

- FRASER, R. D. B., MACRAE, T. P. & SUZUKI, E. (1978). *J. Appl. Cryst.* **11**, 693–694.
- GUINER, M. A. (1939). *Ann. Phys. (Leipzig)*, **12**, 161–237.
- HENDRICKS, S. & TELLER, E. (1942). *J. Chem. Phys.* **10**, 147–167.
- HOLMES, K. C. & BARRINGTON LEIGH, J. (1974). *Acta Cryst.* **A30**, 635–638.
- HOSEMANN, R. & BAGCHI, S. N. (1962). *Direct Analysis of Diffraction by Matter*. Amsterdam: North-Holland.
- INO, T. & MINAMI, N. (1979). *Acta Cryst.* **A35**, 163–170.
- INOUE, H. (1994). *Acta Cryst.* **A50**, 644–646.
- INOUE, H., FRASER, P. E. & KIRSCHNER, D. A. (1993). *Biophys. J.* **64**, 502–519.
- KLUG, A., CRICK, F. H. C. & WYCKOFF, H. W. (1958). *Acta Cryst.* **11**, 199–212.
- KLUG, A. & FRANKLIN, R. E. (1958). *Discuss. Faraday Soc.* **25**, 104–110.
- MAKOWSKI, L. (1978). *J. Appl. Cryst.* **11**, 273–283.
- MAKOWSKI, L. (1982). *J. Appl. Cryst.* **15**, 546–557.
- MARVIN, D. A., SPENCER, M., WILKINS, M. H. F. & HAMILTON, L. D. (1961). *J. Mol. Biol.* **3**, 547–565.
- MILLANE, R. P. (1988). In *Crystallographic Computing 4: Techniques and New Technologies*, edited by N. W. ISAACS & M. R. TAYLOR, pp. 169–186. Oxford Univ. Press.
- MILLANE, R. P. & ARNOTT, S. (1985). *J. Appl. Cryst.* **18**, 419–423.
- MILLANE, R. P. & ARNOTT, S. (1986). *J. Macromol. Sci. Phys.* **B24**, 193–227.
- MILLANE, R. P., CHANDRASEKARAN, R., ARNOTT, S. & DEA, I. C. M. (1988). *Carbohdr. Res.* **182**, 1–17.
- MILLANE, R. P. & STROUD, W. J. (1991). *Int. J. Biol. Macromol.* **13**, 202–208.
- MILLER, A. & PARRY, D. A. D. (1974). *Polymer* **15**, 706–712.
- NAMBA, K. & STUBBS, G. J. (1985). *Acta Cryst.* **A41**, 252–262.
- OKUYAMA, K., ARNOTT, S., MOORHOUSE, R., WALKINSHAW, M., ATKINS, E. D. T. & WOLF-ULLISH, C. (1980). *Fiber Diffraction Methods*, edited by A. D. FRENCH & K. H. GARDNER, pp. 409–427. ACS Symposium Series, Vol 141. Washington: American Chemical Society.
- PAPOULIS, A. (1984). *Probability, Random Variables, and Stochastic Processes*, 2nd ed. New York: McGraw-Hill.
- PARK, H., ARNOTT, S., CHANDRASEKARAN, R., MILLANE, R. P. & CAMPAGNARI, F. (1987). *J. Mol. Biol.* **197**, 513–523.
- RUTLEDGE, G. C. & SUTER, U. W. (1991). *Macromolecules*, **24**, 1921–1933.
- STROUD, W. J. (1993). PhD thesis, Purdue Univ., West Lafayette IN 47907, USA.
- STROUD, W. J. & MILLANE, R. P. (1995a). *Acta Cryst.* **A51**. In the press.
- STROUD, W. J. & MILLANE, R. P. (1995b). *Proc. R. Soc. London Ser. A*. Submitted.
- STUBBS, G. J. (1974). *Acta Cryst.* **A30**, 639–645.
- TANAKA, S. & NAYA, S. (1969). *J. Phys. Soc. Jpn.* **26**, 982–993.
- VAINSHTEIN, B. K. (1966). *Diffraction of X-rays by Chain Molecules*. Amsterdam: Elsevier.
- WELBERRY, T. R. (1985). *Rep. Prog. Phys.* **48**, 1543–1593.
- WILSON, A. J. C. (1942). *Proc. R. Soc. London Ser. A*, **180**, 277–285.
- WORTHINGTON, C. R. & ELLIOT, G. F. (1989). *Acta Cryst.* **A45**, 645–654.

