciprocal unit cell.

(ii) The behaviour of the CO ligands was not related in any simple way to the behaviour of the Fe's. If this were the case, the motion of any one ligand might similarly not be related in any simple way to any other. It then might be expected that with this multiplicity of inter- and intra- CO-to-CO vectors, in terms of both distance and mutual orientation, that these would contribute rather broad featureless diffuse intensity quite unrelated to the Fe scattering.

In either case, it seemed likely that using only the Fe's to calculate the diffraction patterns to be fitted to the observed data should provide at least a good first approximation to the problem and at least give some indication as to whether (i) or (ii) above pertained.

Complete reciprocal sections of data such as those displayed in Fig. 2 contain 400 × 400 pixels. This scale corresponds approximately to that at which an individual pixel matches the resolution of the observed X-ray data. However, this number of data far exceeds that which could feasibly be handled for the envisaged fitting process. One possible way of reducing the number of data would have been to re-bin the data to a lower resolution. However, we opted instead to retain the spatial resolution and to use only selected subsections of the data in the actual fitting process. The areas chosen in each of the observed sections were those regions where the diffuse scattering was strongest and most clearly delineated. The regions of data actually used are indicated in Fig. 2 by the white rectangles. A small region around each Bragg peak was also excluded from the calculations.

5. Least squares

The basic least-squares method seeks to minimize the sum of squares of the differences between a set of observed and calculated quantities. In our present case, the set of observed quantities consists of diffuse scattering intensities measured at individual pixels in the diffuse sections shown in Fig. 2. The corresponding calculated quantities are the suitably scaled values of the intensity obtained at corresponding points in reciprocal space from a MC simulation of a model system having system parameters p_i. That is, we minimize the goodness of fit χ^2:

$$\chi^2 = \sum_{h,k,l,m} \omega_{hklm}\{\Delta I\}^2, \tag{6}$$

where

$$\Delta I = I_{\rm obs} - (b_m + f_m I_{\rm calc}). \tag{7}$$

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 backgound correction applied to section m. [Note that b_m and f_m are determined as described in Proffen & Welberry (1997*a*) 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 $\omega_{hklm} = 1/I_{\rm obs}$.

Increments Δp_i to be applied to the model parameters are calculated using the following expression:

$$\Delta p_i = \sum_{l=1}^{npar} A_{il}^{-1} B_l. \tag{8}$$

The matrix **A** and the vector **B** involve the differentials of ΔI with respect to each of the variables p_i. **A** is a symmetric matrix and **B** is a vector,

$$\mathbf{A}_{ij} = \sum_{hklm} \omega_{hklm} \frac{\partial \Delta I}{\partial p_i} \frac{\partial \Delta I}{\partial p_j} \tag{9}$$

$$\mathbf{B}_{ij} = -\sum_{hklm} \omega_{hklm} \Delta I_{\rm trial} \frac{\partial \Delta I}{\partial p_i}. \tag{10}$$

It is also convenient to define a correlation matrix as a measure of how dependent the parameters used in the least-squares process are on each other. The correlation matrix **C** is defined as

$$\mathbf{C}_{ij} = \mathbf{A}_{ij}^{-1}/(\mathbf{A}_{ii}^{-1}\mathbf{A}_{jj}^{-1})^{1/2}. \tag{11}$$

Note that these correlations are not to be confused with the occupational correlations present in the disordered structure.

5.1. *Estimation of the differentials* $\partial \Delta I/\partial p_i$

We compute the differential $\partial \Delta I/\partial p_i$ as follows. If $\mathbf{p} = (p_0, p_1, p_2, p_3, \ldots, p_i, \ldots, p_n)$ is the current set of system parameters, the differential is estimated by performing two complete MC simulation and diffraction pattern calculations using parameter sets $\mathbf{p}^+ = (p_0, p_1, p_2, p_3, \ldots, p_i + \delta_i, \ldots, p_n)$ and $\mathbf{p}^- = (p_0, p_1, p_2, p_3, \ldots, p_i - \delta_i, \ldots, p_n)$, where δ_i is a suitably chosen small increment. The differential is then taken as

$$\frac{\partial \Delta I}{\partial p_i} = \sum_{hklm} \frac{(\Delta I_{p+} - \Delta I_{p-})}{2\delta_i}. \tag{12}$$

If the calculated diffraction patterns were infinitely accurate then it would be best to choose δ_i as small as possible. However, each calculation of ΔI is only an approximation to the true intensity corresponding to a given model parameter set. Both the MC simulation itself and the ensuing diffraction pattern calculation result in inaccuracies. The size of the model crystal is one limiting factor. For a chosen system size of $32 \times 32 \times 32 \times 2$ molecular sites, the normal statistical variations in the MC simulation lead to lattice averages such as the correlation coefficients c_n having an accuracy of $\sim 1/(32 \times 32 \times 32 \times 2)^{1/2} \approx 0.004$. Though using a larger system size would improve this accuracy, a corresponding increase in computer time (and memory) would be incurred.

Calculation of the diffraction pattern from the results of the MC simulation is performed using the program *DIFFUSE* (Butler & Welberry, 1992). The diffracted intensity from a large simulation array is computed as the average of the intensity from a large number of small subregions (or lots) of the model crystal, chosen at random. The purpose of performing the calculation in this way is to smooth out high-frequency variations in the pattern. The size of an individual lot needs to be

sufficient to include all significant nonzero correlations. Optimal results are obtained when every point in the array is sampled ~1–2 times on average. Further details of this sampling and the effect on the quality of the diffraction pattern can be found in Welberry & Proffen (1998). In the present case, we used throughout an ellipsoidal lot size extending over $5 \times 5 \times 5$ unit cells, *i.e.* each lot included $\sim \frac{4}{3}\pi \times 2.5^3 \simeq 65$ cells. In order to completely cover the whole array once, on average ~500 lots are required. Most of the work presented here was carried out with this degree of coverage. For final calculations, the number of lots was increased to 900, providing some improvement of quality at the expense of proportionately more computer time. In order to obtain still higher quality patterns, it is necessary also to increase the MC simulation array size or average over a number of separate runs.

