STRUCTURAL SCIENCE
CRYSTAL ENGINEERING
MATERIALS

ISSN 2052-5206

Simulation of diffuse scattering in DL-norleucine

Thomas Richard Welberry[a]* and Carl Henrik Görbitz[b]

[a]Research School of Chemistry, Australian National University, Canberra, ACT 2601, Australia, and [b]Department of Chemistry, University of Oslo, PO Box 1033, Blindern, N-0315 Oslo, Norway. *Correspondence e-mail: welberry@rsc.anu.edu.au

Received 5 March 2019
Accepted 9 April 2019

Edited by R. Černý, University of Geneva, Switzerland

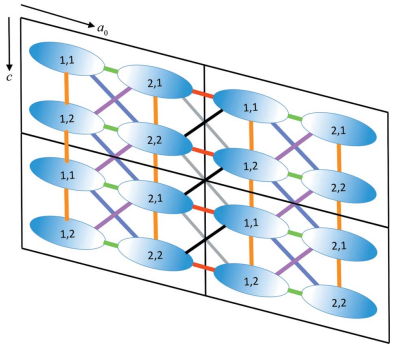
Keywords: DL-norleucine; diffuse scattering; Monte Carlo simulation; racemates; molecular bilayers.

The diffraction patterns of DL-norleucine (*SR*-2-aminohexanoic acid, DL-Nle) crystals may show obvious diffuse scattering, usually described as 'streaking', between the Bragg peaks. This phenomenon is obviously related to the non-ideal behaviour of the crystal. The normal interpretation is disorder in the stacking of weakly interacting 2D layers, known also for a number of other racemates of amino acids with linear hydrophobic side chains, as well as 1:1 complexes between different L- and D-enantiomers (quasi-racemates). Presented here is the first attempt to extract the information hidden in the diffuse scattering for this group of compounds by applying Monte Carlo simulations to the site distributions of two polymorphs in a block of 48 × 48 × 48 unit cells (four sites in each unit cell, 442 368 in total). The results demonstrate that it is indeed possible to model the diffuse scattering and relate it to processes expected to take place during phase transitions, characterized by slipping of molecular bilayers (or parts of them) relative to their neighbours. The understanding of the (intermediate) mixed phases in terms of domain size and defect density is consequently brought to a new level.

1. Introduction

Amino acids with linear side chains, as enantiomers and racemates as well as quasi-racemates display a unique series of phase transitions where molecular bilayers slip with respect to each other, most frequently in a reversible and non-destructive manner. The two end phases of such a transition are sometimes 'clean' in the sense that no traces of the old phase are seen above or below the transition temperature. For other crystals, the transitions may be gradual and/or partial, leading to extensive stacking disorder reflected by 'streaking' in the diffraction patterns, *i.e.* diffuse scattering between more or less smeared out Bragg peaks. This phenomenon was noted as early as 1952 by Mathieson (1952) in his investigation of two crystal forms of DL-Met, called α and β, and has since recurred for numerous other amino acid crystals. For the racemate of norleucine (2-aminohexanoic acid, Nle), a non-natural chain isomer of leucine and isoleucine with an *n*-butyl side chain, Mathieson (1953) determined the structure of one crystal form, corresponding to the $P2_1/c$ (or $P2_1/a$ with a different setting) polymorph of DL-methionine (DL-Met), but also deduced the existence of a structure of a second form, the β form, in space group $I2/a$ (or $C2/c$). He called the two possible types of double layers *L* and *M*, and described the α form as an *L L L L* arrangement [Fig. 1(*a*)] and the β form as *L M L M* [Fig. 1(*b*)]. Diffuse spectra were suggested to arise from the sequence *L L M M* in one crystal, but *L L L M* in another for which the diffuse distribution was different.

After observations by Mnyukh (1975) of a reversible change in DL-Nle crystals at 390 K under a polarizing micro-



scope, Harding *et al.* (1995) collected synchrotron radiation Laue diffraction patterns over a range of temperatures and demonstrated (i) the presence of streaks, indicating disorder in the stacking of layers within the crystal at room temperature, (ii) the complete or almost complete disappearance of these streaks when the crystal was heated to *ca* 373 K, indicating annealing of the crystal, and (iii) the loss of single-crystal characteristics above 390 K so that it was not possible to record diffraction data for the high-temperature form proposed by Mathieson.

The crystal structure of the low-temperature β form of DL-Nle was determined at 120 K by Dalhus & Görbitz (1996),

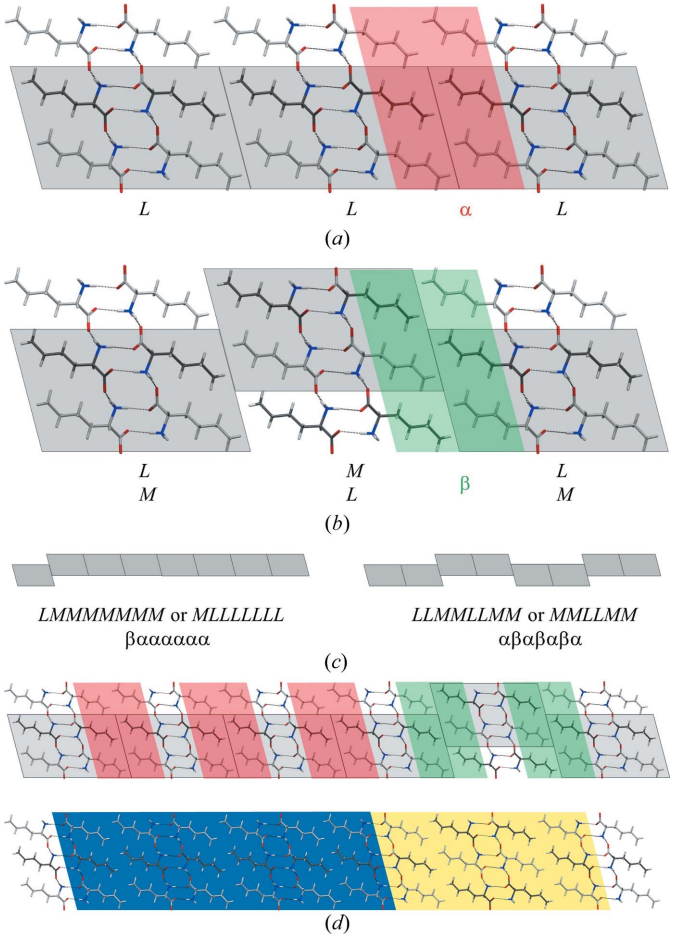


Figure 1
(*a*) The crystal packing of α-DL-Nle. Molecules of L and D chirality have C atoms shown in light and dark grey, respectively. Three molecular bilayers are shown; each grey parallelogram here corresponds to a crystallographic unit cell. This particular packing is called an *LLL* stacking sequence, which indicates that there are no translational components in any direction. (*b*) Alternative stacking in β-DL-Nle. This is called an *LML* stacking, where an *M* bilayer is translated half a unit cell along the vertical axis, readily seen in the illustration, and along the viewing direction (less easily seen here) compared with its neighbour. Alternatively, interactions at the hydrophobic interfaces could be identified as either α-type (red shaded) or β-type (green shaded), yielding a unique description of any stacking sequence (see text for details). (*c*) Two distinct packing sequences described in terms of *L*–*M* units or α–β. The systematic arrangement to the right has been observed as an intermediate state for two quasi-racemates (Görbitz & Karen, 2015). (*d*) The shift of an individual molecular bilayer resulting in an $\alpha\alpha\alpha\alpha\beta\beta$ sequence, depicted using two different colouring schemes: as in panels (*a*) and (*b*) with α and β bilayers identified by red and green shading (top), or by simply using blue and yellow, respectively, for entire molecular layers to show these two types of stacking (bottom). The latter format is used in Fig. 6 and subsequent illustrations.

