research papers

Diffuse scattering resulting from macromolecular frustration

T. R. Welberry,[a]* A. P. Heerdegen,[a] D. C. Goldstone[b] and Ian A. Taylor[b]

[a]Research School of Chemistry, Australian National University, Canberra, ACT 0200, Australia, and [b]Division of Molecular Structure, National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, England

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

Distinctive diffuse scattering in the form of diffuse rings around Bragg positions has been observed in the diffraction patterns of a crystal of the N-terminal fragment of the Gag protein from Feline Foamy Virus. It is shown that these are caused by geometric frustration as molecules try to pack on the triangular b–c mesh of the space group $P6_122$. In order to explain the strong diffuse scattering it is necessary for the crystal to contain occupational disorder such that each unit cell contains one or other of two different molecular arrangements, **A** and **B**. The frustration arises because the nearest-neighbour packing prefers neighbouring cells to be **AB** or **BA**, which cannot be achieved on all three sides of a triangle simultaneously. To explain the observation that reciprocal sections $hk5n$, where n = integer, contain only Bragg peaks it is necessary that **A** and **B** are identical molecular arrangements differing only by a translation of $0.2c$. The implications of the disorder for solving the structure of the protein by conventional techniques as well as the possibility of using the diffuse scattering for this purpose are discussed.

1. Introduction

Studies of diffuse scattering have become relatively common in many areas of crystallography including organic molecular crystals (Chan *et al.*, 2010), inorganic solid-state materials (Paściak *et al.*, 2010), metals and alloys (Du *et al.*, 2010) and quasicrystals (Weidner *et al.*, 2004; Weber *et al.*, 2008). In many examples, using the Monte Carlo (MC) computer simulation of a model structure has become a powerful and well accepted technique for aiding the interpretation and analysis of the diffuse scattering patterns (Weber & Bürgi, 2002; Welberry, 2004). In the most favourable cases an MC model structure can give calculated diffraction patterns in quantitative agreement with the observed data (see *e.g.* Chan *et al.*, 2009).

In the field of macromolecular crystallography, however, studies of diffuse scattering are still relatively rare. Protein crystallography has relied almost entirely on the analysis of Bragg diffraction data alone. In fact, when examined carefully most protein crystals show significant (often substantial) diffuse scattering in addition to the usual Bragg diffraction. This diffuse scattering contains information about the disorder in the crystal and particularly the dynamical behaviour of the molecules that cannot be obtained from the Bragg diffraction data. In studies that have been reported to date two types of diffuse scattering in proteins have been distinguished (Glover *et al.*, 1991).

The first type (I) can be attributed to fluctuations in the atomic positions on a length scale which is shorter than the dimensions of the unit cell. This gives rise to very diffuse, so-

called *variational* scattering which has a cloudy appearance of scattered intensity over broad regions of reciprocal space. Pioneering work on this type of scattering was carried out by Caspar *et al.* (1988) who treated this scattering using a description involving the ordered unit-cell content (structure factor), a displacement factor and a truncation function which took into account the range over which correlated motions occur.

The second type (II) can be attributed to correlated motions that occur at length scales larger than the unit cell. In this case the diffuse scattering is more structured (often in the form of streaks) and appears close to Bragg peaks. In an early paper diffuse scattering of this kind (in orthorhombic lysozyme) was

(a)

(b)

Figure 1
(a) A single frame of data collected from a crystal of *N*-terminal fragment of the Gag protein from Feline Foamy Virus. (b) An enlargement of the rectangular region highlighted in (a) showing the diffuse circular features that occur in the vicinity of the Bragg peaks. A movie showing the complete sequence of raw data frames is included in the supplementary material.

interpreted by Doucet & Benoit (1987) in terms of correlated rigid-body motions of molecules along rows within the crystal.

Further studies have been carried out on a number of different systems including yeast initiator tRNA (Kolatkar *et al.*, 1994), tetragonal lysozyme (Pérez *et al.*, 1996), tropomyosin (Chacko & Phillips, 1992) and calmodulin (Wall *et al.*, 1997). Most recently methods have developed that use elastic network models to describe the motions in proteins (Riccardi *et al.*, 2009).

In this paper we report on some very distinctive diffuse scattering that has been observed in a routine Bragg intensity data collection from a crystal of the *N*-terminal fragment of the Gag protein from Feline Foamy Virus. A single frame of data is shown in Fig. 1(a) and a magnified region taken from this is shown in Fig. 1(b). The enlarged image shows that the form of the scattering is of *diffuse rings* that surround the Bragg positions. An initial assessment of the scattering is that it is quite distinct from either the type I or type II scattering described above.

At the present time an analysis of the Bragg reflection data has not yet yielded a crystal structure solution. It seems a distinct possibility that the presence of the disorder that is giving rise to the diffuse scattering may be a major contributing factor in preventing structure solution using the Bragg reflections. One of the aims of this paper therefore is to determine the origins of the diffuse scattering with a view to providing additional input to the structure solution strategy.

As a consequence, in trying to understand the origin of these distinctive effects, we do not at this stage have recourse to an average structure which can be used as a starting point for model building.

2. Experimental data collection

The crystal was grown from the *N*-terminal 154 residues of the Gag protein from the Feline Foamy Virus. This protein has a molecular weight of 19 kDa and forms a tight dimer in solution. X-ray diffraction data were collected on a Rigaku Micromax-007HF with an R-axis-IV detector using Cu $K\alpha$. 213 individual $0.5°$ oscillation frames were recorded, which represents a total crystal rotation, ω, of $106.5°$. Each frame comprised 3000×3000, 100×100 µm pixels. The Bragg reflection data obtained from these frames were indexed and scaled in the space group $P6_122$ with the dimensions $\mathbf{a} = \mathbf{b} = 73$ and $\mathbf{c} = 109$ Å (see Table 1 for data collection statistics). Cell content analysis establishes the presence of a single monomer in the asymmetric unit with a Matthews coefficient, V_m, of 2.22 Å^3 Da^{-1} and 44.6% solvent. The crystallographic $\mathbf{c}$ axis was found to be inclined to the rotation spindle axis by approximately $66°$.