Initial experiments involving only occupancy correlation variables were carried out in which the value of δ_i was chosen to be 0.025 (this was considered the lower limit of what would provide a significant difference between MC simulations with $\mathbf{p}^+$ and $\mathbf{p}^-$). However, it was found that this value of δ_i lead to rather poor estimates of the differentials and the matrix $\mathbf{A}$ was clearly badly determined, leading to poor convergence of the least-squares process. For later runs, a value of $\delta_i = 0.05$ was adopted and these gave much better convergence.

For the later computer runs in which centre-of-mass displacements were used, the value of δ_i was chosen to be 0.03, corresponding to a 3% change in the intermolecular distances. This value is comparable with the magnitude of size-effect shifts found in other studies. For these later stages of the study, the magnitude of the MC calculation for each cycle was such as to preclude a great deal of further experimentation and, since this value appeared to yield satisfactory results, it was retained for the rest of the study. For the runs in which the angular variables, $\varphi_{i,j,k,l}$, were used, δ_i for these variables was chosen to be 0.01 (in radians). This value gives rise to atomic shifts comparable with those resulting from the centre-of-mass displacements and for similar reasons this value also was retained for the rest of the study.

6. Results

6.1. *Refinement of occupancy correlations*

Initial runs were carried out using only two regions of data taken from the [101] and [100] zero-level sections. These regions are indicated as A1 and B1 in Fig. 2. These two regions contain the strongest and most definitive diffuse features and it seemed feasible that they alone would be sufficient to determine the basic occupational correlation structure. At this stage, correlation variables for vectors 1–7 were used (see Table 2). Although a good fit to this restricted set of data was obtained with this seven-parameter model, it was apparent from the correlation matrix $\mathbf{C}$ that some of the variables were highly correlated, *e.g.* the correlations along vectors 1 and 2. The reasons for this are not difficult to understand since these two vectors are superposed when viewed in projection down [101] and are parallel to each other when viewed down [100]. Similar considerations apply also to vectors 4 and 5. It was found that an equally good fit could be obtained when the correlations along vectors 1 and 2 were constrained to be equal and likewise for the correlations along vectors 4 and 5.

Table 3. *Refined values of the occupancy correlations c_n along the eight near-neighbour vectors defined in Table 2*

Vector	c_n	Vector	c_n
1	−0.120	5	−0.155
2	+0.243	6	+0.202
3	−0.097	7	−0.156
4	+0.114	8	+0.099

In order to try to break the correlation between these parameters, therefore, data from a third reciprocal section normal to [010] was introduced. In this setting, the strongest diffuse scattering appears between the Bragg layers and so data from the (h 0.5 l) section were used. Three additional small regions of data designated C1, C2 and C3 indicated in Fig. 2(*c*) were included in the refinement. It was clear from the outset that the solution obtained from the earlier refinement gave very poor agreement for this third section and although further refinement resulted in some improvement no satisfactory solution could be obtained using the seven-parameter model. It was concluded that additional correlation vectors were required. Inspection of the (h 0.5 l) in Fig. 2(*c*) shows that the diffuse scattering contains broad diffuse peaks that are elongated in a direction normal to [101], suggesting that a direct interaction along this vector is required. An eighth parameter along this vector was therefore added to the model (see Table 2).

With this new eight-parameter model, the least-squares refinement proceeded rapidly towards a minimum of χ^2 in only four cycles. The final overall R factor, $R = (\chi^2 / \sum w I_{\rm obs}^2)^{1/2}$, was 19.2% for 41 256 data points in the five regions used. Further cycles of refinement resulted in some fluctuations around this minimum, which we attribute to the noisy nature of the calculated matrix. The values of the derived correlation coefficients along each of the eight interaction vectors are listed in Table 3. Standard uncertainties for these values estimated from the diagonal elements of $\mathbf{A}^{-1}$ were all ~0.003 to 0.004. This figure is comparable with the rough estimate of the accuracy with which a MC simulation of $32 \times 32 \times 32 \times 2$ molecular sites can be expected to produce desired lattice averages (see §5 above). We consider that this represents an overestimation of the accuracy with which the correlation parameters can be said to have been determined. The

Table 4. *Refined values of the size-effect parameters $\varepsilon_{n,m}$ for centre-of-mass displacements and $\xi_{n,m}$ for orientations*

These are made up of symmetric and antisymmetric components g_m, u_m and γ_m, υ_m, respectively, as given in equations (3) and (5). $\varepsilon_{n,m}$ are unitless quantities and represent a percentage change of intermolecular distances. $\xi_{n,m}$ are in radians.

Vector 1				Vector 2			
Centre of mass		Orientation		Centre of mass		Orientation	
g_m	u_m	γ_m	υ_m	g_m	u_m	γ_m	υ_m
−0.014	−0.028	−0.124	+0.039	−0.023	−0.004	+0.128	+0.108
	(++)	(+−), (−+)	(−−)		(++)	(+−), (−+)	(−−)
$\varepsilon_{n,m}$	−0.042	+0.014	+0.014	$\varepsilon_{n,m}$	−0.027	+0.023	−0.019
$\xi_{n,m}$	−0.085	+0.124	−0.163	$\xi_{n,m}$	+0.236	−0.128	+0.020

correlation matrix **C** for the final cycle of least squares indicated that the correlation between parameters evident in the earlier runs had successfully been broken. Only two off-diagonal elements slightly exceeded 0.5 – those corresponding to parameters 1 and 8, and 4 and 7.

The computation time for four cycles of refinement on a single processor of the Silicon Graphics Power Challenge XL supercomputer was ∼22 h.