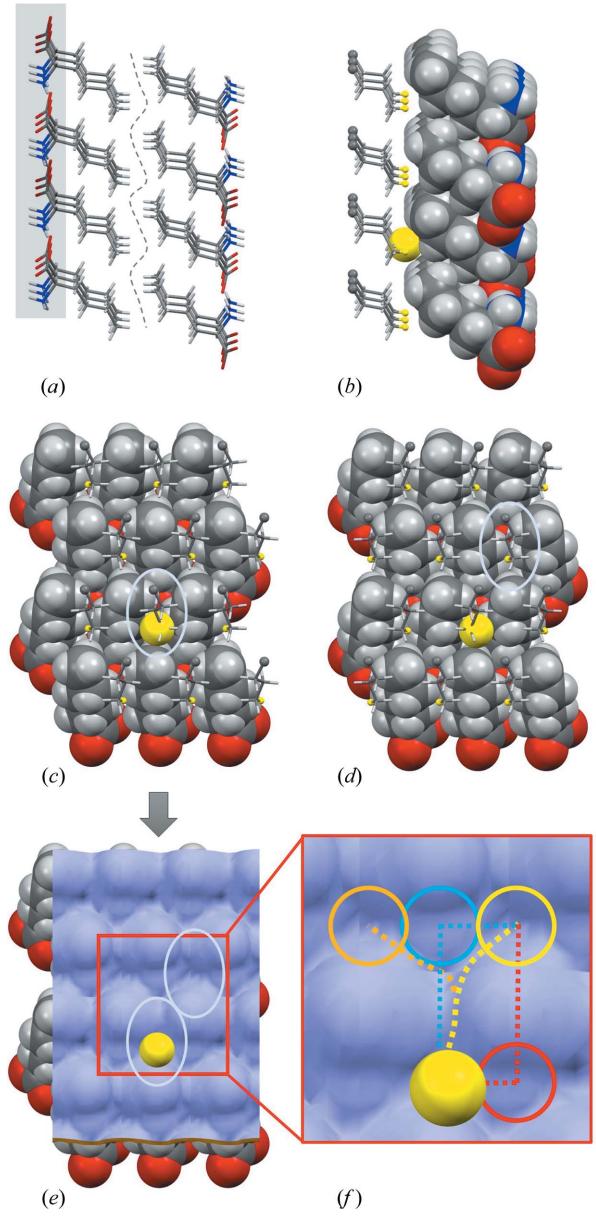


Figure 2
(*a*) A section of the crystal structure of α-DL-Nle (Harding *et al.*, 1995), with the central hydrophobic interface emphasized by a dashed line. (*b*) The same section after removal of the polar heads covered by the grey shade in panel (*a*). C_β atoms (their H atoms not shown) and the ultimate H atoms on C_ε are shown as small spheres, except one H atom depicted in a space-filling representation. Molecules belonging to the monolayer to the right are all in a space-filling representation. (*c*) The same section after a 90° rotation about the *y* axis. (*d*) The equivalent section for β-DL-Nle (Dalhus & Görbitz, 1996). (*e*) The same view as in panel (*c*), but highlighting the solvent-accessible surface of the bottom layer, as calculated by *Mercury* using default settings (Macrae *et al.*, 2008). Only the H atom shown in a space-filling representation in panels (*b*) and (*c*) is retained from the upper layer. (*f*) An enlarged section of the content of the red box in panel (*e*) showing various motion paths during phase transitions (see text for details).

who confirmed a *C*2/*c* (or *I*2/*a*) structure closely related to the β phase of DL-Met, *i.e.* with an *L M L M* layer arrangement. As a point detector was used on that occasion, diffuse scattering was not taken into consideration. Ultimately, the structure of a third phase, γ, in the same space group as β, was determined at 395 K (Coles *et al.*, 2009). The sharp β-to-γ transition at 390 K displays very little hysteresis and is in principle reversible, but specimens usually crack into very small pieces. Bragg reflections are thus hard to observe after cooling. We will not deal with this transition in any further detail.

The α-to-β transition temperature is, by comparison, much less well defined. It has been reported to occur at 253 K upon cooling and ∼295 K upon heating ('room temperature'), but again with large crystal-to-crystal variations and signs that both phases may coexist in some crystals over temperature ranges of up to 100 K (van den Ende *et al.*, 2015). The transition does not involve any conformational changes for the amino acid side chain, which remains fully ordered in an all-*trans* configuration. It is furthermore unique in that there appear to be essentially no differences between the unit-cell volumes of the two phases. Related structures exhibit small to sizeable volume changes during corresponding transitions, *e.g.* 7.1% for DL-norvaline (*n*-propyl side chain) (Görbitz, 2011).

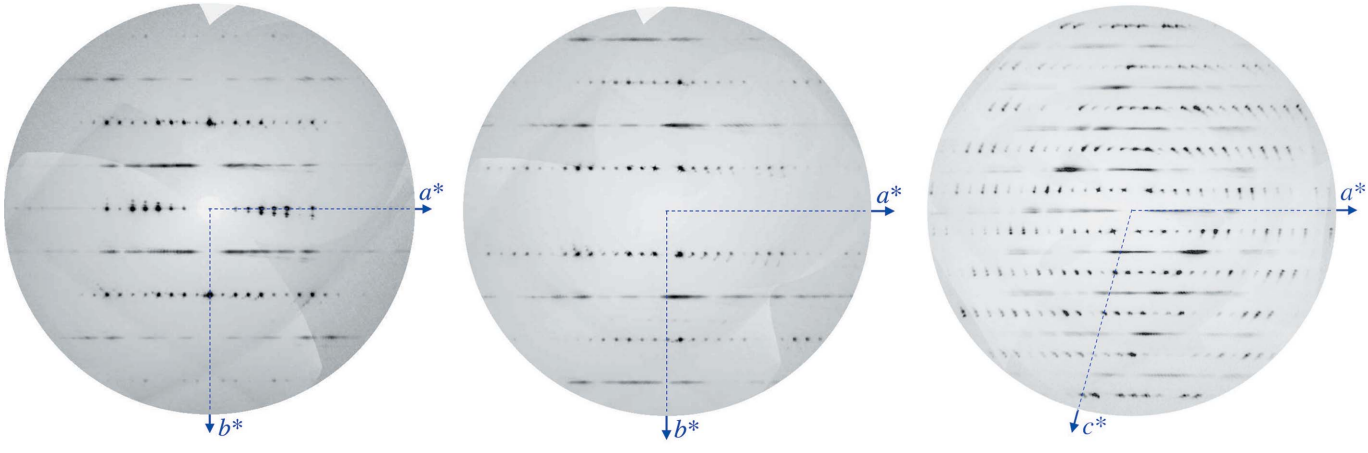
As demonstrated in Fig. 1(*c*), the *LM* terminology has the inherent weakness that it depends on the choice of a reference layer. The sequence shown in Fig. 1(*b*) could consequently also be labelled *LML* or *MLM*. If instead one focuses on interactions at the hydrophobic interfaces between bilayers, their identification as either α (red shading, L-enantiomer side chain facing L-enantiomer side chain, D-enantiomer facing D-enantiomer) or β (green shading, L-enantiomer facing D-enantiomer) is unique, Fig. 1(*c*). We will use the $\alpha\beta$ notation in this article and the yellow–blue colour coding in Fig. 1(*d*).

An important aspect of the α-to-β transition (and indeed all transitions of related compounds) is that the amino acid polar heads and the hydrogen bonds formed between them are unperturbed; everything appears to happen at the interface between neighbouring molecular bilayers. This interface is highlighted in Fig. 2(*a*), with the 'knobs-into-holes' fit between opposing surfaces further emphasized in Fig. 2(*b*). Figs. 2(*c*) and 2(*d*) show rotated views for the α- and β-polymorphs, respectively, revealing how the selected terminal side-chain H atom, in a quite remarkable manner, fits into precisely the same depression regardless of polymorph. What this means is that a phase transition from α to β involves a relocation of this H atom [together with its mother molecule, inside the pale-blue ellipse in Fig. 2(*c*)] to a new position [pale-blue ellipse in Fig. 2(*d*)]. In the simplified view in Fig. 2(*e*), this means starting at the solid sphere and ending at the yellow circle, or potentially at the orange circle. Either movement corresponds to half a unit-cell length along the 9.9 Å vertical axis and along the 4.7 Å horizontal axis, *i.e.* a $(0, \frac{1}{2}, \frac{1}{2})$ shift for a 16.4, 4.7, 9.9 Å primitive unit cell.

The special nature of the DL-Nle α-to-β transition has triggered several investigations trying to elucidate the transition mechanism. Most recently, Cuppen and coworkers have used a range of methods, in particular molecular dynamics (MD) simulations, to study mechanisms and energy barriers, (van den Ende *et al.*, 2015, 2016; Smets *et al.*, 2015). One of their results is a potential shift only along the 9.9 Å axis $(0, 0, \frac{1}{2})$. In Fig. 2(*f*) this would imply moving from the solid-sphere position to the blue circle. There is no hole or depression at this point, but it could still be relevant for a dynamic system.

More recently, Czech *et al.* (2017) used density functional theory (DFT) calculations for models with 100 double layers. Their conclusion is that only $(0, \frac{1}{2}, \frac{1}{2})$ shifts are likely to occur, as static structures after $(0, 0, \frac{1}{2})$ and $(0, \frac{1}{2}, 0)$ shifts were calculated to have significantly higher energies.

With the advent of single-crystal diffractometers that utilize area detectors, it is now a simple matter to use the raw data frames from a crystal structure determination not only to obtain integrated intensities of the Bragg peaks, but also to construct reciprocal-space plots of any desired reciprocal-



(*a*) (*b*) (*c*)

Figure 3
Diffraction patterns of three reciprocal sections of DL-norleucine that show diffuse scattering. These are: (*a*) *hk*0, (*b*) *hk*1 and (*c*) *h*1*l*. Sections 0*kl*, *h*0*l* and 1*kl* show no diffuse scattering and have been omitted.

Table 1
Cell data for DL-Nle.

See Fig. 4(*a*) for the relationship between the original and transformed cells. *T* (K) is the data-collection temperature. Refcodes refer to the Cambridge Structural Database (CSD; Groom *et al.*, 2016).