Individual reciprocal lattice sections were reconstructed from the primary oscillation frames using the software *XCAVATE* (Estermann & Steurer, 1998). After the initial inspection to confirm that the raw diffuse data displayed sixfold symmetry this symmetry was used to populate regions of the patterns not recorded and to improve the overall signal to noise by averaging.
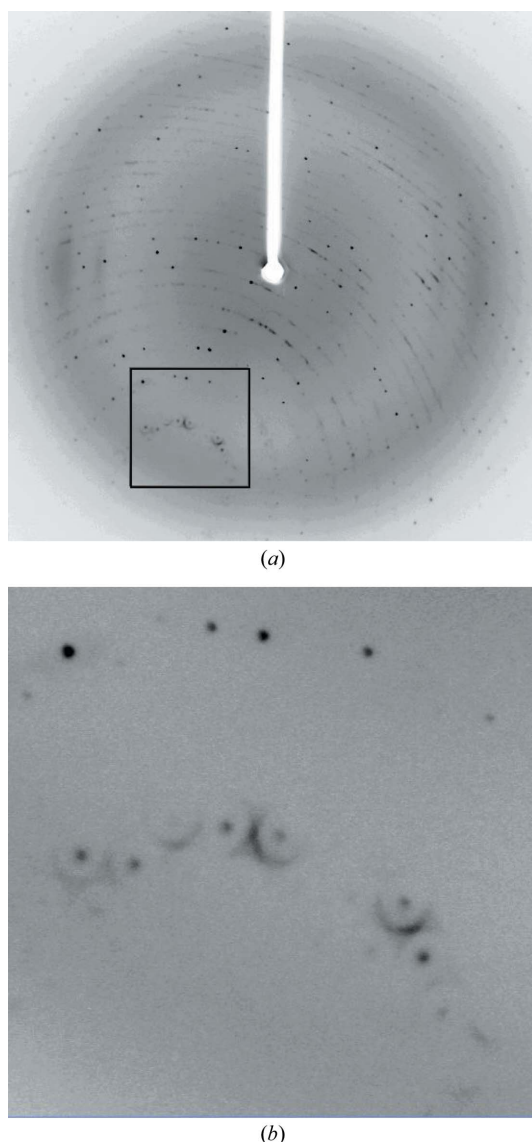
research papers

Table 1
Data collection statistics for FFV-Gag Ntd.

Space group	$P6_122$
Cell dimensions (Å)	$\mathbf{a} = 73$, $\mathbf{c} = 109$
Resolution (Å)	2.3 (2.38–2.3)
R_{merge}	8.3 (27.1)
$\langle I \rangle / \sigma_{(I)}$	27.9 (7.0)
Completeness (%)	99.6 (100)
Redundancy	11.0 (7.1)
M_r (Da)	19 130
Nmol (ASU)	1
V_m (Å^3 Da^{-1})	2.22
Solvent (%)	44.6

V_m is based upon the presence of one molecule in the asymmetric unit. M_r is calculated from the protein sequence including the N-terminal His-tag. Values in parentheses refer to the outermost shell of data.

Initial exploration of the three-dimensional data set indicated that the diffuse scattering features occurred in layers normal to **c**, *i.e. hkn*, where *n* is integral. Fig. 2 shows reconstructed sections for the first six reciprocal layers normal to **c**, *i.e. hk0, hk1, hk2, hk3, hk4, hk5*. Each of these layers has been identically scaled from the original data and the same back-

ground correction has been applied. Similar diffuse features were observed in the region of and outside the solvent ring but to display these would require substantially different treatment of background. Similar data were also extracted for sections *hk6* up to *hk15*.

The diffraction patterns shown in Fig. 2 are all of sections normal to **c** with *l* = integer. The diffuse scattering is more or less confined to these *l* = integer sections. This can be seen by taking a slice through the three-dimensional data perpendicular to these sections. An example of such a section is shown in Fig. 3.

2.1. Initial assessment of diffuse scattering.

It is clear from the plots shown in Fig. 2 that, to a good first approximation, sections *hk0* and *hk5* contain only Bragg peaks and no diffuse scattering; sections *hk1* and *hk4* contain fairly strong Bragg peaks and weak diffuse scattering but sections *hk2* and *hk3* contain strong diffuse scattering and only very weak Bragg peaks. This pattern is repeated in the sections *hk6* upwards. *hk10* and *hk15* contain only Bragg peaks; *hk6*, *hk9*, *hk11* and *hk14* contain fairly strong Bragg peaks and weak
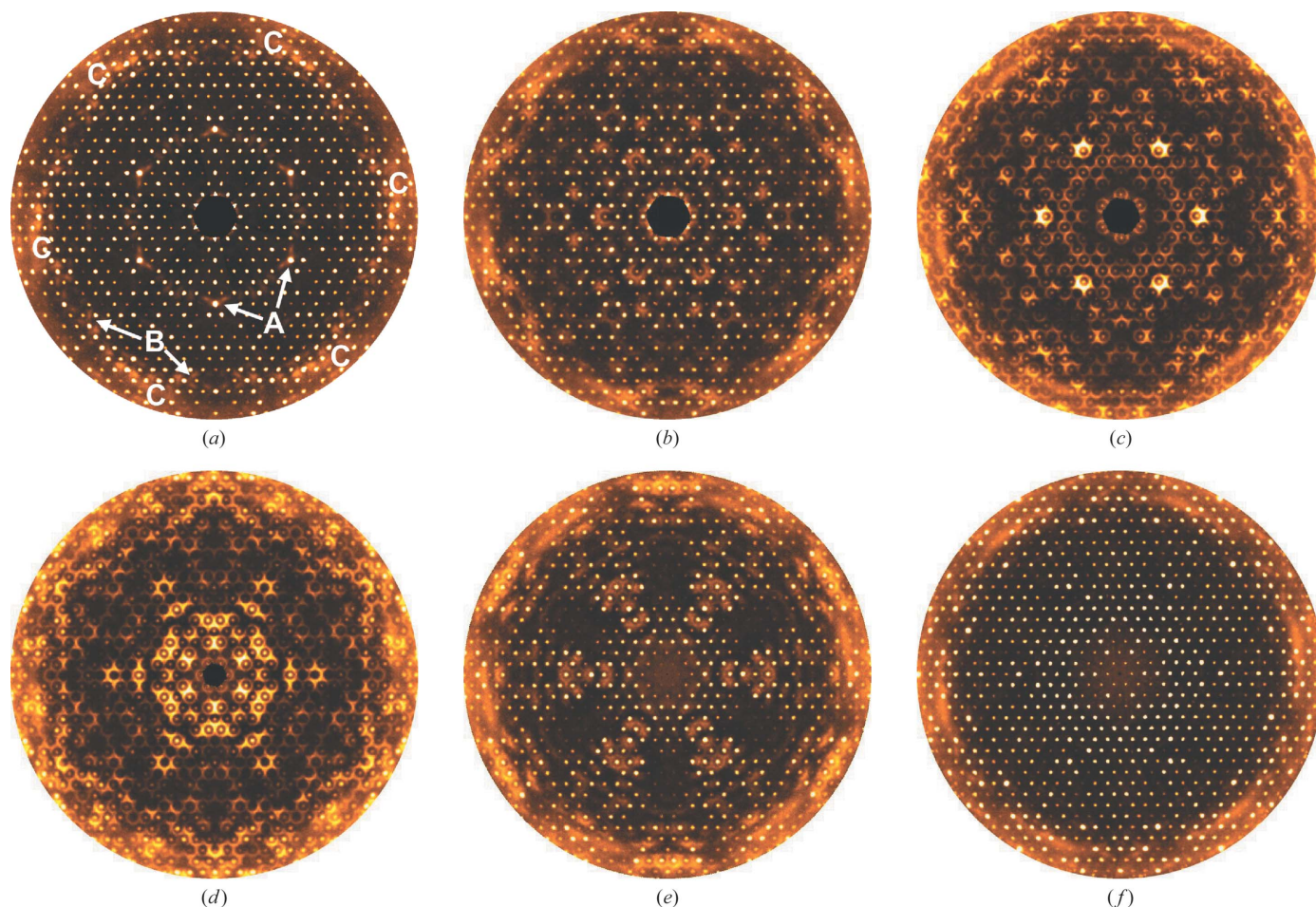