Acta Cryst. (1995). **A51**, 790–800

Analysis of Disorder in Biopolymer Fibers

BY W. J. STROUD AND R. P. MILLANE*

Whistler Center for Carbohydrate Research, Purdue University, West Lafayette, Indiana 47907-1160, USA

(Received 5 September 1994; accepted 15 May 1995)

Abstract

X-ray diffraction patterns from oriented polycrystalline fibers of some biopolymers show that the molecules are disordered within the microcrystallites. Quantifying the disorder in such specimens is a necessary step for the use of their diffraction patterns for accurate structure determination. Theory and algorithms for calculating diffraction patterns from such fibers have recently been described [Stroud & Millane (1995). *Acta Cryst.* **A51**, 000–000]. Here the application of these methods to determining the kind and degree of disorder in two polynucleotide fibers is described. The more ordered system shows random screw disorder accompanied by small lattice distortions, and the more disordered system shows larger lattice distortions and significant rotational disorder. These results show the potential of these methods for determining disorder in polycrystalline fibers;

uniqueness of the solutions and implications for structure determination are discussed.

1. Introduction

The molecular and crystal structures of many of biopolymers and rod-like macromolecular assemblies have been determined by X-ray fiber diffraction analysis (Arnott, 1980; Millane, 1988). In most cases, diffraction data from *polycrystalline* specimens, which are made up of small crystallites that are randomly rotated relative to each other, have been used for structural analysis (Arnott, 1980; Millane, 1988). The diffraction patterns from these specimens are equivalent to the cylindrical projection of the pattern from a single crystal and are used to determine full crystal structures. In other cases, structure determination has used diffraction data from *non-crystalline* fibers, in which the molecules are merely oriented but not otherwise organized in the speci-

* To whom all correspondence should be addressed.

men (Namba & Stubbs, 1985; Millane, Chandrasekaran, Arnott & Dea, 1988). The data are equivalent to the cylindrical average of the intensity diffracted by a single molecule and can be used to determine the structure of an isolated molecule.

In some cases, however, the best specimen that can be obtained gives a diffraction pattern indicative of ordering that is intermediate between the extremes present in polycrystalline and non-crystalline fibers (Miller & Parry, 1974; Arnott, 1980). The pattern may either contain sharp Bragg reflections at low resolution that give way to continuous layer-line intensities at high resolution or contain Bragg and diffuse diffraction dispersed over its entirety. In either case, the specimen is essentially polycrystalline but the packing of the molecules in the crystallites is disordered in some way. The diffracted intensity then depends not only on the molecular and crystal structures but also on the kind and degree of disorder within the crystallites. Since structure determination essentially involves finding a model structure whose calculated diffraction amplitudes match those that are measured, methods for calculating the diffraction from an appropriate model of a disordered polycrystalline fiber are necessary for accurate structure determination using data from these kinds of specimens.

In a few cases, where the disorder is simple and easily characterized (usually involving two distinct positions or orientations of the molecules), the effects of disorder have been incorporated into the determination of biopolymer structures (Miller & Parry, 1974; Arnott & Selsing, 1974). However, in other cases where the disorder is more complicated and not as easily characterized, structures have been determined without taking the disorder into account (Arnott, Chandrasekaran, Millane & Park, 1986; Park, Arnott, Chandrasekaran, Millane & Campagnari, 1987). In these latter cases, the higher-resolution data were treated as being due to a non-crystalline specimen and the low-resolution Bragg data as coming from a perfect polycrystalline specimen. Since disorder affects both of these components, a straightforward approach such as this may introduce errors into the structures so determined.

We have recently reported a statistical model of disordered polycrystalline fibers together with theory and computational procedures for calculating fiber diffraction patterns from the model (Stroud & Millane, 1995). Simulations using these methods show how various parameters describing disorder can be related to different features on diffraction patterns (Stroud & Millane, 1995). Here we describe the application of these methods to determining the kind and degree of disorder in two polynucleotide fibers. The various disorder parameters in models are adjusted so that the features on the calculated patterns match those of the observed patterns. This allows the primary forms of disorder in the fibers to be identified and quantified.