Reference	Phase	*a* (Å)	*b* (Å)	*c* (Å)	*β* (°)	*V* (Å^3)	CSD refcode	*T* (K)
Mathieson (1953)	*α*	16.56	4.74	9.84	104.5	748	DLNLUA	293
Harding *et al.* (1995)	*α*	16.382	4.737	9.907	104.68	743.7	DLNLUA01	293
Dalhus & Görbitz (1996)	*β*	31.067	4.717	9.851	91.37	1443.2	DLNLUA02	120
Original centred cell	Mixed	31.713	4.726	9.901	93.09	1481.8	This work	293
Transformed cell	Mixed	16.354	4.726	9.901	75.50	740.94	This work	293
From current CSD data	Mixed	16.383	4.745	9.934	75.45	747.43	This work	293

lattice section. Such plots can reveal diffuse scattering due to the presence of disorder. Although diffuse scattering obtained in this way is not of the same quality as that which may be obtained in the best experiments using synchrotron radiation (see *e.g.* Chan *et al.*, 2010), it can nevertheless reveal useful structural information not available from the Bragg peaks alone.

Monte Carlo (MC) computer simulation of a model structure has become a powerful and well accepted technique for aiding the interpretation and analysis of diffuse scattering patterns (Welberry & Goossens, 2008; Welberry, 2004; Weber *et al.*, 2001; Weber & Bürgi, 2002). In this method, a computer model is set up based on known physical and chemical principles, with adjustable parameters describing the basic interatomic or inter- and intramolecular interactions. Diffraction patterns are calculated from the results of the simulations and these are compared with the observed data. Different models can be tested. Those that are incorrect can be definitely eliminated, while substantial confidence can be placed in a model that gives good agreement with the observation. In this article we show how such MC methods can be used to yield an unprecedented and informative structural model of an intermediate step in the transition between the two end structures *α*- and *β*-DL-Nle that reproduces the observed diffuse scattering.

2. Diffraction data

The diffraction images used in the present work were derived from a complete set of CCD frames recorded in a conventional crystal structure determination experiment using Mo *Kα* radiation from a fine-focus sealed X-ray tube source. The images were obtained from a crystal that had first been cooled to 100 K before being returned to 293 K for data collection. The plots shown in Fig. 3 are reciprocal-lattice sections extracted from these data frames using *CrysAlisPRO* (Oxford Diffraction/Agilent Technologies UK Ltd, Yarnton, England.).

It is clear from these images that the crystal sample was by no means perfect and clear signs of crystal splitting are evident. Nevertheless, the peak-finding and -indexing routines worked quite straightforwardly with peaks from the major crystal fragment to yield the unit-cell dimensions given in the bottom row of Table 1. A setting in which the *β* angle was acute (75.50°) was chosen so that the structure could be readily compared to the *C*-centred cell given in some publications. Fig. 4 shows the relationship between the two cells. The cell determined here is drawn in heavy lines.

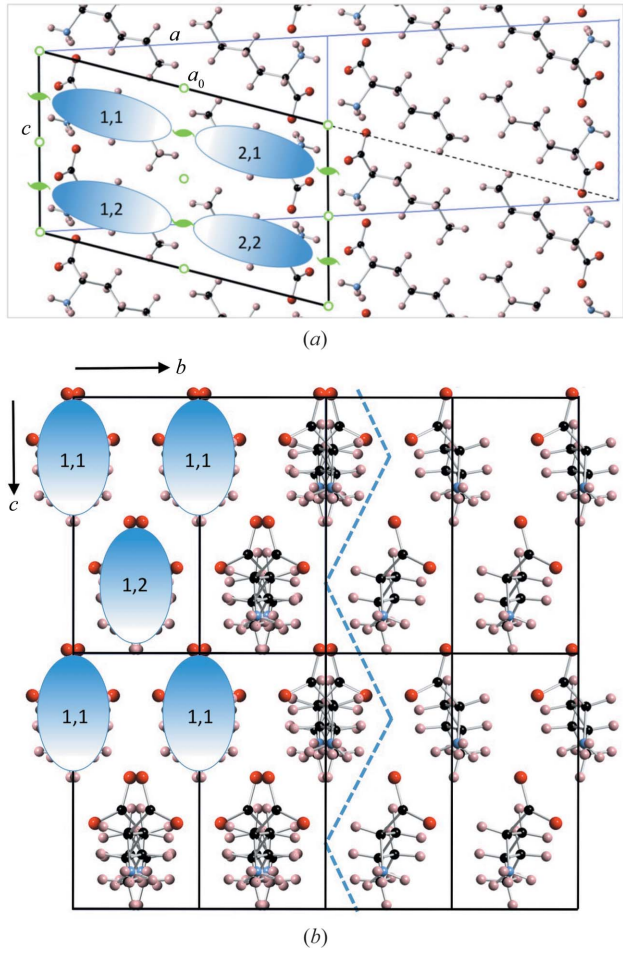


Figure 4
(*a*) The structure of DL-norleucine, viewed down *b*. The symmetry elements for space group $P2_1/c$ are shown in green. (*b*) The structure of a single molecular layer, comprising sites 1, 1 and 1, 2, viewed down a_0. Sites 2, 1 and 2, 2 have been omitted for clarity. In panel (*b*), for the molecular sites to the left of the dashed blue line both *A* and *B* molecules are shown superimposed, while to the right only *A* molecules are shown. In both panels (*a*) and (*b*) elliptical shapes are drawn as a simple motif to identify the molecular sites. These are used in Fig. 5.

Table 2
Fractional coordinates (x, y, z) for the molecular orientations A and B referred to the transformed cell given in Table 1 and shown in Fig. 4.

The coordinates for the other three sites in the unit cell are given by $(1-x, \frac{1}{2}+y, \frac{1}{2}-z)$, $(1-x, 1-y, 1-z)$ and $(x, \frac{1}{2}-y, \frac{1}{2}+z)$, corresponding to $P2_1/c$ symmetry.

	Molecule A			Molecule B		
Atom	x	y	z	x	y	z
C1	0.1097	−0.1142	0.1138	0.1097	0.1142	0.1138
C2	0.1489	0.0626	0.2103	0.1489	−0.0626	0.2103
C3	0.2372	−0.0574	0.2039	0.2372	0.0574	0.2039
C4	0.2827	0.0766	0.3045	0.2827	−0.0766	0.3045
C5	0.3700	−0.0429	0.2890	0.3700	0.0429	0.2890
C6	0.4128	0.0665	0.3976	0.4128	−0.0665	0.3976
O1	0.0667	−0.3222	0.1656	0.0667	0.3222	0.1656
O2	0.1266	−0.0450	−0.0120	0.1266	0.0450	−0.0120
N	0.0949	0.0478	0.3540	0.0949	−0.0478	0.3540

Distinctive diffuse streaking can be seen in the three low-index sections shown in Fig. 3: $hk0$, $hk1$ and $h1l$. The presence of crystal splitting does not detract significantly from the appearance of these. In fact, in the $hk0$ section a secondary weaker image of one of the streaks, clearly coming from the minor crystal fragment, can be seen inclined to the main streak. The relative intensities of these are a clear indication of the relative volumes of the major and minor crystal fragments. [Note that other low-index sections, $0kl$, $h0l$ and $1kl$ (not shown), do not contain diffuse scattering.]

3. Monte Carlo simulation

We assume that the crystal has $P2_1/c$ symmetry and that each molecular site contains one or other of two molecular conformations A and B. Coordinates x, y, z for these two conformations are given in Table 2. Molecules A and B differ only in their y coordinates; it is seen that A is the mirror image of B. In order to represent a disordered crystal, the occupancy of each molecular site is defined by a binary random variable $\sigma_{ijklm} = \pm 1$. The indices i, j, k designate the unit cell along the a_0, b and c directions, while l and m point to sites within the unit cell, as shown in Fig. 4(a). We define the energy of the system as

$$E_{\rm MC} = \sum_{pq} J_v \sigma_p \sigma_q. \qquad (1)$$

Here, p and q are neighbouring pairs of molecular sites (i, j, k, l, m) and (i', j', k', l', m') connected by a vector v with interaction parameter J_v (see Fig. 5).

Monte Carlo (MC) simulations were carried out using a model crystal comprising $48 \times 48 \times 48$ unit cells each containing four molecular sites, making 442 368 sites in total. Initially the random variable σ_{ijklm} at each site was set at random to be, with equal probability, 1 to represent occupancy by an A molecule or -1 to represent occupancy by a B molecule. Each MC run comprised 1000 individual MC cycles, where a cycle is defined as that number of individual MC steps required to visit each site once on average. Diffraction patterns were calculated from the resulting model crystal using the program *DIFFUSE* (Butler & Welberry, 1992).