Figure 2
The first six reciprocal lattice sections normal to **c** plotted on the same intensity scale and with the same background correction. The plots are restricted to the region falling within the protein *solvent* ring. Similar diffuse features have been observed in the region of and outside the solvent ring, but to display the scattering in these regions a substantially different teatment of background would be required. Original quality files of the data displayed are included in the supplementary material.

diffuse scattering but $hk7$, $hk8$, $hk12$ and $hk13$ contain strong diffuse scattering and only very weak Bragg peaks.

The diffuse scattering, where it is visible, is seen to be comprised of diffuse rings surrounding each Bragg position. The diameter of the rings is $\sim a^*$ so that rings around neighbouring Bragg peaks merge. There is no indication that the degree of diffuseness varies from ring to ring nor that the intensity within the rings varies with azimuthal angle. However, the actual intensity of the pattern of rings varies both from reciprocal cell to reciprocal cell and even within a given reciprocal cell. Such variation of intensity is clearly related to the scattering factor of the unit-cell contents and not the basic effect giving rise to the rings.

A number of general features of the diffuse patterns are worth comment.

(i) The fact that there is virtually no diffuse scattering in the $hk0$ section means that when viewed in projection down c the crystal does not appear disordered.

(ii) For all the sections from $hk1$ up to $hk14$ that show diffuse scattering the intensity does not diminish near the middle of the pattern, i.e. for $h = k = 0$. This means that the disorder must be occupational (or substitutional) in nature rather than caused by atomic or molecular displacements in the ab plane, since a characteristic of displacement disorder is that the scattering goes to zero at low $\mathbf{Q}$.

(iii) The very weak diffuse scattering features labelled 'A' and 'B' in Fig. 2(a) violate the statements in (i) and (ii) above and most likely arise because of some small displacements that occur in addition to the basic occupancy disorder. For the purposes of this paper these weak features will be ignored.

(iv) The features labelled 'C' in Fig. 2(a) are scattering from the mounting loop in which the crystal was held. Similar features occur in each of the sections shown.
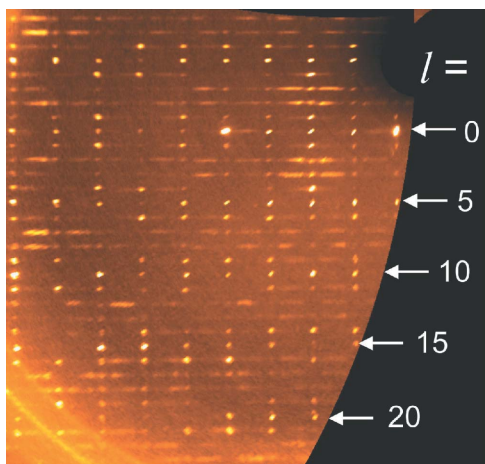

Figure 3
Part of a section of data normal to those in Fig. 2 showing how the diffuse scattering is confined to $l =$ integral layers. Note also how, to a good first approximation, the layers $l = 5n$, where n is an integer, contain only Bragg peaks. Similarly, the plot shows how the diffuse scattering is strongest mid-way between these $l = 5n$ layers.

In §3 we outline the theory required for substitutional disorder in order to demonstrate the significance of the fact that no diffuse scattering appears on $hk0$, $hk5$, $hk10$, $hk15$ etc.

3. Diffraction from a binary disordered crystal

Equation (1) gives the intensity of scattering from a disordered crystal that contains two types of scatterer, $\mathbf{A}$ and $\mathbf{B}$, which occupy the nodes of a regular lattice (i.e. we assume no atomic or molecular displacements, see Welberry, 2004). $\mathbf{A}$ and $\mathbf{B}$ may be single atoms, whole molecules or collections of molecules within a unit cell. In this expression m_A and m_B are the concentrations of the two species and F_A and F_B are the corresponding atomic or molecular scattering factors. The C_n are a set of correlation coefficients that describe the distribution of the two species. Here $\mathbf{R}$ is a real-space vector and $\mathbf{S}$ is a reciprocal space vector.

$$I(\mathbf{S}) = (m_A F_A + m_B F_B)(m_A F_A^* + m_B F_B^*) \sum \exp(2\pi i \mathbf{S} \cdot \mathbf{R}) + m_A m_B (F_A - F_B)(F_A^* - F_B^*) \sum C_n \exp(2\pi i \mathbf{S} \cdot \mathbf{R}) \tag{1}$$

The first term in (1) corresponds to the Bragg peaks and it is seen that their intensities are proportional to the average scattering factor $(m_A F_A + m_B F_B)(m_A F_A^* + m_B F_B^*)$. The second term corresponds to the diffuse scattering and shows that this intensity is proportional to the difference in scattering factors, $(F_A - F_B)(F_A^* - F_B^*)$.