Diffraction patterns for the three complete sections calculated from a simulation using the final set of parameters (see Table 3) are shown in Figs. 3(*a*), (*b*), (*c*). It is seen that a generally good fit has been achieved where the data were fitted. Some features elsewhere, however, are not reproduced at all. The [101] section is seen to consist of a series of vertical wavy bands of intensity. If these are labelled from the centre as 1, 2, 3, 4 *etc.* in the b^* direction, whereas bands 1 and 2 are of comparable intensity and are reproduced well in the calculation, band 3 is rather weak in the observed data but is strongly present in the calculated pattern. This intensity variation may be considered as an asymmetry of intensity across a vertical plane passing through the (020) reciprocal point and is characteristic of the sort of effect that results from local relaxation in which the displacement of an atom or molecule is coupled to an occupancy variable. In order to proceed further with the analysis, the MC model was therefore extended to allow such relaxation effects, as decribed in §3.2. For this phase of the analysis, the occupancy correlations given in Table 3 were used throughout without change.

6.2. *Refinement of relaxation displacements*

In order to include the region of data poorly modelled by the occupancy refinement described above, the regions of data A1 and B1 were replaced in the analysis by regions A2 and B2 (see Fig. 2). Then, as a first stage in attempting to refine relaxational displacements, a model was tested that allowed only rigid-body translations of the molecular centre of mass. In this case, two size-effect parameters g_m and u_m were allowed for each of the four nearest-neighbour vectors 1, 2, 4 and 5, making a total of eight parameters in all. If only the [101] data were used for the fit, the refinement was able to achieve reasonably good agreement, with the relative intensity of the third wavy band much reduced in intensity. However, when all the data were included, the refinement was unable to achieve a satisfactory fit. It became apparent on inspection that any improvement in the [101] fit was offset by a worsening of the [100] fit. It was therefore deemed necessary to move to a model that allowed the molecules to rotate as well as translate.

This abortive attempt to use only centre-of-mass translations, however, seemed to indicate that the most important relaxation effects involved only the vectors 1 and 2 and resulted in atomic displacements predominantly in the crystallographic *b* direction. It was clear that displacements resulting from size-effect shifts along vectors 4 and 5 were necessarily not along **b**. In proceeding to a model involving both translation and rotation, therefore, only effects along vectors 1 and 2 were considered in order to keep the number of refinable parameters to a minimum. The new (and final) model therefore involved two centre-of-mass parameters, g_m and u_m, and two orientation parameters, γ_m and υ_m, for each of the two vectors 1 and 2, making eight parameters in all. As outlined in §3, the $\varphi_{i,j,k,l}$ were used to represent a rotation about the pseudo-twofold axis of the Fe_3 triangle. One further constraint was added to the model. During refinement, the average value of the rotational variables, $\langle\varphi_{i,j,k,l}\rangle$, tended to drift away from the observed position. This resulted in the average molecular site losing its centre of symmetry. In order to overcome this, the $\varphi_{i,j,k,l}$ variables were renormalized after each cycle to maintain a zero mean.

This model refined smoothly to a minimum in χ^2 after 14 cycles. This was rather slow convergence compared with the occupancy refinement but there may have been some contributing factors. The refinement was started using the parameter set obtained from the attempted centre-of-mass refinement. The final centre-of-mass shifts were quite different from these starting values and so a more appropriate starting point might have been the position of zero displacements. The final values for the γ_m and υ_m variables were rather larger than anticipated and the slow convergence might have been improved with the use of rather larger values of δ_i for these variables. The values of the parameters obtained

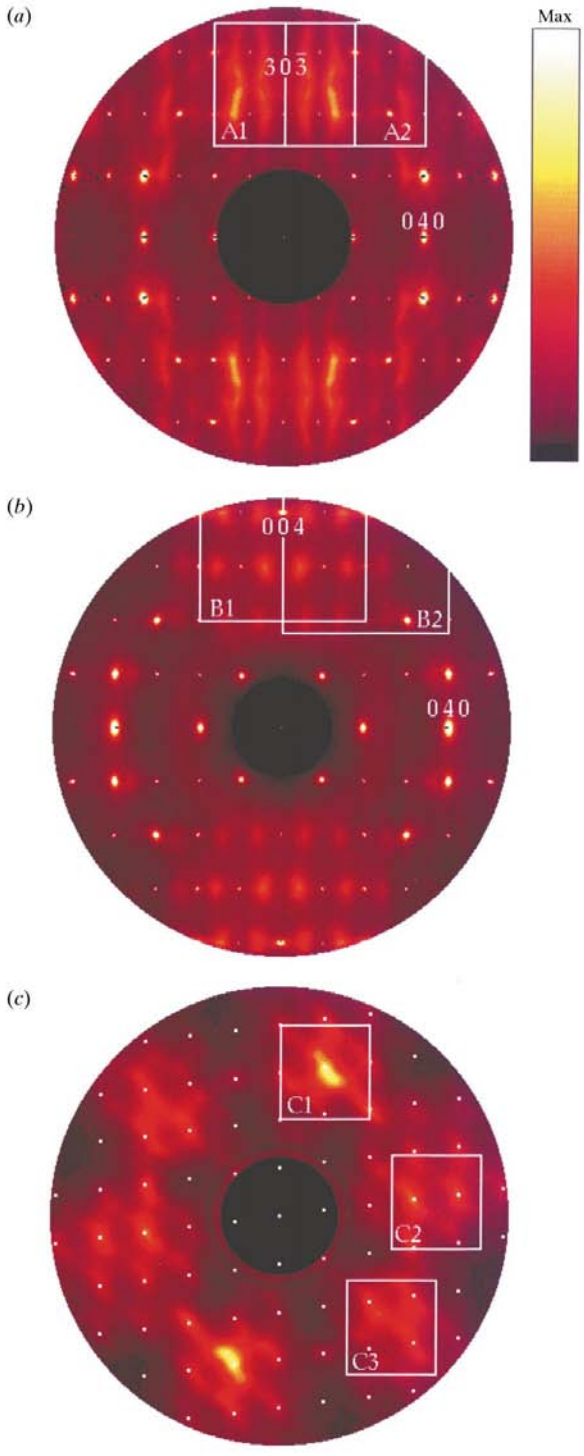

Fig. 2. Observed X-ray diffraction patterns of $Fe_3(CO)_{12}$. (*a*) Zero level section normal to [101]. (*b*) Zero level section normal to [100]. (*c*) 0.5$\mathbf{b}^*$ level section normal to [010]. The white rectangles indicate the regions of data actually used for fitting (see text for details). In (*c*), the white dots have been inserted to mark the positions of integral h and l.

are given in Table 4. The final R factor was 21.0% for 40 037 data points. The computation time for 14 cycles of refinement on a single processor of the Silicon Graphics Power Challenge XL supercomputer was ~56 h.