2. Methods

2.1. Diffraction data

Diffraction patterns were scanned on an Optronics Photoscan P-1000 rotating drum microdensitometer, with a raster spacing of 50 μm . The center of each pattern, the orientation of the pattern relative to the raster, the film-to-fiber distance and the tilt of the fiber relative to the incident X-ray beam were determined using standard methods (Fraser, MacRae, Miller & Rowlands, 1976; Millane & Arnott, 1986). Intensities measured at positions between the recorded layer lines were used to calculate a two-dimensional global background [expanded as a Fourier-Bessel series (Millane & Arnott, 1985)], which was subtracted from the diffraction pattern. The background-corrected patterns were mapped to reciprocal space and averaged over the four quadrants. Layer-line amplitudes were obtained as traces of the amplitude along the centers of layer lines.

2.2. Calculation of diffraction patterns from models of disordered specimens

Our model of disordered polycrystalline fibers and the theory of diffraction from these fibers are described in detail by Stroud & Millane (1995). Fourier-Bessel structure factors were calculated [using water-weighted atomic scattering factors (Fraser, MacRae & Suzuki, 1978)] from the atomic coordinates of the structures previously determined from the diffraction patterns under study. As described above, these coordinates cannot be expected to be exactly correct since the procedure used to determine them did not take into account the effects of disorder. Assuming, however, that any errors in the coordinates are small, their effect will be to introduce small changes into the calculated Fourier-Bessel structure factors. This will influence the amplitudes of the Bragg and continuous components but not the distribution of these components in different regions of the diffraction pattern, which is mainly affected by disorder. Small errors in the atomic coordinates are not, therefore, expected to significantly affect the analysis of packing disorder.

Layer-line amplitudes were calculated for ideal specimens in which the constituent disordered crystallites are perfectly oriented and are composed of structurally regular molecules of infinite length. Full two-dimensional diffraction patterns were obtained from the layer-line amplitudes by first broadening them with a Gaussian profile to account for the effect of the coherence length, l_c , of the molecules (Stubbs, 1974). The broadened layer lines were then substituted into the disorientation integral of Holmes & Barrington Leigh (1974), which was evaluated for a normal distribution of crystallite orientations with variance α_0^2 . The resulting two-dimensional pattern was then convolved with a circularly symmetric

Gaussian function with standard deviation $8 \times 10^{-4} \text{ \AA}^{-1}$ (corresponding, approximately, to a Gaussian beam profile with a standard deviation of $25 \mu\text{m}$ on the film). Traces were made through the centers of the layer lines of the final patterns for comparison with the corresponding traces made from the measured patterns. These traces show more detail than do continuous tone representations of a whole diffraction pattern.

Our overall approach to quantifying disorder in polycrystalline fibers is as follows. Crystallite size, coherence length and the disorientation angle (which are independent of the exact kind of disorder present) are first estimated by matching the lateral and angular profiles of discrete reflections on the calculated and measured diffraction patterns. The lateral profile of a reflection is the profile of a cut through the center of the reflection made along the layer line, while the angular profile is that measured along an arc centered on the origin of reciprocal space. The calculated profiles are scaled to the corresponding observed profiles so as to remove the influence of the particular molecular models used in their calculation and the relevant disorder parameters adjusted to minimize the r.m.s. error between the two sets of profiles.

Layer-line amplitudes are then calculated using these parameters and compared with the measured amplitudes. Discrepancies between the two sets of amplitudes are used to infer a component of disorder that is present, and a series of diffraction patterns are calculated to estimate parameters describing this component. Residual discrepancies are then used to identify further components of the disorder and the process repeated until the

best fit with data is obtained. Since we are interested in the uniqueness of solutions, different possible kinds of disorder were investigated in all cases. The exact procedure used depends on the problem at hand and is described in detail for each diffraction pattern studied.