In order to explain the fact that every fifth reciprocal layer contains no Bragg peaks consider a situation in which $\mathbf{A}$ and $\mathbf{B}$ are in fact the same molecule or arrangement of molecules but one is translated along $\mathbf{c}$ by $0.2\mathbf{c}$. Then we can write

$$F_B = F_A \exp(2\pi i \mathbf{S} \cdot \mathbf{c}/5) \tag{2}$$

whence

$$F_A - F_B = F_A(1 - \exp(2\pi i \mathbf{S} \cdot \mathbf{c}/5)). \tag{3}$$

Equation (3) shows that when $\mathbf{S} = 5n \times \mathbf{c}^*$, for n an integer the value of $F_A - F_B$ is zero and there is no diffuse intensity. Similarly the maximum intensity occurs when $\mathbf{S} = 5(n + \frac{1}{2}) \times \mathbf{c}^*$. For the Bragg peaks, assuming $m_A = m_B$

$$F_A + F_B = F_A(1 + \exp(2\pi i \mathbf{S} \cdot \mathbf{c}/5)). \tag{4}$$

This equations shows that the intensity of the Bragg peaks is maximum for $\mathbf{S} = 5n \times \mathbf{c}^*$, for n integral and a minimum when $\mathbf{S} = 5(n + \frac{1}{2}) \times \mathbf{c}^*$. This explains the variation of the relative intensity of the Bragg peaks and diffuse scattering in the reciprocal sections shown in Fig. 2.

4. Disorder and frustration on a triangular lattice

Diffuse rings or circles of scattering similar to those observed here (see Fig. 2) have been observed previously in a number of quite different disordered materials, ranging from the aluminosilicate ceramic mullite (Welberry & Withers, 1990) to the organic inclusion compound didecylbenzene/urea inclusion compound (Welberry, 2001). In these cases the effects were

attributed to a competition between the forces tending to produce a particular local structure and the strain that progressively builds up as this structure is incorporated into the lattice of the average long-range structure.

In the present case, although the phenomena are clearly related to those examples mention above, a rather simpler explanation can be given. This can be done by invoking the concept of 'geometric frustration', which may occur when two different 'types' of scatterer attempt to pack together on the triangular mesh that forms the *ab* plane of the hexagonal crystal. The archetypical example of such frustration is the Ising antiferromagnet on a triangular lattice (Wannier, 1950). Although the effect has thus been known for a long time it still has an important application in many present day fields of research. For example, the structure of sodium cobalt oxide, NaCoO, features a triangular lattice in which the electron spins cannot align in the regular arrangement seen in rectangular lattices. The geometric frustration this causes leads to a variety of unusual electronic and magnetic properties (Ong & Cava, 2004).

To explain the idea of frustration in more detail, suppose that each unit cell contains one or other of two different possible molecular configurations, say **A** or **B**. [In fact it is known that the unit cell contains 12 monomer units of the *N*-terminal fragment of the Gag protein, so **A** and **B** might represent two different arrangements of these.] We further suppose that if a particular unit cell contains **A** then a **B** in a neighbouring cell is energetically favourable while an **A** is energetically unfavourable. On a square lattice this is easily achieved by alternating **A** and **B** to produce a superlattice. However, on a triangular lattice (as we have here) two sides of any triangle may be **AB**, but the third side must then be either **AA** or **BB**. This is known as 'geometric frustration'. Fig. 4 shows this graphically. If the blue and red circles represent the **A** and **B** it is seen that in each of the (equilateral) triangles two of the sides are **AB** or **BA**, but then the third side is **AA** or **BB**.

Lattice realisations illustrating this type of frustration have been obtained using Monte Carlo (MC) simulation. Suppose $\sigma_{i,j}$ are $(+1, -1)$ random variables on a two-dimensional triangular lattice (see Fig. 4). Then we use an MC energy of the form

$$E_{MC} = \sum_i \sigma_i \left[\sum_{n.n.} J_1 \sigma_k + \sum_{n.n.n.} J_2 \sigma_l \right]. \quad (5)$$

Here J_1 is the interaction energy for nearest neighbours and J_2 the corresponding energy for the next-nearest neighbours (as shown in Fig. 4). For values of $J_1 > 0$ nearest neighbours tend to be different, *i.e.* **AB** or **BA**, although this cannot be achieved on all sides of a triangle which leads to the frustration mentioned above.

Simulations were carried out for a series of examples using the normal Metropolis algorithm (Metropolis *et al.*, 1953). A value of $J_1 = 1.0$ was kept constant for the whole series and J_2 was varied. Three examples are shown in Fig. 5. The values of J_1 and J_2 are relative to $k_B T$ in the Boltzmann transition probability, $P = \exp(-\Delta E / k_B T)$, where T is the simulation

temperature and k_B is Boltzmann's constant. The lattice realisations are shown in two different magnifications in order to show both the overall textural qualities of the three patterns and the detail of the near-neighbour contacts. The corresponding diffraction patterns computed from these distributions are shown in Figs. 5(*d*), (*e*) and (*f*). All three patterns show a dark circular region surrounding each Bragg peak that arises from the positive value of J_1, but the diffuse scattering shows some secondary structure that varies with the value of J_2. In Fig. 5(*d*) the diffuse scattering has a maximum at the centroid of the triangle of Bragg peaks, whereas in Fig. 5(*f*) the maxima are midway between each pair of Bragg peaks. In Fig. 5(*e*) the diffuse scattering shows no peaks and the scattering appears as featureless diffuse rings. Since this figure shows characteristics very similar to the observed patterns, the distibution in Fig. 5(*b*) was adopted for use in subsequent investigations that are described in the remainder of the paper.