The form of the Monte Carlo energy given by equation (1) was made sufficiently general so that different possible ordering schemes could be explored. The summation is over all near-neighbour intermolecular vectors, as shown in Fig. 5. The interaction parameters J_v are generally unkown but can be adjusted to induce the required correlation effects. We have generally found it most convenient to use an iterative feedback procedure in order to obtain particular target correlation values, C_v. Starting with initial arbitrary values for the J_v values, the induced correlations are measured after each MC cycle. These are compared with the target values and then the J_v values are adjusted accordingly by an amount proportional

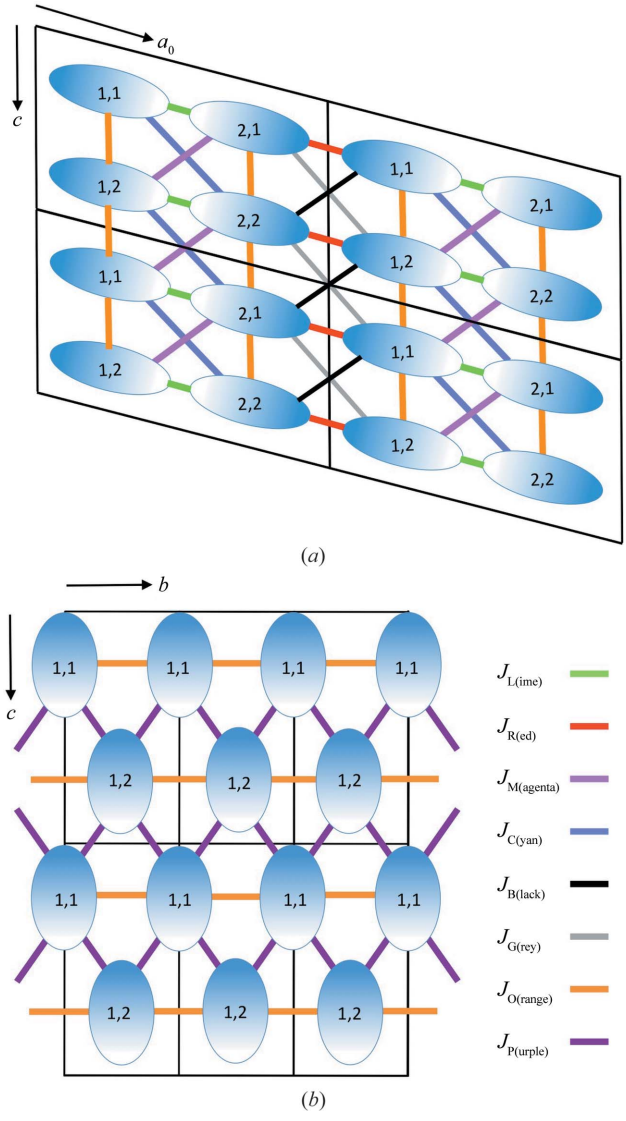


Figure 5
The structure of DL-norleucine, viewed down b and a_0 and represented schematically using the elliptic motifs shown in Fig. 4. The differently coloured vectors connecting pairs of molecular sites correspond to the interaction vectors used in the MC simulations. Panel (b) includes a key identifying the eight different types of interaction, J_v.

to the discrepancy. For the $\sigma = \pm 1$ variables, the nearest-neighbour correlation coefficient is given by

$$C_v = \frac{P_{11} - m_A^2}{m_A(1 - m_A)}. \quad (2)$$

Here, P_{11} is the joint probability, $P(\sigma_p = 1, \sigma_q = 1)$, and $m_A = (1 - m_B)$ is the concentration or fraction of A molecules in the crystal.

It should be noted that the designation of molecules as A or B (which is used in formulating the Monte Carlo model) is quite independent of the labelling of leucine molecules as D or L. A crystal comprising a perfectly ordered arrangement of A molecules generated by the $P2_1/c$ symmetry elements will contain two D- and two L-molecules in each unit cell. Neighbouring pairs of molecular sites are related by either the centre of symmetry or the screw rotation, as shown in Fig. 4(*a*). In a crystal containing all A molecules, the sites labelled (1, 1) and (2, 1) will be occupied by D-leucine, while sites (1, 2) and (2, 2) will be occupied by L-leucine.

To demonstrate how this carries over into the lattice plots which will be used later in the paper, we show in Fig. 6 a small region of a single *ac* plane taken from a simulated crystal. In Fig. 6(*a*) a blue motif is used to depict A molecules and a yellow motif to depict B molecules. In this mode of plot, the molecular bilayers are easily seen as long vertical columns with the hydrophilic interfaces in the centre. In Fig. 6(*b*) the same example is displayed using a different colouring scheme to identify whether each molecule is D or L. A black molecular motif corresponds to a D-molecule, while a grey motif corresponds to an L-molecule. In this mode of plot, the molecular bilayers are not so easily discernible. However, it readily allows the addition of background shading that indicates whether the local structure conforms to the α-phase structure (blue) or the β-phase structure (yellow). See also Fig. 1.

4. Example simulations I

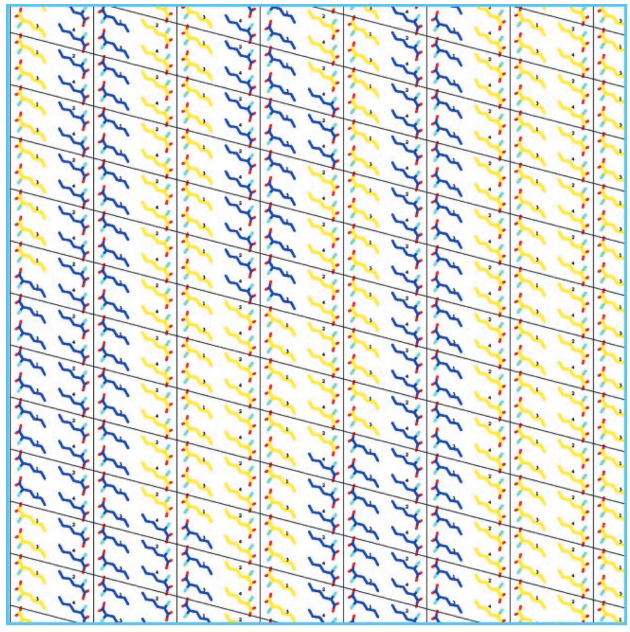
Some example simulations and their calculated (*h*1*l*) diffraction patterns are shown in Fig. 7. Values of the correlations C_v induced by application of a chosen set of J_v for each example are given in Table 3. The numbers given in bold are correlations that have been specified as input and the lighter italicised numbers are correlations that have consequently resulted along other vectors.

4.1. Figs. 7(*a*)–7(*c*)

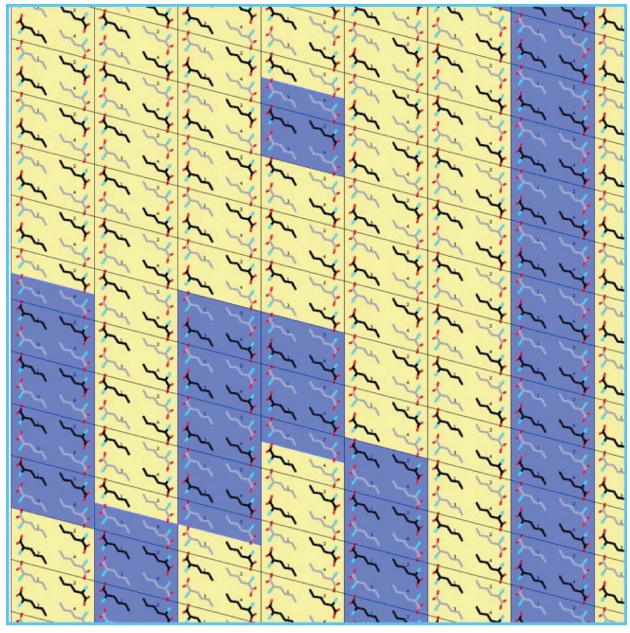
For this example each molecular site is occupied by an A or a B molecule at random (with 50% probability). There are no correlations between any pairs of sites. The diffraction pattern shows a continuously smooth distribution of diffuse scattering, reflecting the molecular scattering factor. This is the so-called Laue monotonic scattering.

4.2. Figs. 7(*d*)–7(*f*)

Now some short-range order has been introduced along nearest-neighbour vectors, both within *bc* layer planes (C_O and C_P; orange and purple vectors) and between pairs of these planes (C_L and C_R; lime-green and red vectors). The diffuse scattering has the same overall form but is modulated such that the intensity is concentrated in diffuse peaks or blobs.



(*a*)



(*b*)

Figure 6
(*a*) A small region of a single layer of a simulated crystal in which A and B molecules are displayed as blue or yellow motifs, respectively. (*b*) The same region of the crystal but with the molecular motifs now displayed as either black or grey, respectively, according to whether they are D- or L-leucine molecules. The background shading in panel (*b*) indicates whether the local structure is characteristic of the α-phase structure (blue) or β-phase structure (yellow). See also Fig. 1.