In Figs. 3(*d*), (*e*), (*f*), we show diffraction patterns for the three complete sections calculated from a MC simulation using the occupancy correlation variables given in Table 3 and the final displacement variables given in Table 4. The agreement with the observed patterns has clearly improved relative to the occupancy-only patterns shown in Figs. 3(*a*), (*b*), (*c*). In particular, the intensity of the wavy bands 3 and 5 are much reduced in intensity.

In order to be able to visualize exactly what this local relaxation means, it is convenient to make some statistical plots of the positions and orientations of the molecules averaged over the whole simulation. Fig. 4 shows such a probability plot for all the (+) molecules in the structure. The solid triangle represents the position of the average coordinates obtained from the crystal structure determination. This plot does not show any correlation information but simply shows the one-body distribution function for each atomic site. This may be compared with the plots of the anisotropic displacement parameters given by Braga *et al.* (1994). The refinement clearly indicates that the atomic site Fe(1) is reasonably localized and isotropic while sites Fe(2) and Fe(3) show a very large anisotropy. As a quantitative measure of this anisotropy, we can compute the variance $\langle\varphi^2_{i,j,k,l}\rangle$. The value obtained for our final solution is $\sim(15^\circ)^2$. This is a little larger than the value of $130 = (11.4^\circ)^2$ estimated by Braga *et al.* (1994) from a rigid-body analysis of their anisotropic displacement parameters.

To see the effect of the local relaxation that has resulted from the analysis, similar plots to Fig. 4 may be made by averaging over only those cells that contain (++), (+ −), (−) or (− −). Such plots are shown in Fig. 5. The pair of sites within one unit cell corresponds to the intersite vector 2. Similar plots for pairs of sites joined by vector 1 are shown in Fig. 6. Here it is clearly seen that the value of $\varphi_{i,j,k,l}$ for a particular molecule is strongly influenced by the occupancy of the neighbouring site. For some combinations of neighbouring occupancies, *e.g.* Figs. 5(*a*), 6(*d*), the effect is quite dramatic while, for other combinations, *e.g.* Figs. 5(*d*), 6(*a*), it is rather less so.

In order to interpret these figures in terms of the size-effect-like parameters, $\xi_{n,m}$, given in Table 4, it is important to realise that the sense of the $\varphi_{i,j,k,l}$ rotation for the molecular site at the origin appears opposite to that for its neighbouring molecule. This is because the two molecular sites are related by the space group 2_1 screw axis, which in projection appears to reverse the direction of the $\varphi_{i,j,k,l}$ rotation. For example, in Fig. 5(*a*) the molecules in both sites appear to have undergone a strong clockwise rotation. However, a clockwise rotation for the molecule at the origin corresponds to a positive