The diffraction patterns in Fig. 5 and subsequent figures in this paper were calculated using the program *DIFFUSE* (Butler & Welberry, 1992). For the calculations in Fig. 5 the values of F_A and F_B were taken as $F_A = f_C$ and $F_B = 0$, where f_C is the atomic scattering factor for carbon.

5. A simple three-dimensional helical model

In this section we present results for a simple three-dimensional model that incorporates the various different aspects of the basic scattering problem described above, *i.e.* we require:
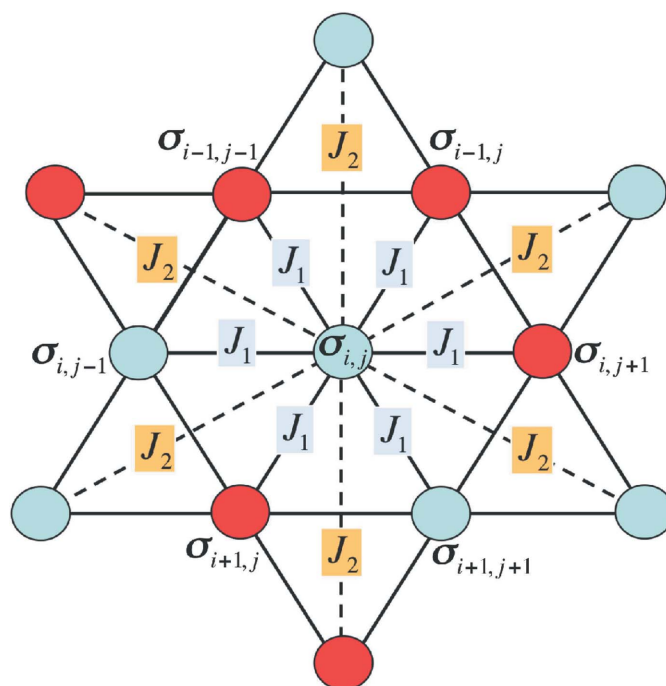

Figure 4
Frustration on a triangular lattice. The two differently coloured discs (red and blue) represent molecular configurations **A** and **B**. If two sides of a triangle contain **AB** then the third side must be **AA** or **BB**. The occupants of different sites on the lattice are represented by the random variables $\sigma_{i,j}$. J_1 and J_2 are interactions between $\sigma_{i,j}$ and its neighbouring variables.

(i) A model which exhibits frustration on the *hk*0 triangular mesh.

(ii) A model that satisfies the requirement that the two different scattering objects **A** and **B** are in fact the same molecule or arrangement of molecules but one is translated along **c** relative to the other by 0.2**c**.

(iii) A model that conforms to the average crystal symmetry in the space group $P6_122$.

A model satisfying these criteria is shown in Fig. 6. This has a single atom in the asymmetric unit with the fractional coordinate (0.25,0.05,0.0) in the space group $P6_122$. The 12 sites per unit cell thus form a helical chain running along **c**. Figs. 6(*a*) and (*b*) show the chain in its normal **A** position. The
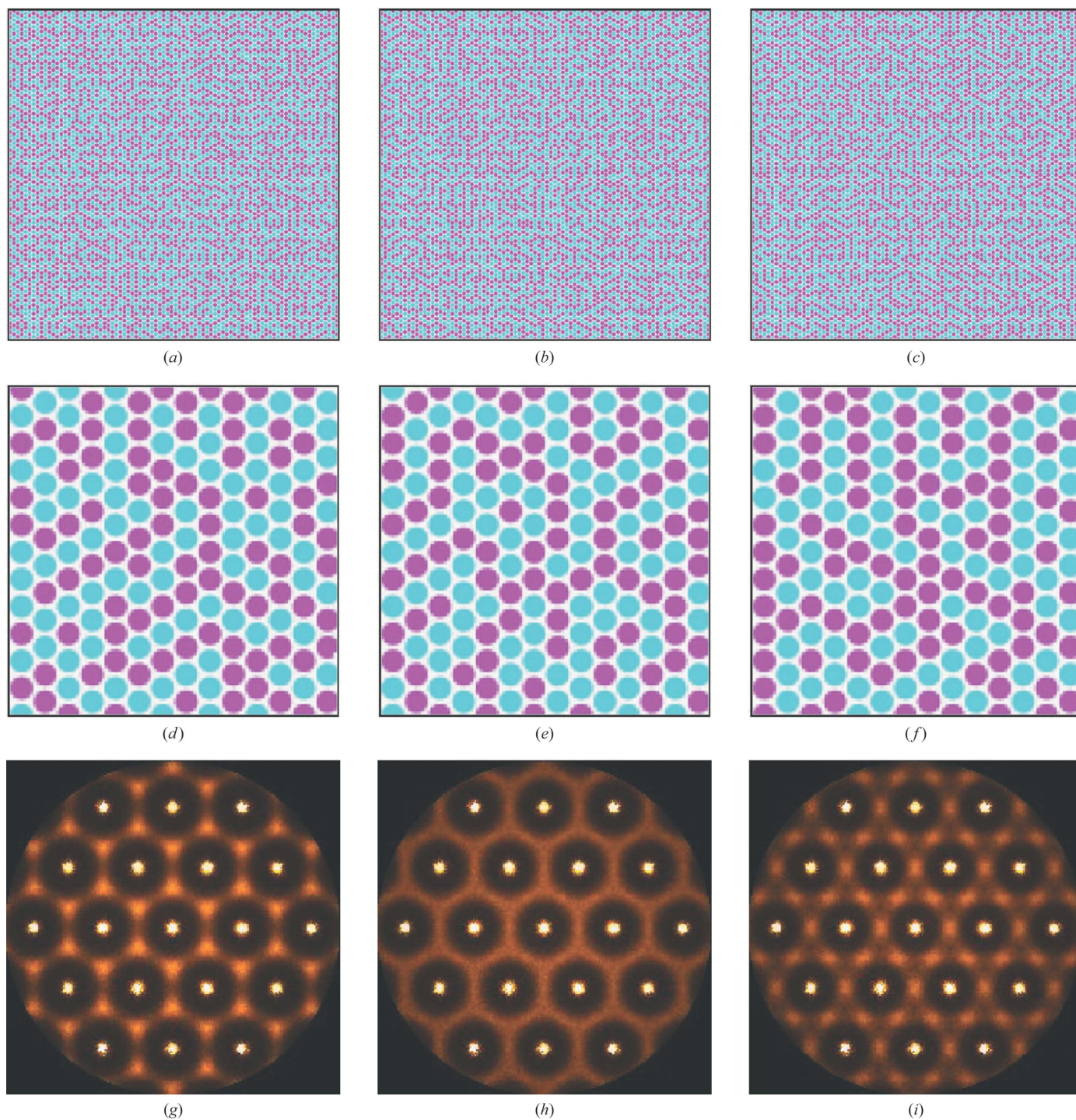