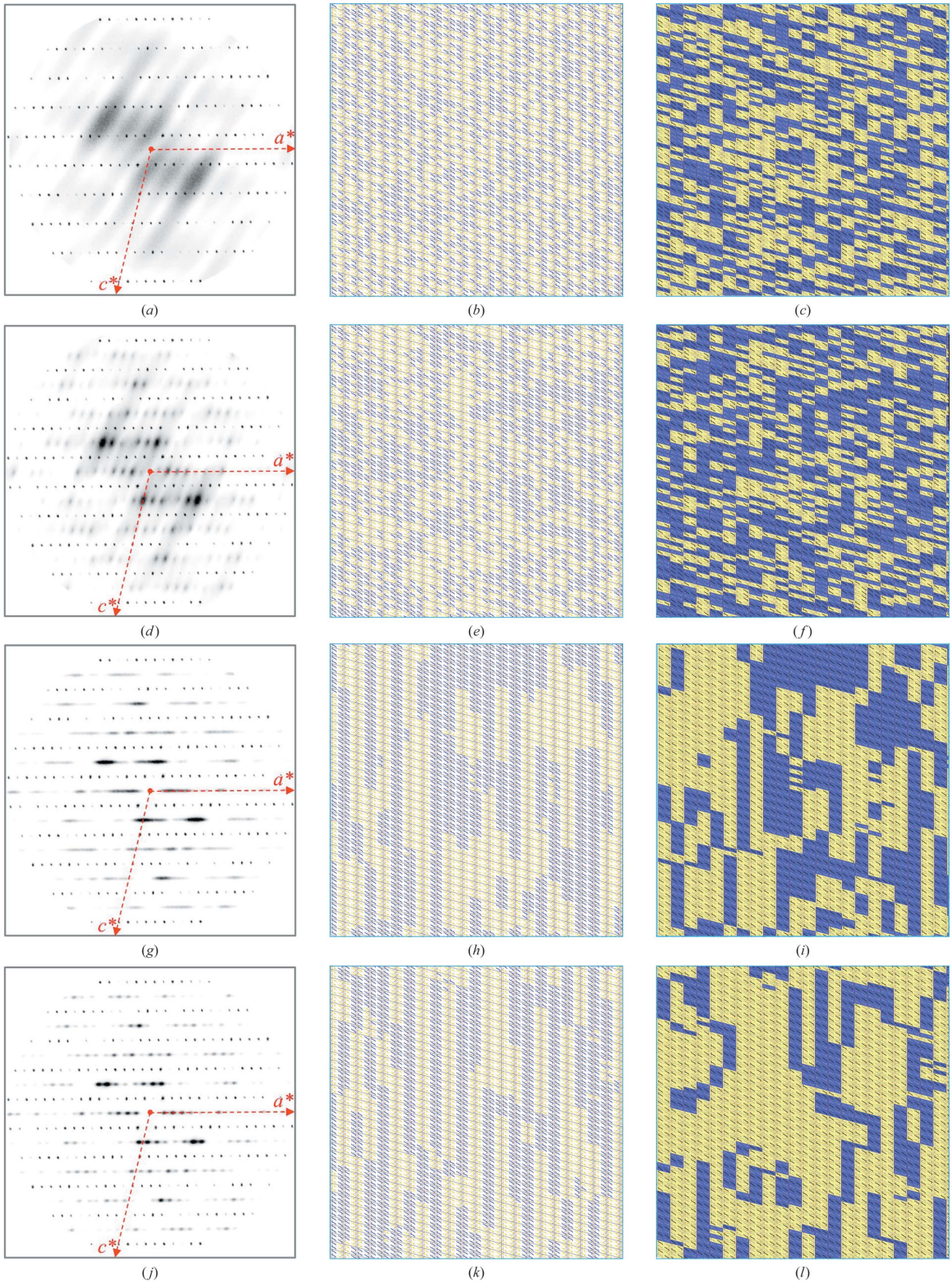


Figure 7
Example simulations of the disordered structure of DL-norleucine and their corresponding calculated (*h*1*l*) diffraction patterns. The central panels show plots of part a single layer of the simulation in which blue and yellow motifs represent molecules of types *A* and *B*, respectively. The panels on the right show the same distributions as the central panels but with the molecules now displayed as black or grey motifs to identify D- or L-enantiomers, respectively. The blue and yellow background shading identifies local regions that conform to the α or β phases, respectively. Note how the example of panel (*l*) has a higher proportion of the β phase than the example of panel (*i*). The four examples shown in this figure all have a 50:50 concentration of *A* and *B* molecules and correspond to the first four examples listed in Table 3. See text for further details.

Table 3
Key to the examples shown in Figs. 7 and 8.

The table shows the correlations C_i between the occupancies of pairs of molecular sites separated by the differently coloured vectors shown in Fig. 5. **Bold** numbers correspond to those vectors along which a direct interaction J_i was applied, while *italicised* values correspond to vectors along which correlation resulted consequentially. mA is the overall occupancy used, *i.e.* the probability that any site is occupied by an A molecule.

Figure	mA	C_O	C_P	C_L	C_M	C_C	C_R	C_B	C_G
7(*a*)	**0.50**	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
7(*d*)	**0.50**	**0.40**	**0.40**	**0.40**	*0.34*	*0.34*	**0.40**	*0.34*	*0.34*
7(*g*)	**0.50**	**0.91**	**0.91**	*−0.01*	*−0.01*	*−0.00*	**0.95**	*0.91*	*0.91*
7(*j*)	**0.50**	**0.90**	**0.90**	**−0.40**	*−0.39*	*−0.40*	**0.95**	*0.91*	*0.91*
8(*a*)	**0.50**	**0.88**	**0.88**	*−0.00*	*−0.00*	*−0.00*	**−0.94**	*−0.88*	*−0.88*
8(*d*)	**0.50**	**0.88**	**0.88**	*−0.00*	*−0.00*	*−0.00*	**0.00**	*−0.00*	*−0.00*
8(*g*)	**0.20**	**0.91**	**0.91**	*0.03*	*0.03*	*0.03*	**0.95**	*0.92*	*0.92*
8(*j*)	**0.50**	**0.90**	**−0.91**	*−0.01*	*−0.01*	*0.00*	**0.95**	*−0.91*	*−0.91*

4.3. Figs. 7(*g*)–7(*i*)

For this example the range of order within the *bc* layer planes has been greatly increased via high values of C_O and C_P. This results in long vertical columns of like-occupied molecular sites. The diffraction pattern consequently displays scattering in the form of horizontal streaks, the widths of which are inversely related to the correlation lengths in the vertical columns. Application of the strong correlation C_R between pairs of *bc* layers (along the red vectors) results in the segregation of these vertical columns into pairs or bilayers. These bilayers are ones which contain the hydrophilic interface in which the $-COO^-$ and $-NH_3^+$ groups are in close proximity. The probability of having *AA* or *BB* pairs in the horizontal direction is very high for sites separated by the red vectors but is quite random for sites separated by the lime-green coloured vectors. These bilayers are easy to see in the central panel (*i*). Using the second mode to colour panel (*h*) reveals which parts conform locally to the α-phase structure (blue) or β-phase structure (yellow).

The segregation into layer pairs affects the distribution of intensity along the horizontal diffuse streaks. It can be seen that the distribution is quite smooth and the variation of intensity derives solely from the molecular transform of a bilayer. This simulation is considered to provide the best agreement with the observed scattering shown in Fig. 3. Calculated diffraction patterns and plots of the structure are shown later in the paper for other sections of this example. Note that for this random stacking of bilayers there appear to be approximately equal amounts of α and β, in agreement with the traditional crystallographic refinement.

4.4. Figs. 7(*j*)–7(*l*)

This example is identical to the previous example [Figs. 7(*g*)–7(*i*)], with the only difference being the inclusion of a negative correlation $C_L = -0.4$ along the lime-green vectors. This relatively small amount of correlation has a significant effect on the distribution of intensity along the diffuse streaks. The previous smoothly varying diffuse streak is now broken up into a row of diffuse peaks. Note that the negative correlation results in these peaks appearing at the $h = n + \frac{1}{2}$ positions. In the real-space image [panel (*k*)] it is seen that there is a tendency for *AA* and *BB* bilayers to alternate. This becomes even more apparent in the alternative colouring plot [panel (*l*)] which shows that the proportion of the whole plane that conforms to the β structure is now much higher than the proportion of α.

5. Example simulations II

Some further example simulations and their calculated ($h1l$) diffraction patterns are shown in Fig. 8. In selecting these examples the aim has been to show the range of different ordering patterns that are possible but that clearly differ from the observed patterns.

5.1. Figs. 8(*a*)–8(*c*)

This example shows the effect of reversing the strong correlation C_R that resulted in the formation of the *AA* and *BB* bilayers evident in Figs. 7(*e*) and 7(*g*). Now the large negative value of C_R means that bilayers of types *AB* or *BA* are favoured. There is still no correlation along the green vectors. Note now how the variation in intensity along the diffuse lines is quite different and, in particular, how there is strong intensity around the origin.

5.2. Figs. 8(*d*)–8(*f*)

In this example only the correlations C_O and C_P within a single molecular layer are non-zero. Each molecular plane is therefore quite independent of the others. There is no tendency to form bilayers, either of types *AA*/*BB* or *AB*/*BA*. The C_O and C_P correlations within the planes still result in long vertical columns in the lattice plots and these give rise to the horizontal diffuse streaks, but the intensity variation along them is now quite different.

5.3. Figs. 8(*g*)–8(*i*)

This example is designed to show the effect of having a non-50:50 proportion of *A* and *B* molecules while maintaining the same correlation structure as the example in Figs. 7(*g*) and 7(*h*). The fraction of *A* molecules here is only 20%. The diffuse streaking has the same profile as that in Fig. 7(*g*) but the intensity is reduced by a factor $(0.2 \times 0.8)/(0.5 \times 0.5) = 0.64$ (see *e.g.* Welberry, 1985). On the other hand, the Bragg peaks are proportional to the average structure factor which will be different from the 50:50 mixture. As a result, weak additional Bragg peaks can be seen along the diffuse scattering streaks.