3. A DNA duplex

A fiber diffraction pattern from the α form of the polynucleotide duplex poly(dA) · poly(dT) is shown in Fig. 1 (Arnott & Selsing, 1974). The pattern is dominated by Bragg reflections at resolutions less than 5 \AA , which give way to continuous diffraction at higher resolution. This diffraction pattern was first analyzed by Arnott & Selsing (1974), who determined the structure using 20 of the low-resolution Bragg reflections. The structure was re-determined in more detail by Park, Arnott, Chandrasekaran, Millane & Campagnari (1987) using a larger data set consisting of the continuous layer-line amplitudes measured in the region $0.18 \lesssim \rho \lesssim 0.32 \text{ \AA}^{-1}$ and the amplitudes of 33 Bragg reflections in the region $\rho \lesssim 0.18 \text{ \AA}^{-1}$, ρ being the distance from the origin of reciprocal space. The atomic coordinates from the later study are used in the diffraction calculations presented here. The average unit cell is monoclinic with $a = b = 23.2$, $c = 32.3 \text{ \AA}$ and $\gamma = 120^\circ$, a single molecule with 10_1 helix symmetry passes through the unit cell, and the space group is $P1$.

For the current study, the diffraction pattern in Fig. 1 was processed as described in §2. Traces of the diffracted amplitude along the centers of the layer lines of the processed pattern are shown in Fig. 2. A series of

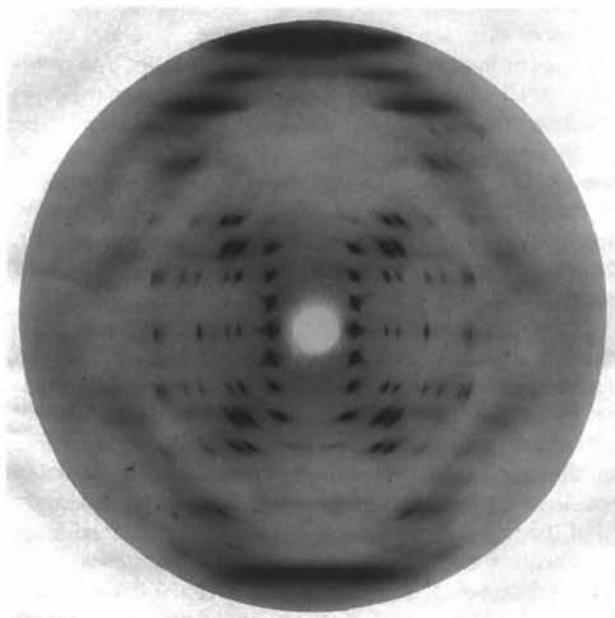


Fig. 1. Fiber diffraction pattern from α -poly(dA) · poly(dT) (Arnott & Selsing, 1974).

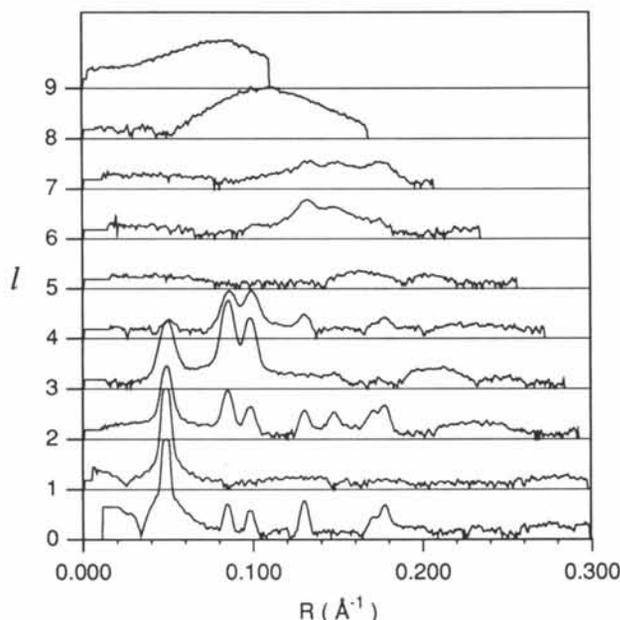


Fig. 2. Amplitudes along the centers of the layer lines of the diffraction pattern from α -poly(dA) · poly(dT).