Figure 5
(*a*), (*b*) and (*c*) show example realisations of the Ising model defined by (5) with (*d*), (*e*) and (*f*) showing magnifications of (*a*), (*b*) and (*c*). For all three examples $J_1 = 1.0$ and the three differ in the value of J_2. These are: (*a*) $J_2 = 0.0$; (*b*) $J_2 = 0.125$; (*c*) $J_2 = 0.25$. (*g*), (*h*) and (*i*) show the corresponding computed diffraction patterns.

model assumes that each column of atoms seen in projection in Fig. 6(*a*) is either in the **A** or **B** position. Figs. 6(*c*) and (*d*) show such a disordered structure in which chains in the normal **A** (blue) position are intermingled with chains in the alternative **B** (pink) position, which is displaced by 0.2**c** relative to **A**. The distribution of the **A** and **B** columns was assumed to be identical to that of the two-dimensional model shown in Fig. 5(*b*).

Diffraction patterns computed from this model are shown in Fig. 7. For these calculations the distribution in the *ab* plane of the **A** and **B** molecular chains was as in Fig. 5(*b*) and each of the chains was considered to be essentially a perfect one-dimensional crystal along **c**. This means that the diffuse scattering only occurs within the Bragg layers at integral values of *l*. In practice it might be supposed that the chains extend only over a finite distance along **c**. In this case the thickness of the diffuse layers in **c*** will be proportional to the reciprocal of the correlation length along **c** (see Fig. 3).

5.1. Discussion of calculated diffraction patterns

It should first be noted that the patterns in Fig. 7 cover a much larger area of reciprocal space than those shown in Fig. 5, with *h*, *k* indices in excess of 7 included. These patterns now clearly show many of the attributes that were observed in the X-ray data. Reciprocal layers *hk*0 and *hk*5 contain Bragg peaks only. Layers *hk*1 and *hk*4 show Bragg peaks and diffuse scattering. Layers *hk*2 and *hk*3 show predominantly diffuse scattering with only a few weak Bragg peaks discernible. The pattern of diffuse rings around each Bragg position that was shown by the single atom model in Fig. 5(*e*) is clearly retained, but now the intensity varies from cell to cell as a result of modification by the molecular scattering factor term, $(F_A - F_B)$, in (3). Given the overall similarity of the patterns produced by this simple model and the observed data it is not difficult to imagine that the only difference is that for the real protein structure a more complicated molecular scattering factor term would pertain.
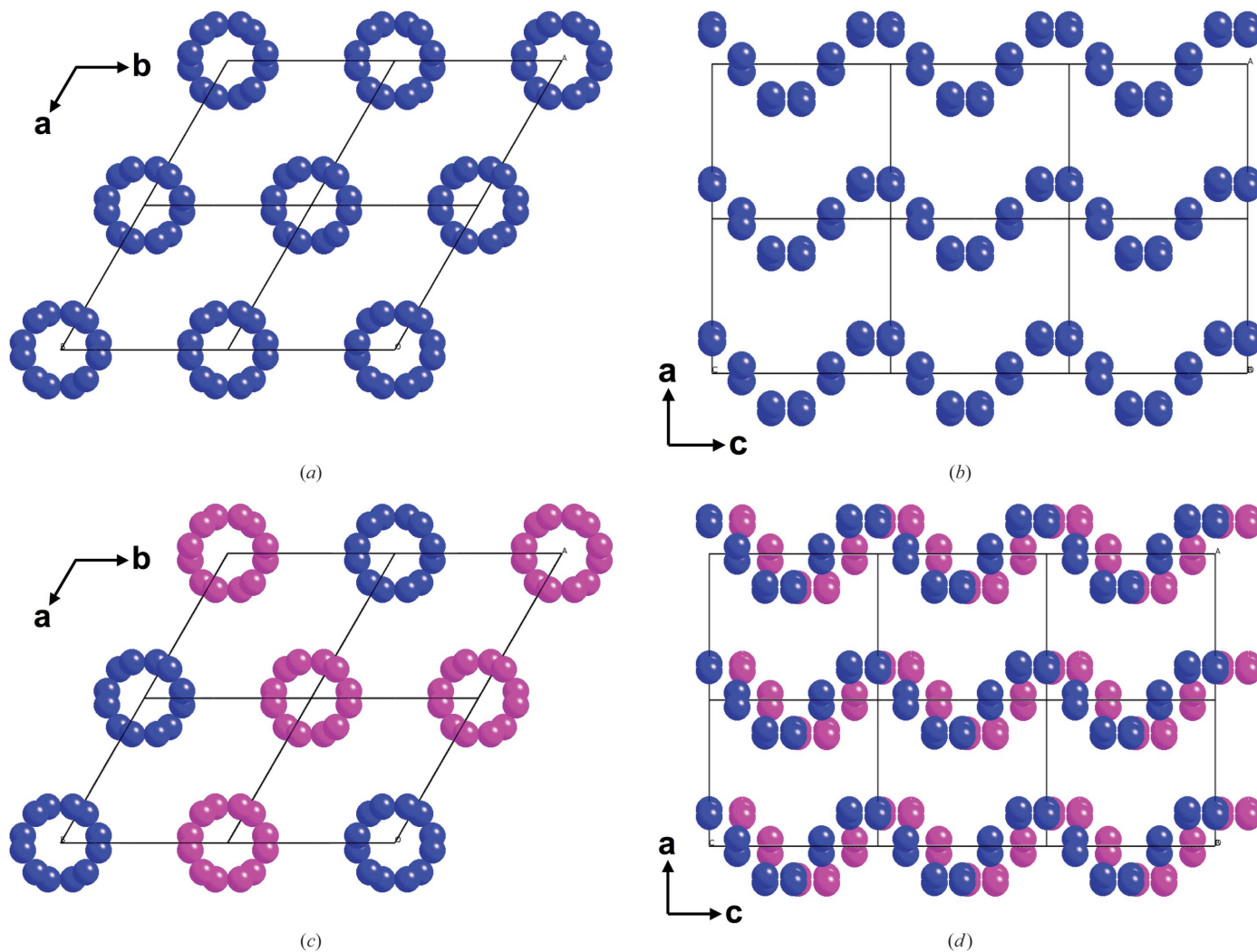