5.4. Figs. 8(*j*)–8(*l*)

In this final example we show the effect of reversing the sign of the correlation C_P. This has the effect in the *bc* plane of flipping the molecule at the face-centring position so that, instead of being the opposite enantiomer of the molecules at the corner of the cell, it is now the same enantiomer. The

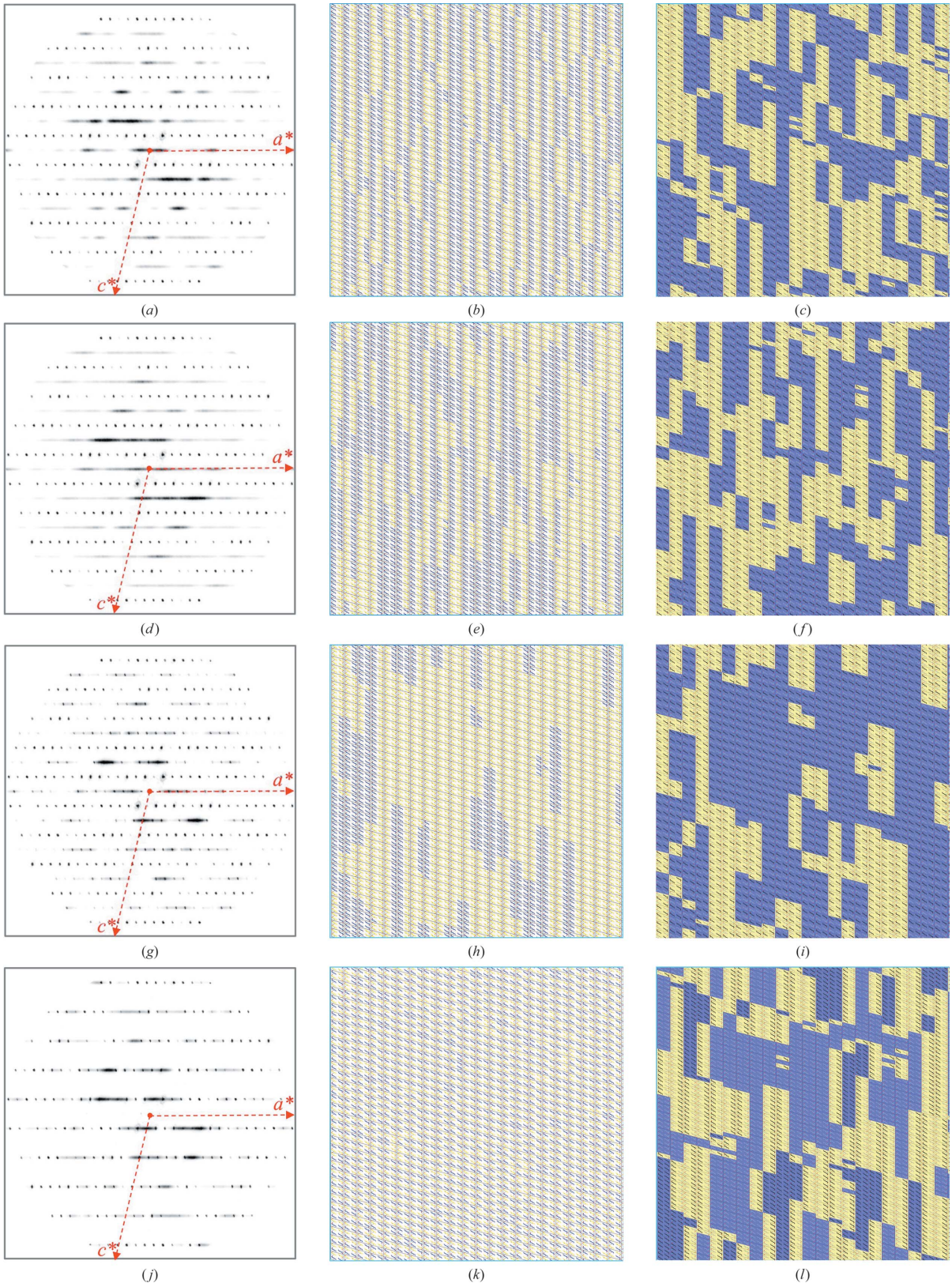


Figure 8
Further example simulations of the disordered structure of DL-norleucine and their corresponding calculated (*h*1*l*) diffraction patterns. (*a*)–(*c*) Patterns for an example simulation which contains predominantly *AB* or *BA* bilayers instead of *AA* or *BB*. (*d*)–(*f*) Patterns for an example containing only monolayers with no tendency to form bilayers. (*g*)–(*i*) The effect of reducing the fraction (occupancy) of *A* molecules in the sample while maintaining the same correlation structure as Figs. 7(*g*)–7(*i*). (*j*)–(*l*) The effect of reversing the sign of the C_P correlation, which has the effect of inducing regions in each *bc* plane to contain all D- or all L-enantiomers. See text for more details.

strong C_R value is maintained, so maintaining the bilayers. The blue/yellow motif colouring plot in Fig. 8(*k*) is now difficult to interpret, but the background shading plot in Fig. 8(*l*) shows an interesting structure with three different types of local region visible. The blue-shaded areas have two different types of region containing either all D- or all L-enantiomers. In contrast, the yellow shaded areas have alternating bilayers that are all L- or all D-enantiomers. The diffuse scattering, while still comprising similar streaking, now occurs along the rows of Bragg peaks instead of at the $(n+\frac{1}{2})c^*$ positions.

6. Final best-fit model

The diffraction pattern shown in Fig. 7(*g*) corresponds closely to that of the observed (*h*1*l*) section data shown in Fig. 3(*c*). After exploration of many different disordered structures in the present study, this example is considered to provide the 'best fit' to the observed data. As confirmation of the veracity of this model, we show in Fig. 9 additional diffraction patterns computed from this model for the (*hk*0) and (*hk*1) sections. These may be compared with the observed data for these sections, Figs. 3(*a*) and 3(*b*), respectively. The blue/yellow motif colouring of the lattice sample in Fig. 9(*a*) shows clearly the *AA* and *BB* bilayers extending (vertically) in the *c* direction. These columnar features are in fact two-dimensional domains in *bc* planes, as shown in Fig. 9(*b*). The lengths of the columns in Fig. 9(*a*) correspond to line segments through the domains seen in Fig. 9(*b*). The whole structure is produced by just three correlation parameters, C_O and C_P within the *bc* planes and the correlation C_R which forces the individual planes to form bilayers. In the absence of any other correlation these bilayers are quite independent.

7. Significance of the final best-fit model

Starting with a clean end phase such as α [Fig. 10(*a*)], there are a number of different mechanisms that can lead to a gradual change to another end phase. In Fig. 10(*b*) these have been categorized as the four different models, namely *B* for bilayer slip, *F* for a fault-type slip, *M* for slipping of multiple layers and *D* for a coherent process leading to clear domains for two (dual) phases. Model *M*, with slippage at every second hydrophobic interface, was incidentally observed to be a uniform and stable phase for the two quasi-racemates L-norvaline:D-norleucine and L-norvaline:D-methionine over temperature ranges of 8 K and 22 K, respectively (Görbitz & Karen, 2015). In addition to slipping along the 9.9 Å vertical axis and along the 4.7 Å axis of the viewing direction, the introduction of defects complicates the mixing and balance