diffraction patterns for a polycrystalline structure with no disorder were calculated and the crystallite radius r_c , coherence length l_c and disorientation angle α_0 adjusted to obtain the best match of the calculated angular and lateral profiles of the Bragg reflections to those observed. The estimate for r_c was obtained by matching the lateral profiles of equatorial reflections, while estimates of l_c and α_0 were obtained by matching the angular profiles of all of the strong Bragg reflections on the pattern. This gave the estimates $r_c = 165 \text{ \AA}$, $\alpha_0 = 1.8^\circ$ and $l_c = 175 \text{ \AA}$. The diffraction pattern calculated for an

ideally polycrystalline fiber with these values is shown in Fig. 3(a).

As is evident in Fig. 3(a), the effect of disorientation is to broaden reflections on the upper layer lines, particularly those closer to the meridian. This removes much of the structure from the upper layer lines of the diffraction pattern. Comparison of Figs. 2 and 3(a) shows that, despite this, the calculated pattern still displays too much Bragg structure at high resolution. This indicates that there is significant disorder in the specimen. The layer-line traces in Fig. 2 show Bragg reflections on layer lines

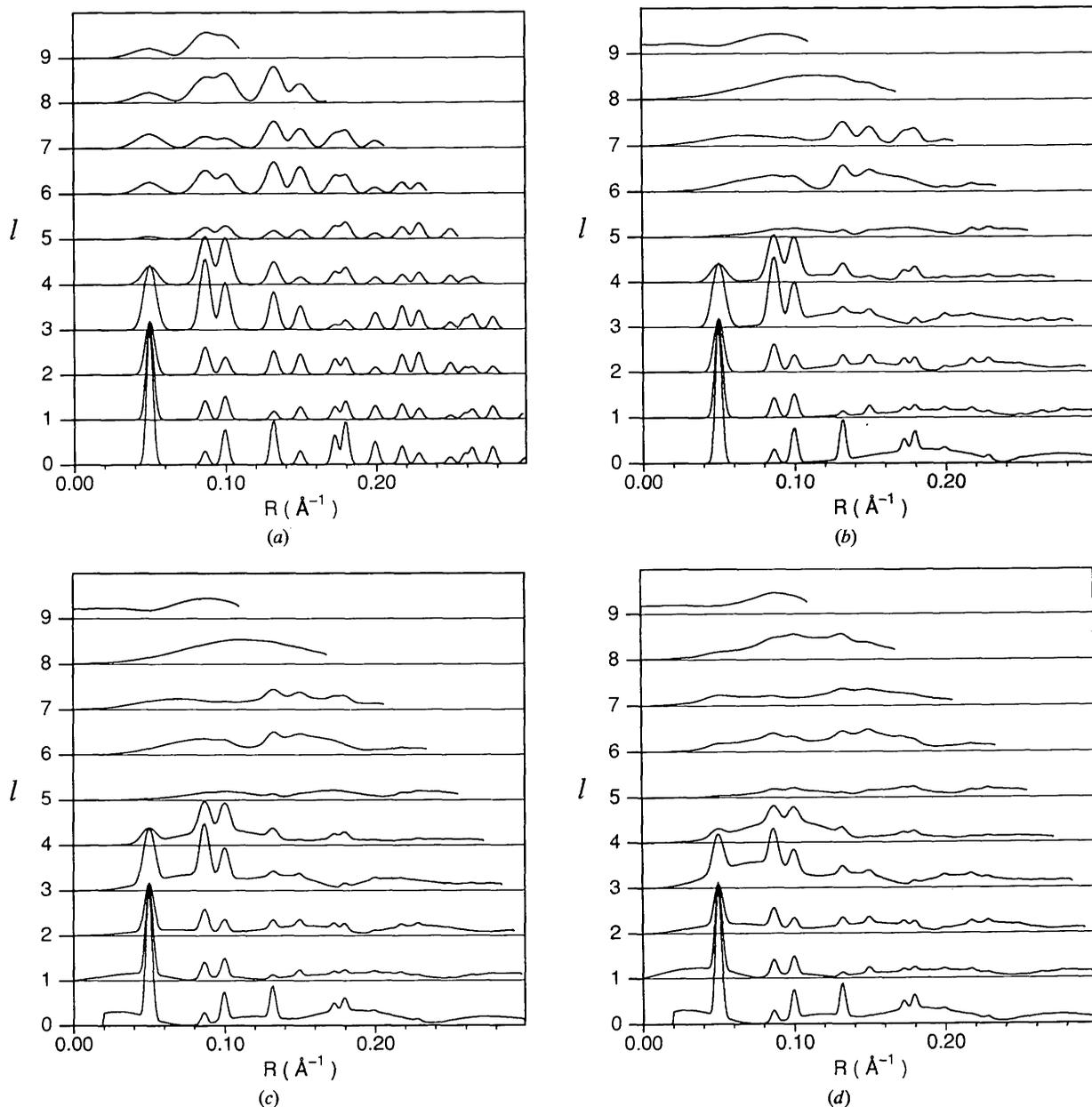


Fig. 3. Calculated layer-line amplitudes for polycrystalline models of α -poly(dA)·poly(dT) with (a) no disorder, (b) random screw disorder, (c) random screw disorder and lattice distortions with $\sigma_{\text{lat}} = 0.6 \text{ \AA}$ and $\sigma_{\text{axial}} = 0.5 \text{ \AA}$ and (d) small independent rotations and translations of the molecules with $\sigma_{\phi} = 5^\circ$ and $\sigma_{\text{axial}} = 0.7 \text{ \AA}^{-1}$ and lateral disorder with $\sigma_{\text{lat}} = 0.6 \text{ \AA}^{-1}$.

0 through 4 and vestiges of reflections on layer lines 6 and 7. The absence of Bragg interference for reciprocal-space cylindrical polar radii $R \lesssim 0.1 \text{ \AA}^{-1}$ on layer lines 6 and 7, but its presence for larger values of R , is evidence for screw disorder (Stroud & Millane, 1995). Introducing random screw disorder into the model gives the diffraction pattern shown in Fig. 3(b). Overall, the calculated pattern agrees rather well with the observed pattern but differs from it in the following respects: (1) there is too much Bragg structure for $0.12 \lesssim R \lesssim 0.18 \text{ \AA}^{-1}$ on layer lines 6 and 7; (2) there is too much Bragg structure for $R \gtrsim 0.2 \text{ \AA}^{-1}$ on the lower layer lines; and (3) the two Bragg peaks on layer line 1 at $R \simeq 0.08 \text{ \AA}^{-1}$ and $R \simeq 0.10 \text{ \AA}^{-1}$ are absent from the observed pattern. Discrepancies 1 and 2 can be addressed by introducing small amounts of lateral and axial disorder into the model.

Lateral disorder reduces the intensity of Bragg reflections with increasing R , while axial disorder reduces the intensity of the reflections with increasing distance, Z , from the equator. The weighting that applies to a reflection is

$$w(R, l) = \exp[-4\pi^2(R^2\sigma_{\text{lat}}^2 + l^2\sigma_{\text{axial}}^2/c^2)], \quad (1)$$

where σ_{lat}^2 and σ_{axial}^2 are, respectively, the variances of small normally distributed lateral and axial displacements of the molecules away from their ideal positions in a crystallite (Stroud & Millane, 1995). An effective strategy for estimating σ_{lat} and σ_{axial} is to set one of these parameters to zero and vary the other until the calculated pattern reproduces a particular feature of the observed pattern. A second feature of the observed pattern is then reproduced by varying the parameters so that the weight given by (1) remains approximately constant in the region occupied by the first feature. In the current example, setting $\sigma_{\text{axial}} = 0.7 \text{ \AA}$ reduces the Bragg intensity on layer lines $l = 5$ and above. Holding $w(R, Z)$ constant for $R \simeq 0.14 \text{ \AA}^{-1}$ and $l = 6$ and reducing σ_{axial} to 0.5 \AA while increasing σ_{lat} to 0.6 \AA keeps the residual Bragg structure on layer line $l = 6$ but eliminates the undesirable structure at large R on the lower layer lines, introduces more continuous intensity at small values of R on these layers and results in the best fit (Fig. 3c) to the observed pattern. The overall match of this pattern to that observed is now quite good, aside from some weak residual Bragg structure at high resolution, with discrepancies 1 and 2 noted above taken care of. Discrepancy 3, the presence of Bragg reflections for $R \simeq 0.08$ and 0.10 \AA^{-1} on layer line 1 on the calculated pattern, remains however.