Figure 6
The simple three-dimensional helical model comprising 12 atoms per unit cell. (*a*) View down **c**; (*b*) view down **b**; (*c*) same as (*a*) but showing how the two alternative types of unit cell containing molecules in the **A** (blue) and **B** (pink) configuration can form a disordered crystal. (*d*) Corresponding view down **b**.

The explanation of diffuse scattering in terms of frustration on a triangular lattice, as described by the Ising model of (5), may seem somewhat unphysical in crystals of the N-terminal fragment of Gag protein since the distances involved are many times larger than in the original example (Wannier, 1950). In the present case nearest-neighbours are 73 Å apart so that the interaction between 'spin-like' variables might seem physically inappropriate. Nevertheless, the formulation does appear to describe the observed phenomena very well.

The most important feature of the distributions that have been produced by the model is that there is a very strong tendency for nearest-neighbouring cells to be of the type **AB** rather than **AA** or **BB**. Such a tendency does not require interactions on the scale of 73 Å but can arise just through steric hindrance effects between the outer regions of the molecular clusters within each unit cell. These will only be on the scale of van der Waals radii. However, it should be noted that in order to produce smooth featureless diffuse rings the Ising formulation required an albeit relatively small contribution from second-nearest neighbours. These are ~ 125 Å apart and are not in direct contact, so steric arguments cannot be easily invoked. Any second neighbour involvement must result indirectly from successive nearest-neighbour contacts.

It is possible that the steric effects that result in the preference for **AB** rather than **AA** or **BB** neighbour cell configurations are simply invoked during crystal growth as protein monomers are added to the crystal. However, since the protein is present in solution in dimer form it might be considered that during crystal growth it is these dimers that are being incorporated into the crystal. If this is the case then it might be envisaged that the dimers themselves are the source of the strong nearest-neighbour correlations. For example, if the two halves of the dimer correspond to a pairing of (say, twofold related) monomers that span the boundary of a blue and a pink unit cell in Fig. 6(c), then having such pre-formed **AB** pairs could facilitate growth of the kind of frustrated distributions shown in Fig. 5. At present such a possibility is purely speculative and how things work out in practice must await further study.

6. Discussion

In this paper we have shown that the diffuse ring features that have been observed in the diffraction patterns of a crystal of the N-terminal fragment of the Gag protein from Feline Foamy Virus are caused by geometric frustration as the indi-
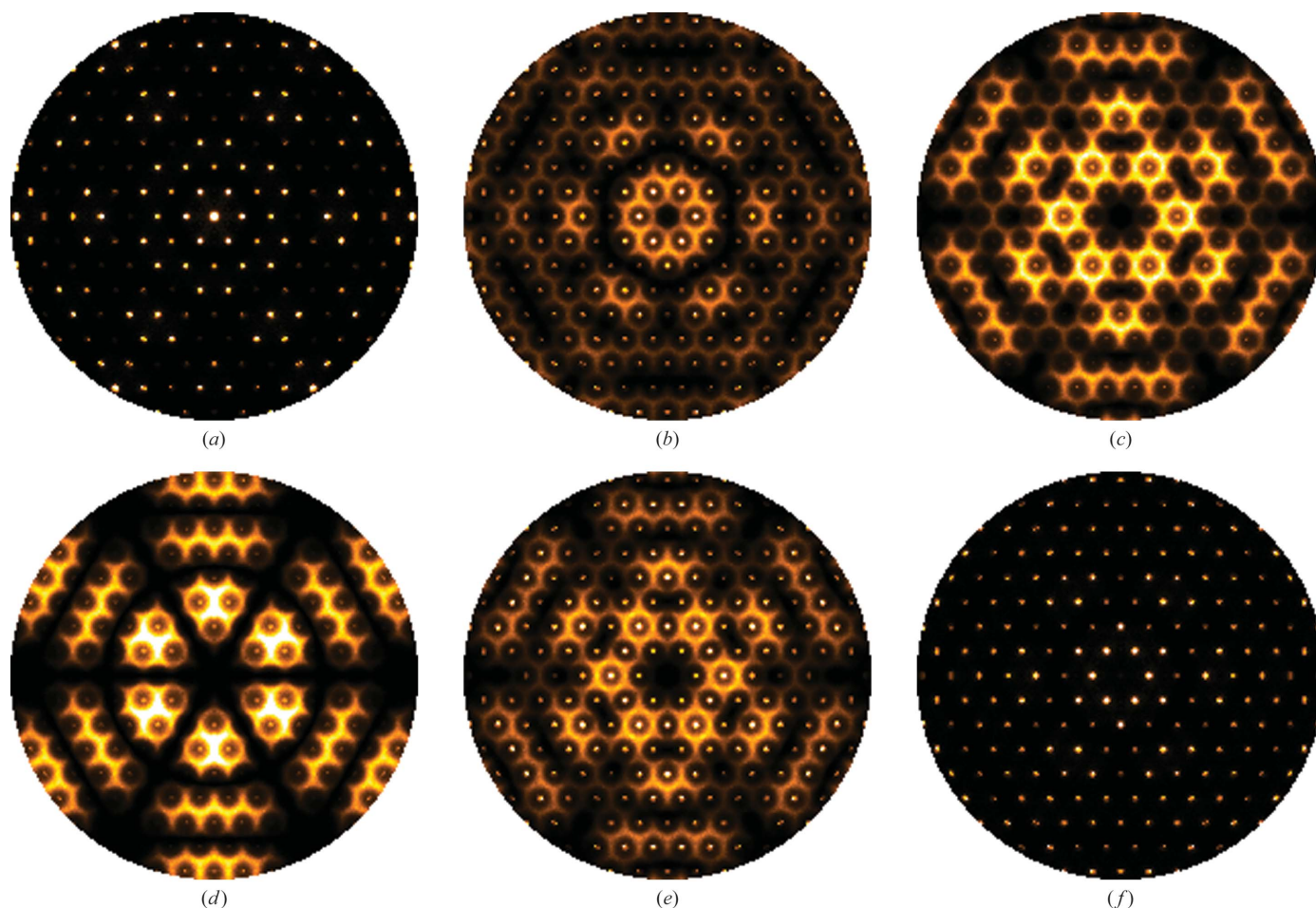