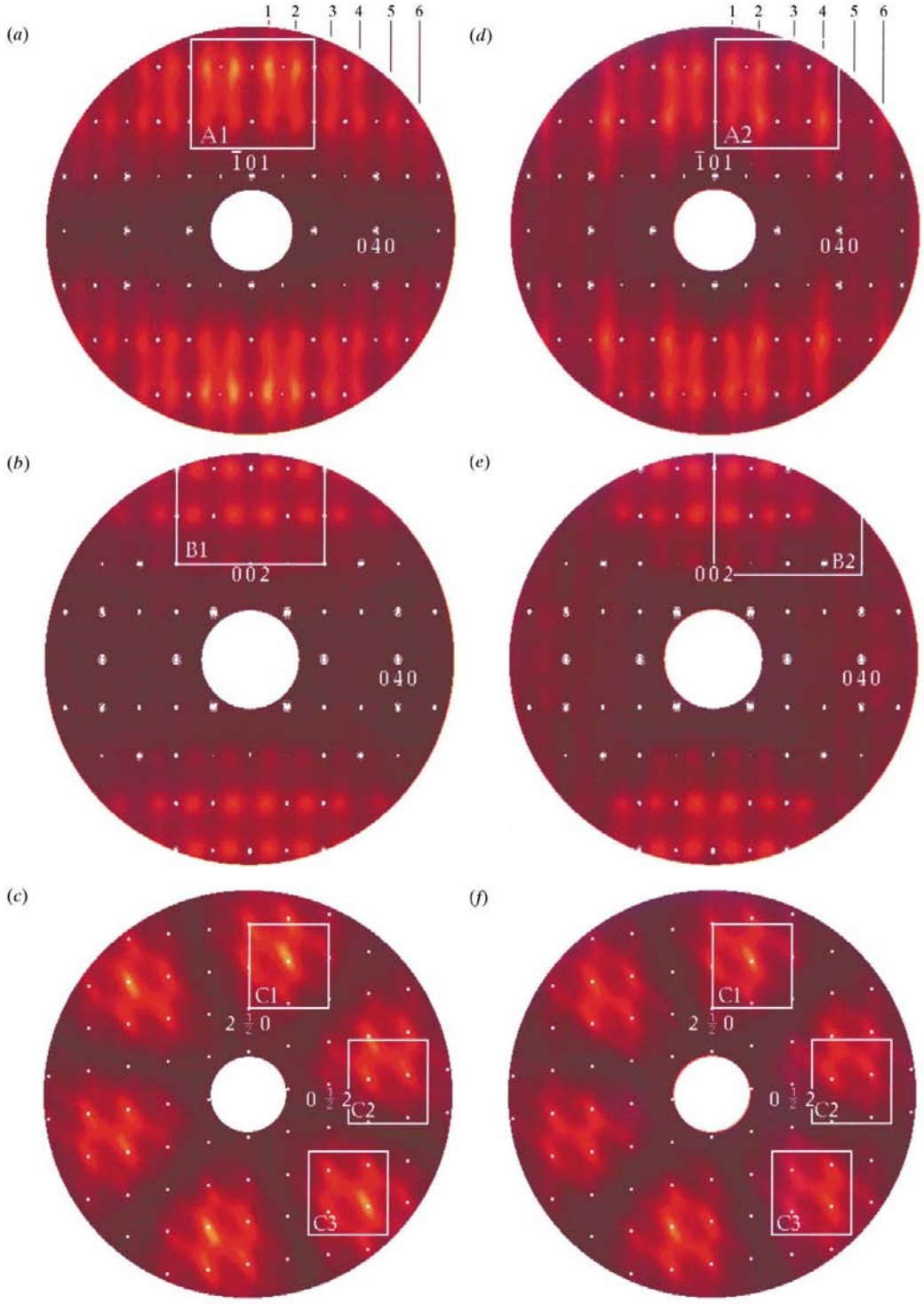

Fig. 3. Diffraction patterns of $Fe_3(CO)_{12}$ calculated from Monte Carlo simulations using the model parameters given in Tables 3 and 4. (*a*)–(*c*) correspond to an occupancy-only model; (*d*)–(*f*) correspond to the final model including relaxation. (*a*), (*d*) Zero-level section normal to [101]. (*b*), (*e*) Zero-level section normal to [100]. (*c*), (*f*) 0.5$\mathbf{b}^*$ level section normal to [010]. The numbers at the top of the figure identify the wavy bands of diffuse intensity referred to in the text. In (*a*), (*b*), (*c*) and (*d*), the positions of the Bragg peaks have been marked by superposing patterns calculated separately from the average structure using a large 'lot' size. In (*c*) and (*f*), the white dots have been inserted to mark the positions of integral *h* and *l*. The white rectangles indicate the regions actually used in the fitting.

$\varphi_{i,j,k,l}$ rotation while that in the neighbouring site correspond to a negative $\varphi_{i,j,k,l}$ rotation. Thus, the value of $\Delta\varphi_{n,m} = \varphi_{n,1} - \varphi_{n,2} = (+) - (-)$ is strongly positive as indicated by the value $\xi_{n,m}$, $(++) = 0.236$ in Table 4.

6.3. *Alternative model for relaxation displacements*

For comparison, we carried out a refinement using a model in which the $\varphi_{i,j,k,l}$ variables were used to describe a rotation about an axis in the plane of the Fe_3 triangle and normal to that described for the model above. Simple observation of the average molecular site as given by the Bragg experiment is not able to distinguish these two possibilities. This is clearly demonstrated by consideration of Fig. 5 since from this alone it is impossible to tell whether Fe(2) and Fe(3) are moving in phase or antiphase. The result of this attempted refinement was however quite conclusive. This alternative parametrization was unable to adjust to produce any significant improvement in χ^2. We therefore can conclude quite definitely that the relaxation takes place by rotation about the pseudo-twofold axis as described above.

6.4. *Calculations using whole rigid molecule*

For comparison with the final calculated diffraction patterns shown in Figs. 3(*d*), (*e*), (*f*), patterns were also computed using all the atoms (3 Fe, 12 C and 12 O atoms per molecule). These calculations simply used the centre-of-mass coordinates and the orientational variables $\varphi_{i,j,k,l}$ for the final model, together with the atomic coordinates for all of the atoms relative to a local axial system obtained from the room-temperature crystal-structure determination (Cotton & Troup, 1974). As anticipated, the computed patterns resembled the patterns of Fig. 3, but they were much noisier with higher (spatial) frequency modulations. These latter presumably coming from the larger interatomic distances within the rigid molecule. This tends to destroy the rather smooth appearance of the patterns compared with those in Fig. 3. In addition, the effect of the different molecular structure factor had caused the pattern of strong and weak vertical wavy bands in the [101] section to be quite different, no longer agreeing with the observations. This result would seem to confirm that option (ii) in §4 is the more likely scenario.

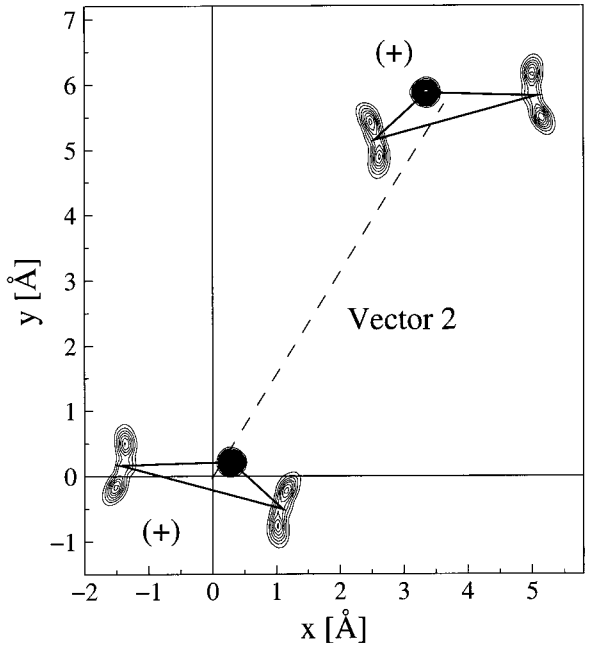

Fig. 4. Plot of the distribution of atomic coordinates for one orientation of the molecular triangle (+) obtained by averaging over all unit cells in the structure. The black triangle indicates the average positions determined by Cotton & Troup (1974).