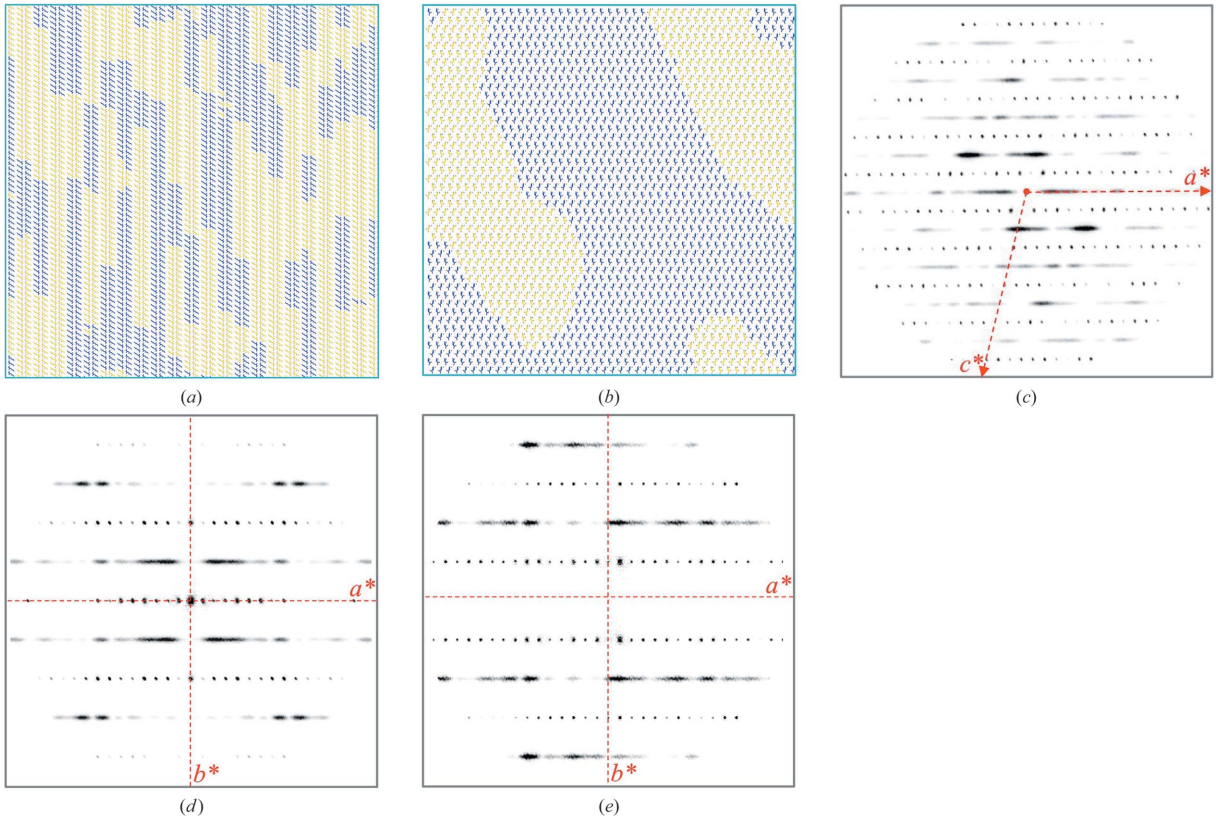


Figure 9
The final best-fit model corresponding to example 7(*g*) given in Table 3. Blue motifs correspond to *A* molecules and yellow to *B* molecules. (*a*) A plot of a single *ac* layer with the elongated *AA* bilayer columns evident. (*b*) A plot of a single *bc* molecular layer, revealing the 2D nature of the bilayers. Only molecular sites 1, 1 and 1, 2 are shown; sites 2, 1 and 2, 2 have been omitted for clarity [see Fig. 4(*b*)]. The vertical bilayer columns in panel (*a*) correspond to sections through such distributions. Panels (*c*)–(*e*) show the corresponding calculated diffraction patterns for this model: (*c*) *h*1*l*, (*d*) *hk*0 and (*e*) *hk*1.

between the domains in the intermediate phases even further, as illustrated in Fig. 10(*c*).

Before analysing the best-fit model, it is imperative to note that, according to Figs. 1(*a*) and 1(*b*), there are two important packing principles in the plane shown that are pertinent to both the α and the β phases:

(i) In the horizontal direction every single molecule forms a homodimer with another molecule of the same chirality.

(ii) In the vertical direction the neighbours above and below any molecule have opposite chirality to it.

The simulation cell is a perfect rigid lattice with no dislocations other than regular slips converting an *A* to a *B* and *vice versa*, and has, in particular, no voids or vacancies as indicated by black boxes in Fig. 10(*c*). An important implication of these restrictions is that any crystal defect must come about by breaking one or both of the above principles.

Fig. 11 shows a detailed view of the best-fit model shown on a smaller scale in Fig. 7(*i*). Here only part of a single layer of the 48 × 48 × 48 model crystal is included, but the observations are representative of (but not repetitive in) the full crystal. The two grey boxes labelled (1) show the normal packing mode, and in the enlarged segment to the right the red ellipses surround two sets of homodimers.

White boxes highlight defects where principle (ii) is broken but not (i). Depending on the surroundings, such defects give rise to phase mixing of two types, (2) and (3). Both appear in two versions with respect to α and β phases (with inverted yellow and blue colours), but while motif (2) shown at the top right of Fig. 11 occurs with about the same frequency as its alternative (8:10), all drawn with dashed lines, there is a 16:5 imbalance between the version of (3) drawn with solid lines *versus* its alternative drawn with dotted lines. The significance (if any) of this observation is not clear. It is likely that this is just a statistical fluctuation in this relatively small sample, away from the average value for the whole crystal.

There are two identical examples of three consecutive homodimers along the vertical axis all having the same chirality. These have been identified by green boxes (4).

When principle (i) is broken, it follows that so too is (ii). Black boxes surround occurrences in Fig. 11. The introduction of a single heterodimer (14 cases) may either introduce a blue speck in a yellow region (or the opposite) or lead to a more substantial change in packing. Examples of both situations are labelled (5) in the illustration. There are nine examples of pairs of heterodimers (6), five of them leading to the type of defect shown enlarged at the bottom right. Two nearby boxes (7) have a series of three heterodimers, while four heterodimers occur once (8) (not shown in detail).

Layer translations or slips in a real crystal certainly do not have the power to change the chiralities of the molecules involved. The formation of perfect heterodimers or interactions between homochiral molecules along the vertical axis in our simulation model should thus be seen to represent slightly different defects in a real crystal, such as imperfect heterodimers, misaligned homochiral neighbours and layers, vacancies [as the black boxes in Fig. 10(*c*) could represent], and reorientations of one or more molecules to form non-standard hydrogen-bond interactions. Even with these caveats, the exceptionally good fit between the diffuse scattering predicted by the model and the experiment provides valuable insight, not only into the types of slip, which obviously occur in a non-systematic manner with examples of all the distinct models described in Fig. 10, but also into the density of defects along all unit-cell axes. Prior to this investigation, the domain sizes of mixed-phase crystals were unknown. From our simulations, it appears that defects may occur with very high density during phase transition, typically separated by only 1–20 unit cells, rather than hundreds or thousands. This truly emphasizes the dynamic nature of the α-to-β phase transition for DL-norleucine and the plasticity of the crystalline network in order to accommodate a large number of faults, bilayer slips and regular defects.

It should be noted that the simulations carried out for this work were somewhat restricted, both in the size of the simulation crystal that could be used (48 × 48 × 48 unit cells) and also in the number of MC cycles that were carried out for each example (typically ∼1000). It is expected that increasing these attributes could lead to model crystals containing rather fewer of the defects described above and give bilayers with larger domain sizes. However, it is considered that neither of these would materially affect the general observations made.

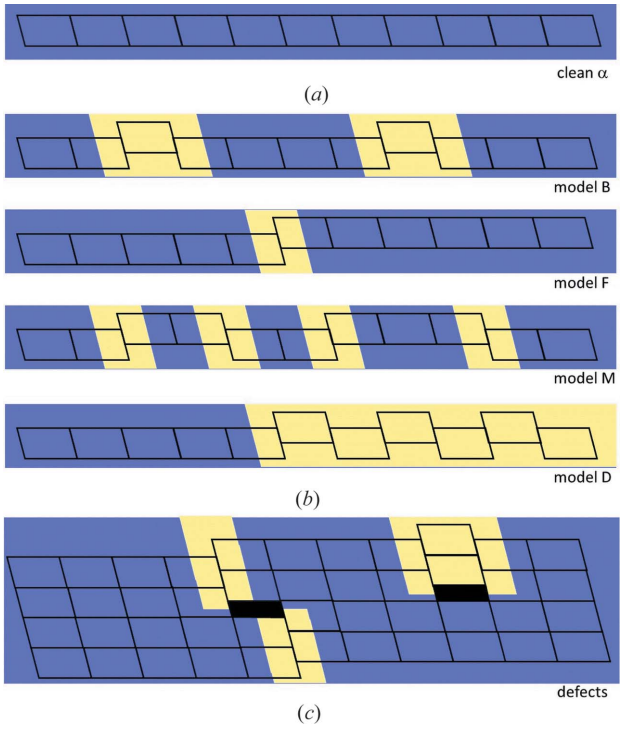


Figure 10
(*a*) A display of the clean α phase, with parallelograms used as in Fig. 1. (*b*) The principally distinct ways that slipping translations can occur. Model *B* shows translation of individual bilayers, model *F* shows a fault-type movement, model *M* shows coordinated movement of multiple bilayers, here shown as one pair and one triplet, and model *D* shows coherent translations that result in two domains with a clear division between phases. (*c*) The effect of the introduction of defects along the short, almost vertical, axis on phase designation. See text for details.

8. Conclusion

Monte Carlo simulation of a model structure has been used to investigate the disorder and associated diffuse X-ray scattering from a crystal of DL-norleucine. The diffuse scattering data used were obtained from CCD data frames measured in a conventional crystal structure determination. Though diffuse scattering data obtained in this way are not of the same quality as could be obtained in the best experiments using synchrotron radiation (see *e.g.* Chan *et al.*, 2010), they nevertheless provide sufficient detail to allow the nature of the disorder to be probed by MC methods.

The MC model was initially set up in very general terms where each site on a $P2_1/c$ lattice contained a binary random variable that defined the occupancy of that site. Correlations between the occupancies were induced via a set of interactions acting along all near-neighbour intermolecular vectors (as shown in Fig. 5). This allowed different ordering schemes to be explored, several of which resulted in disordered structures with diffraction patterns displaying diffuse streaks extending in the a^* direction. Of these, the example shown in Fig. 9 is considered to provide the best fit to the observations. The model consists of molecular bilayers, each containing the hydrophilic interface at its centre. Each such bilayer is independent of neighbouring bilayers, which may or may not have slipped relative to it as a result of very weak interaction across the hydrophobic interface.