Figure 7
Diffraction patterns computed from the simple three-dimensional helical model shown in Fig. 6. (a) hk0; (b) hk1; (c) hk2; (d) hk3; (e) hk4; (f) hk5.

vidual molecules try to pack in the triangular *bc* mesh in the hexagonal space group $P6_122$. In order to explain the strong diffuse scattering it is necessary for the crystal to contain occupational disorder such that each unit cell contains a molecular arrangement of one or other of two different kinds, **A** and **B**. The frustration arises because the nearest-neighbour packing prefers neighbouring cells to be **AB** or **BA** (and not **AA** or **BB**). This cannot be achieved on all three sides of any triangle.

In order to explain the fact that reciprocal sections $hk5n$, where n = integer, contain only Bragg peaks it is necessary that **A** and **B** are in fact identical molecular arrangements and differ only by a translation of 0.2**c**. This also explains why the diffuse scattering is strongest on sections midway between these Bragg-only layers (*e.g. hk2* and *hk3*).

At the time of writing, attempts at conventional structure solution from the Bragg reflection data set have not so far been successful. Compared with a non-disordered structure with the same size unit cell the number of available Bragg reflections is substantially reduced. Only one fifth of the reciprocal layers normal to **c** have normal Bragg peak intensities; two fifths have Bragg peaks with much reduced intensity and two fifths have virtually no Bragg peak intensities. This means that overall the number of observables per structural parameter is also greatly reduced, at least by a factor of two, and this on its own might well prevent solution using the Bragg peaks alone. In this case it may be necessary to undertake experiments to modify exposed loops or reduce surface entropy (Cooper *et al.*, 2007) to prevent adverse packing interactions, with the aim of generating a new crystal form.

While the Bragg peaks contain much less structural information than usual, the diffuse scattering clearly contains a great deal. The modulation of the basic diffuse ring pattern by the molecular scattering factor term $(F_A - F_B)$, in (3) clearly gives rise to a great deal of extra information over and above anything that is obtainable from the Bragg peaks. In fact, because of the observed special condition that F_A and F_B only differ in phase due to the 0.2**c** shift, the distribution of intensity within the diffuse rings can be directly related to F_A. Since this information is available over the whole of the $h, k, 5nl + 2$ and $h, k, 5nl + 3$ sections and not just at integral values of h and k, it would seem quite feasible to use this information directly to determine the molecular

shape. Further work to investigate this possibility is planned.

References

Butler, B. D. & Welberry, T. R. (1992). *J. Appl. Cryst.* **25**, 391–399.
Caspar, D. L., Clarage, J., Salunke, D. M. & Clarage, M. (1988). *Nature*, **332**, 659–662.
Chacko, S. & Phillips, G. N. (1992). *Biophys. J.* **61**, 1256–1266.
Chan, E. J., Welberry, T. R., Goossens, D. J., Heerdegen, A. P., Beasley, A. G. & Chupas, P. J. (2009). *Acta Cryst.* B**65**, 382–392.
Chan, E. J., Welberry, T. R., Heerdegen, A. P. & Goossens, D. J. (2010). *Acta Cryst.* B**66**, 696–707.
Cooper, D. R., Boczek, T., Grelewska, K., Pinkowska, M., Sikorska, M., Zawadzki, M. & Derewenda, Z. (2007). *Acta Cryst.* D**63**, 636–645.
Doucet, J. & Benoit, J. P. (1987). *Nature*, **325**, 643–646.
Du, Y., Huang, M., Chang, S., Schlagel, D. L., Lograsso, T. A. & McQueeney, R. J. (2010). *Phys. Rev. B*, **81**, 054432-1–054432–9.
Estermann, M. A. & Steurer, W. (1998). *Phase Transitions*, **67**, 165–195.
Glover, I. D., Harris, G. W., Helliwell, J. R. & Moss, D. S. (1991). *Acta Cryst.* B**47**, 960–968.
Kolatkar, A. R., Clarage, J. B. & Phillips, G. N. (1994). *Acta Cryst.* D**50**, 210–218.
Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H. & Teller, E. (1953). *J. Chem. Phys.* **21**, 1087–1092.
Ong, N. P. & Cava, R. J. (2004). *Science*, **305**, 52–53.
Paściak, M., Wołcyrz, M., Pietraszko, A. & Leoni, S. (2010). *Phys. Rev. B*, **81**, 014107.
Pérez, J., Faure, P. & Benoit, J.-P. (1996). *Acta Cryst.* D**52**, 722–729.
Riccardi, D., Cui, Q. & Phillips, G. N. (2009). *Biophys. J.* **96**, 464–475.
Wall, M. E., Clarage, J. B. & Phillips, G. N. (1997). *Structure*, **5**, 1599–1612.
Wannier, G. H. (1950). *Phys. Rev.* **79**, 357–364.
Weber, T. & Bürgi, H.-B. (2002). *Acta Cryst.* A**58**, 526–540.
Weber, T., Schaub, P. & Steurer, W. (2008). *Acta Cryst.* A**64**, C131–C132.
Weidner, E., Frey, F., Lei, J.-L., Pedersen, B., Paulmann, C. & Morgenroth, W. (2004). *J. Appl. Cryst.* **37**, 802–807.
Welberry, T. R. (2001). *Acta Cryst.* A**57**, 244–255.
Welberry, T. R. (2004). *Diffuse X-ray Scattering and Models of Disorder.* IUCr Monographs on Crystallography. Oxford University Press, Oxford.
Welberry, T. R. & Withers, R. L. (1990). *Phys. Chem. Miner.* **17**, 117–124.