7. Conclusions

In this paper, we have demonstrated the feasibility of automatic least-squares refinement of a Monte Carlo model of a disordered structure by quantitative comparison of its calculated diffraction pattern with observed diffuse scattering data. The method has been applied to the material $Fe_3(CO)_{12}$, which shows strong diffuse scattering caused by the fact that the triangular Fe_3 group can occupy a given molecular site, in either of two different orientations [designated (+) and (−)]. At each stage of the refinement, the goodness-of-fit parameter χ^2 is used as a quantitative measure of the agreement of the model with the observed data. Numerical estimates of the differentials of χ^2 with respect to the model parameters are calculated and these are used to provide automatic updating of the parameters *via* a least-squares algorithm. A model involving 16 independent parameters has been refined using a subset of the observed data consisting of in excess of ~40 000 data points. Of the 16 parameters used, 8 describe the correlation of the (+) and (−) occupancy variables along different neighbouring intersite vectors, 4 describe the dependence of centre-of-mass relaxation on the occupancy of neighbouring sites and the final 4 parameters describe, similarly, the dependence of the molecular orientational relaxation on the occupancy of neighbouring sites.

The particular course that was followed during this investigation was influenced by certain decisions made at a quite early stage of the study and the fact that limited (albeit generous) computational resources precluded the possibility of simultaneously following alternative routes. The decision to use the limited regions of data A1, B1 *etc*. indicated in Fig. 2 was influenced by the desire to maintain the full resolution of the observed data and by our prior experience in making qualitative comparisons between patterns,

where we tended to concentrate on the most intense features. In retrospect, it seems that it would perhaps have been more appropriate to use a data set of lower resolution but covering more of reciprocal space. It is our intention to pursue this possibility in future work. A similar choice to break the problem into two parts by first refining the occupancy correlations and then developing a relaxation model was again influenced by our prior experiences with qualitative modelling. This in turn dictated that we first used regions of data A1, B1

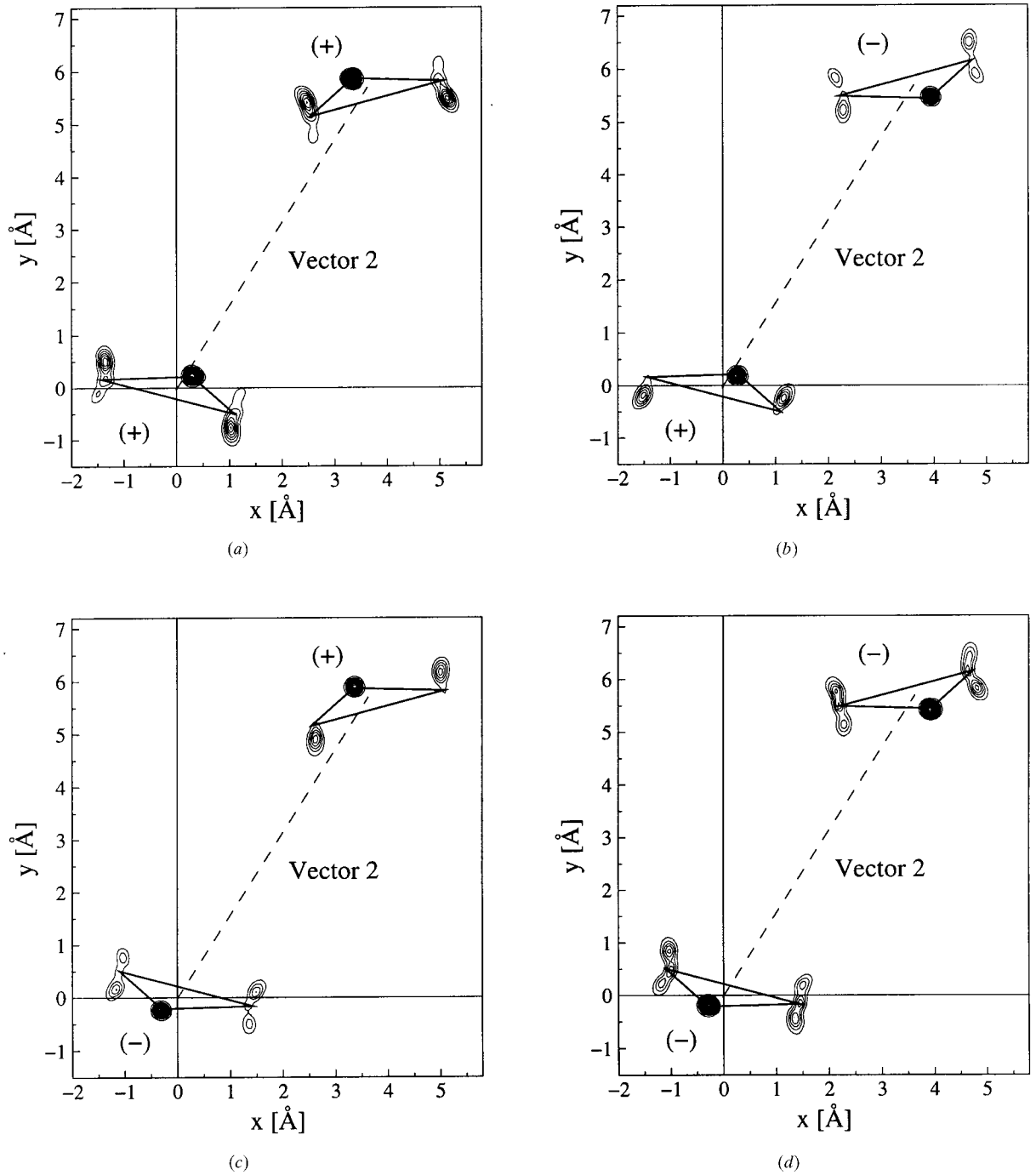

Fig. 5. Plots of the distribution of atomic coordinates for *pairs* of molecules separated by vector 2. These were obtained by averaging over all unit cells in the structure. (*a*) Both molecules are in orientation (+). (*b*) Molecular orientations (−) and (+). (*c*) Molecular orientations (+) and (−). (*d*) Both molecules are in orientation (−).

seemingly not affected by relaxation, and later introduced data A2, B2 when relaxational displacements were involved. [Note that this results in a slightly higher R factor in the latter case than that obtained for the more restricted data set used for the occupancy refinement alone.] Again, it is our intention in future work to pursue the possibility of refining simultaneously both occupational and relaxational parameters.