Very large values for just two (J_R and J_O) of the eight different intermolecular interactions depicted in Fig. 5 are primarily responsible for the formation of the structure. The large value of J_R ensures that the molecules pack in pairs across the hydrophilic interface and J_P (together with J_O) promotes the stacking of these pairs into bilayers that extend a considerable distance in the lateral directions. The width of the diffraction streaks depends on the range of order within these *bc* bilayers. It is clear that the finite size of the simulation ($48 \times 48 \times 48$ unit cells) limits the range of order than can be achieved, and a larger size could allow larger domains to form with consequent narrowing of the diffuse streaks. However, the present model gives a reasonable approximation to the widths of peaks in the observed diffraction patterns. If much better experimental patterns were available, a larger simulation crystal might be needed to allow more detail to be extracted.

The zero value of J_L means that successive bilayers are independent. Even a small correlation (either positive or

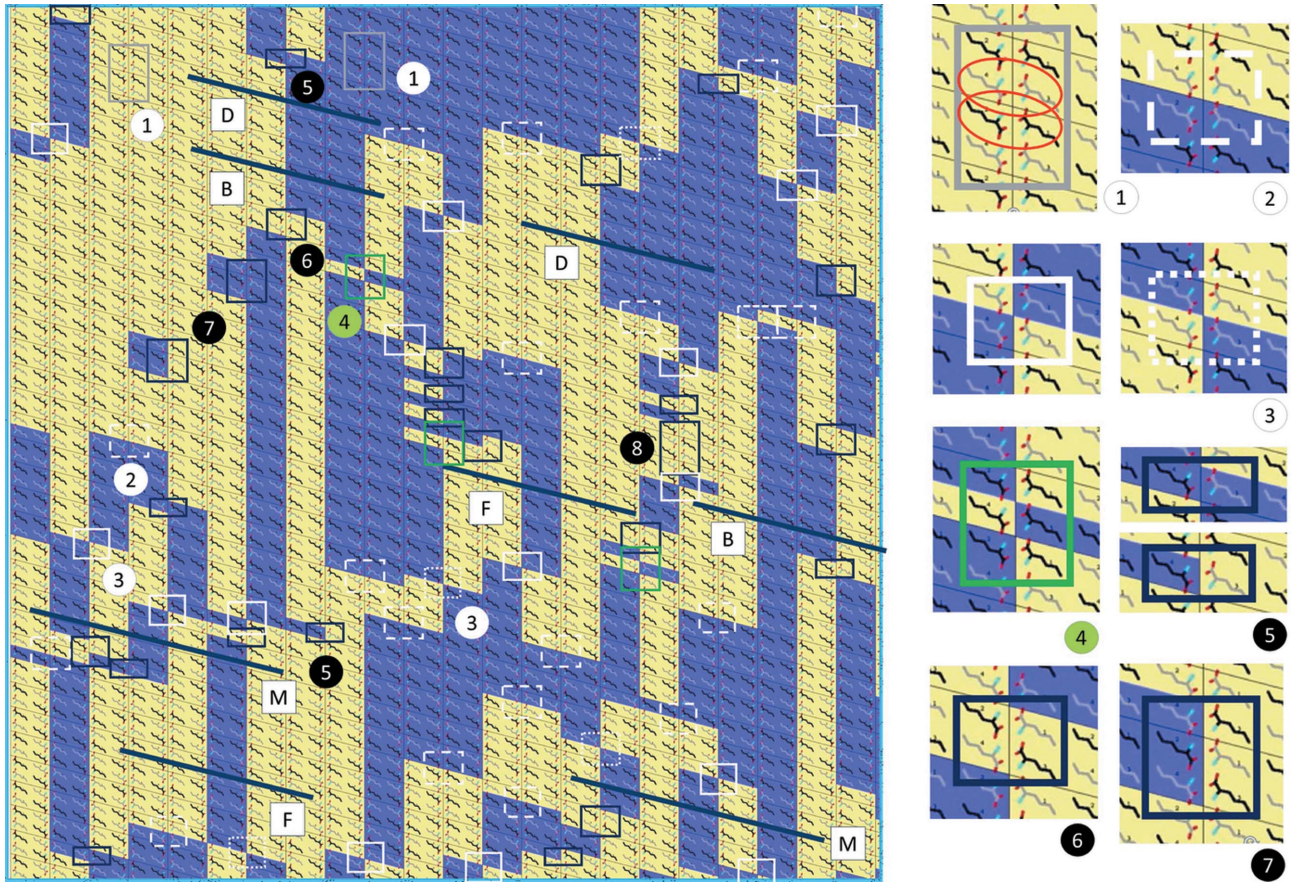


Figure 11
A detailed version of Fig. 7(*i*). Prototypic slipping translations are identified by black lines and letters corresponding to the codes in Fig. 10(*b*). Segments of the main illustration, labelled (1)–(7), are enlarged on the right-hand side. The numbers highlight the different types of defect discussed in the text. As before, L- and D-enantiomers can be distinguished by having light-grey and dark C atoms, respectively.

negative) induced by J_L results in significant changes to the intensity distribution along the diffuse streaks. With $J_L = 0$ the intensity variation along the streaks is entirely due to the molecular structure factor. It should be noted here that each molecule is considered to be a rigid body and is fixed to the average site position. Consequently, the calculated diffraction contains only information about the occupancy disorder, and omits any details that might be expected to occur in a real crystal due to variations in the molecular geometry, local variations in the orientation and position of molecules due to thermal fluctuations, and local size-effect relaxations around defects. Such detailed information would require much better diffraction data, as was available for example in the study of aspirin (Chan *et al.*, 2010) which also involved stacking faults.

Perhaps the most useful outcome of the present work is the insight that the MC model gives into the possible nature of the local structure during the α-to-β phase transition. This was achieved by identifying in the model crystals regions that conformed locally to the α or the β structures. As described in Section 7 the model provides valuable insight, not only into the types of molecular slip that must occur, but also into the densities of defects that occur along all unit-cell axes, typically separated by only 1–20 unit cells, rather than hundreds or thousands. This emphasizes the dynamic nature of the α-to-β phase transition for DL-norleucine and the plasticity of the crystalline network in order to accommodate a large number of faults, bilayer slips and regular defects.

Acknowledgements

The authors would like to thank Dr Thomas Weber for his help with processing the X-ray data using *CrysAlisPRO*.

References

Butler, B. D. & Welberry, T. R. (1992). *J. Appl. Cryst.* **25**, 391–399.
Chan, E. J., Welberry, T. R., Heerdegen, A. P. & Goossens, D. J. (2010). *Acta Cryst.* B**66**, 696–707.
Coles, S. J., Gelbrich, T., Griesser, U. J., Hursthouse, M. B., Pitak, M. & Threlfall, T. (2009). *Cryst. Growth Des.* **9**, 4610–4612.
Czech, C., Glinnemann, J., Johansson, K. E., Bolte, M. & Schmidt, M. U. (2017). *Acta Cryst.* B**73**, 1075–1084.
Dalhus, B. & Görbitz, C. H. (1996). *Acta Cryst.* C**52**, 1761–1764.
Ende, J. A. van den, Ensing, B. & Cuppen, H. M. (2016). *CrystEngComm*, **18**, 4420–4430.
Ende, J. A. van den, Smets, M. M. H., de Jong, D. T., Brugman, S. J. T., Ensing, B., Tinnemans, P. T., Meekes, H. & Cuppen, H. M. (2015). *Faraday Discuss.* **179**, 421–436.
Görbitz, C. H. (2011). *J. Phys. Chem. B*, **115**, 2447–2453.
Görbitz, C. H. & Karen, P. (2015). *J. Phys. Chem. B*, **119**, 4975–4984.
Groom, C. R., Bruno, I. J., Lightfoot, M. P. & Ward, S. C. (2016). *Acta Cryst.* B**72**, 171–179.
Harding, M. M., Kariuki, B. M., Williams, L. & Anwar, J. (1995). *Acta Cryst.* B**51**, 1059–1062.
Macrae, C. F., Bruno, I. J., Chisholm, J. A., Edgington, P. R., McCabe, P., Pidcock, E., Rodriguez-Monge, L., Taylor, R., van de Streek, J. & Wood, P. A. (2008). *J. Appl. Cryst.* **41**, 466–470.
Mathieson, A. McL. (1952). *Acta Cryst.* **5**, 332–341.
Mathieson, A. McL. (1953). *Acta Cryst.* **6**, 399–403.
Mnyukh, Y., Panfilova, N. A., Petropavlov, N. N. & Uchvatova, N. S. (1975). *J. Phys. Chem. Solids*, **36**, 127–144.
Smets, M. M. H., Brugman, S. J. T., van Eck, E. R. H., van den Ende, J. A., Meekes, H. & Cuppen, H. M. (2015). *Cryst. Growth Des.* **15**, 5157–5167.
Weber, T. & Bürgi, H.-B. (2002). *Acta Cryst.* A**58**, 526–540.
Weber, T., Estermann, M. A. & Bürgi, H.-B. (2001). *Acta Cryst.* B**57**, 579–590.
Welberry, T. R. (1985). *Rep. Prog. Phys.* **48**, 1543–1594.
Welberry, T. R. (2004). *Diffuse X-ray Scattering and Models of Disorder*. *IUCr Monographs on Crystallography*, Vol. 16. Oxford University Press.
Welberry, T. R. & Goossens, D. J. (2008). *Acta Cryst.* A**64**, 23–32.