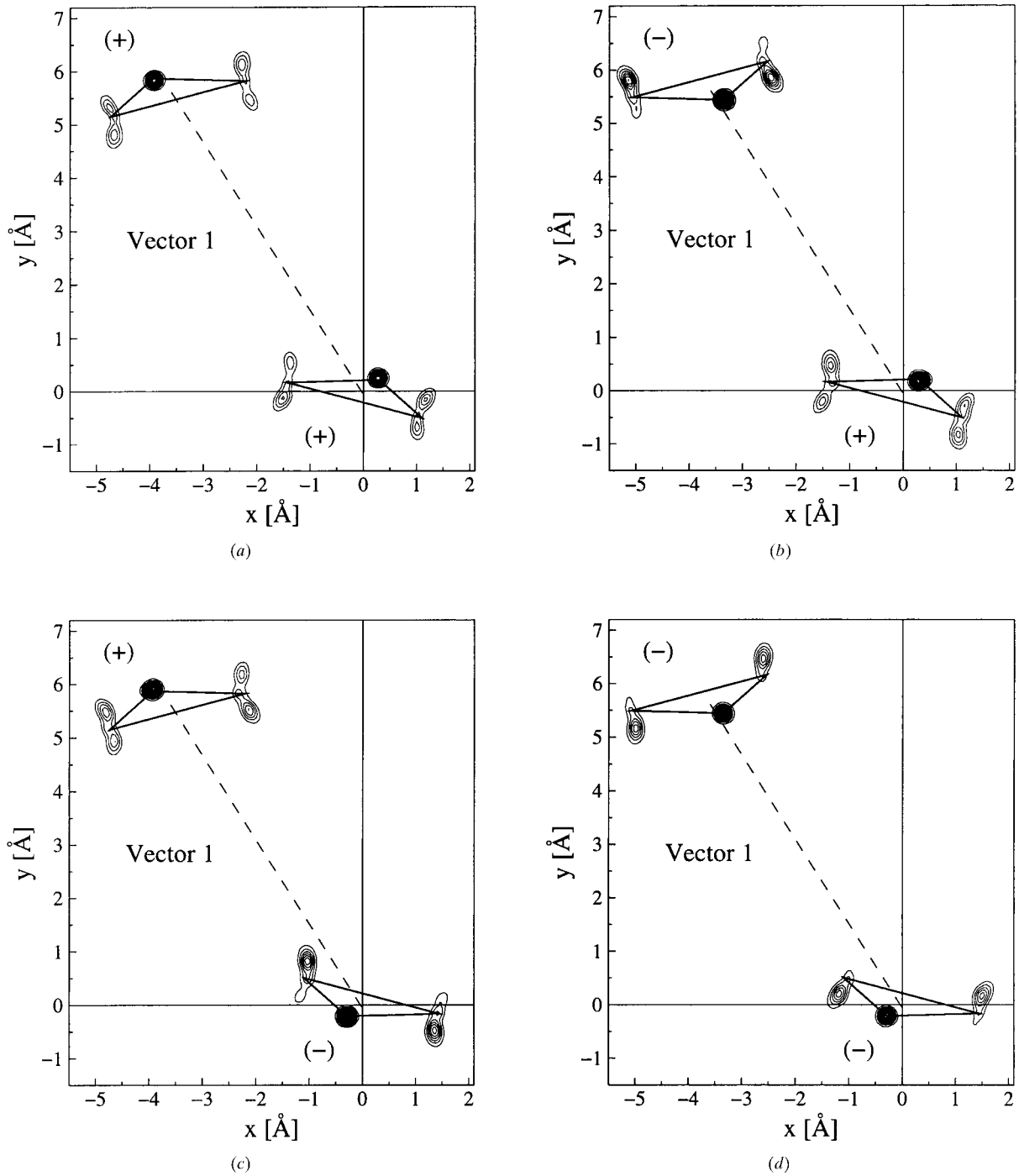

Fig. 6. Plots of the distribution of atomic coordinates for *pairs* of molecules separated by vector 1. These were obtained by averaging over all unit cells in the structure. (*a*) Both molecules are in orientation (+). (*b*) Molecular orientations (−) and (+). (*c*) Molecular orientations (+) and (−). (*d*) Both molecules are in orientation (−).

The methodology described here is a natural extension of our previous more qualitative approach to the interpretation of diffuse scattering patterns. In that method, we made visual comparisons of patterns calculated from a Monte Carlo model and made adjustments to the model parameters using more subjective criteria based on experience. Apart from the more obvious benefits of a quantitative treatment, one significant advantage of the new approach is that it allows more objective assessment of how the analysis is progressing. For example, it was readily apparent from the behaviour of the early refinements that the two sections of data chosen for the initial study were inadequate to determine the model correlation parameters uniquely, and that further data were required. Similarly, when additional data were added it was immediately apparent that the initial set of seven model parameters was unable to describe these additional data and a further parameter was required. Such judgements would have been difficult to make using our previous qualitative approach.

As a precursor to the work described in this paper, we have explored the possibility of using reverse Monte Carlo (RMC) simulation as a means of analysing diffuse scattering from single crystals (Welberry & Proffen, 1998; Proffen & Welberry, 1998). In RMC (McGreevy & Pusztai, 1988), the model used consists of a collection of atoms in a box in a similar way to MC, but the system variables are the atomic (or molecular) coordinates themselves. As the simulation progresses, these coordinates are adjusted in order to minimize χ^2. The main problem we encountered in that work was that, when the system size was large enough to provide a calculated diffraction pattern of sufficient quality to make a meaningful comparison with the observed pattern, the number of parameters was so large that the fitted model appeared quite spurious. Although attempts have been made to overcome this problem (Proffen & Welberry, 1997*b*), difficulties still remain. In contrast, for the present work using direct MC, a model involving only 16 parameters was used and even if this were to be extended considerably (should computational resources allow) the number of parameters would still remain much smaller than the number of observations. Furthermore, the set up of the initial model in terms of near-neighbour interactions guarantees a physically and chemically plausible description of the disordered structure. We believe therefore that the method as described here is a much more satisfactory approach to the modelling of diffuse scattering data.

Even for the relatively simple model described in this paper, a single refinement represents a formidable computational task and consumed ~80 h of CPU time on the fastest computer available to us (a single processor of the Silicon Graphics Power Challenge XL supercomputer). In developing the methodology and developing and testing a suitable model, many times this amount of computing was consumed. Nevertheless, we believe that the present work has effectively demonstrated the viability of the automatic MC approach and, as computational resources become ever faster, more and more complex problems should be amenable to solution. In this context, it should be pointed out that the method is ideally suited for parallel computer processing.

In this paper, we have concentrated on presenting a description of the method of analysis. The chemical implications of the results obtained for the compound $Fe_3(CO)_{12}$ are still the subject of further study and a more detailed discussion of these will be published elsewhere. For the present, it may be said that the results obtained here clearly demonstrate that the motion of the iron triangle within its icosahedron of carbonyl ligands can be described in terms of a libration of the iron triangle about the pseudo-twofold axis passing through the unique Fe atom, Fe(1) (Johnson & Bott, 1990; Braga *et al.*, 1990; Johnson *et al.*, 1992, 1993).

We gratefully acknowledge receipt of a grant of computer time from the Australian National University Supercomputer Facility.

References

Borie, B. & Sparks, C. J. (1971). *Acta Cryst.* A**27**, 198–201.
Braga, D., Anson, C. E., Bott, A., Johnson, B. F. G. & Marseglia, E. (1990). *J. Chem. Soc. Dalton Trans.* pp. 3517–3524.
Braga, D., Grepioni, F., Farrugia, L. J. & Johnson, B. F. G. (1994). *J. Chem. Soc. Dalton Trans.* pp. 2911–2918.
Butler, B. D. & Welberry, T. R. (1992). *J. Appl. Cryst.* **25**, 391–399.
Clarage, J. B., Clarage, M. S., Phillips, W. C., Sweet, R. M. & Caspar, D. L. D. (1992). *Proteins Struct. Funct. Genet.* **12**, 145–157.
Clarage, J. B. & Phillips, G. N. (1997). *Methods Enzymol.* **277**, 407–432.
Cotton, F. A. & Troup, J. M. (1974). *J. Am. Chem Soc.* **96**, 4155–4159.
Doucet, J., Benoit, J.-P., Faure, P. & Durand, D. (1992). *J. Phys. I*, **2**, 981–993.
Epperson, J. E., Anderson, J. P. & Chen, H. (1994). *Metall. Mater. Trans.* A**25**, 17–35.
Frey, F. (1995). *Acta Cryst.* B**51**, 592–603.
Frey, F. (1997). *Z. Kristallogr.* **212**, 257–282.
Hayakawa, M. & Cohen, J. B. (1975). *Acta Cryst.* A**31**, 635–645.
Johnson, B. F. G. & Bott, A. (1990). *J. Chem. Soc. Dalton Trans.* pp. 2437–2444.
Johnson, B. F. G., Parisini, E. & Roberts, Y. V. (1993). *Organometallics*, **12**, 233–235.
Johnson, B. F. G., Roberts, Y. V & Parisini, E. (1992). *J. Chem. Soc. Dalton Trans.* pp. 2573–2578.
Kolatkar, A. R., Clarage, J. B. & Phillips, G. N. (1994). *Acta Cryst.* D**50**, 210–218.
McFarlane, W. & Wilkinson, G. (1966). *Inorg. Synth.* **8**, 181–183.
McGreevy, R. L. & Pusztai, L. (1988). *Mol. Simul.* **1**, 359–367.

Osborn, J. C. & Welberry, T. R. (1990). *J. Appl. Cryst.* **23**, 476–484.
Proffen, Th. & Welberry, T. R. (1997*a*). *Acta Cryst.* A**53**, 202–216.
Proffen, Th. & Welberry, T. R. (1997*b*). *Z. Kristallogr.* **212**, 764–768.
Proffen, Th. & Welberry, T. R. (1998). *J. Appl. Cryst.* **31**, 318–326.
Tibballs, J. E. (1975). *J. Appl. Cryst.* **8**, 111–114.
Warren, B. E., Averbach, B. L. & Roberts, B. W. (1951). *J. Appl. Phys.* **22**, 1493–1496.
Welberry, T. R. (1986). *J. Appl. Cryst.* **19**, 382–389.
Welberry, T. R. & Butler, B. D. (1994). *J. Appl. Cryst.* **27**, 205–231.
Welberry, T. R. & Butler, B. D. (1995). *Chem. Rev.* **95**, 2369–2403.
Welberry, T. R., Butler, B. D. & Heerdegen, A. P. (1993). *Acta Chim. Hung. Models Chem.* **130**, 327–345.
Welberry, T. R., Butler, B. D., Thompson, J. G. & Withers, R. L. (1993). *J. Solid State Chem.* **106**, 461–475.
Welberry, T. R. & Christy, A. G. (1997). *Phys. Chem. Miner.* **24**, 24–38.
Welberry, T. R. & Glazer, A. M. (1994). *J. Appl. Cryst.* **27**, 733–741.
Welberry, T. R. & Mayo, S. C. (1996). *J. Appl. Cryst.* **29**, 353–364.
Welberry, T. R. & Proffen, Th. (1998). *J. Appl. Cryst.* **31**, 309–317.