The maximum radius of α -poly(dA)·poly(dT) is such that, on layer line $l = 1$, only the first-order Fourier-Bessel structure factor contributes to the diffracted intensity for $R \lesssim 0.13 \text{ \AA}^{-1}$. Removing the reflections at $R \simeq 0.08$ and $R = 0.10 \text{ \AA}^{-1}$ by introducing additional rotation and lattice disorder into

the model also attenuates the reflection at $R \simeq 0.04 \text{ \AA}^{-1}$ and produces excessive continuous intensity at low resolution. Attempts to reduce the amplitudes of these reflections must, therefore, consider other forms of disorder.

The systematic absence of reflections from particular layer lines and particular values of R can result from discrete rotations and/or axial translations of the molecules (Stroud & Millane, 1995). However, calculations for models in which the molecule in each unit cell could occupy either of two discrete positions, made over the full range of possible rotations and axial translations, showed that reflections could not be removed from layer line 1 without producing serious discrepancies between the calculated and observed patterns elsewhere. In fact, since there are no absent reflections for $R \lesssim 0.12 \text{ \AA}^{-1}$ on layer lines 0, 2, 3 and 4 of the observed pattern, it is likely that the extra reflections on layer line 1 of the calculated pattern are due not to shortcomings in the model of disorder but rather to an overestimate (as a result of small errors in the molecular model or of associated ordered solvent) of the molecular transform on the first layer line in this region. Errors in the molecular transform could also account for the reflection at $R \simeq 0.18 \text{ \AA}^{-1}$ on layer line 2 of the calculated pattern being somewhat weaker than observed.

The model incorporating random screw disorder with small axial and lateral shifts gives, therefore, a good match to the observed diffraction pattern. However, this solution cannot be considered unique without considering alternative kinds of disorder in which the screw displacements of the molecules are not completely random (*i.e.* small screw displacements) or the axial displacements of the molecules are not correlated to their rotations (*i.e.* independent small rotations and axial translations) or the molecules are directionally disordered.

To investigate the first alternative, we kept the lattice-distortion parameters fixed at the values determined above and calculated diffraction patterns for models with normally distributed screw disorder for a range of standard deviations σ_φ (Stroud & Millane, 1995). For $\sigma_\varphi > 10^\circ$, there is little change to the diffraction pattern compared to that for a random screw because aliasing of the distribution function of molecular positions for a normally distributed screw leads to a distribution rather similar to that for random screw disorder (Stroud & Millane, 1995). Reducing σ_φ below 10° modifies the pattern but introduces additional Bragg intensity close to the meridian on the upper layer lines, in contrast to observation. These Bragg reflections can be removed by increasing the axial disorder but this then attenuates Bragg reflections on the lower layer lines too much. Small screw disorder ($\sigma_\varphi < 10^\circ$) is therefore not consistent with the observed diffraction pattern.

The second alternative model involves small independent rotations and translations of the molecules. The effect of small rotations is to reduce, with increasing

Bessel order, the contribution of the Fourier–Bessel structure factors to the amplitude of the Bragg reflections. The location of the reflections most affected depends on which Bessel orders satisfy the helix selection rule on the different layer lines. For a molecule with 10_1 helix symmetry, the reflections most affected are those on layer lines $l = 4, 5$ and 6 and those at large values of R on the remaining layer lines. Small rotations with a normal distribution and standard deviation σ_φ were introduced into a model with no lattice disorder. Rotations with $\sigma_\varphi = 5^\circ$ remove some Bragg intensity from layer lines $l = 5, 6$ and 7 and weaken the Bragg reflections on layer lines $l = 0$ and 1 for $R \gtrsim 0.14 \text{ \AA}^{-1}$. Larger rotations suppress the Bragg reflections on layer line $l = 4$ too much and introduce excessive diffuse intensity into this region. They are not, therefore, consistent with observation. Small rotations, by themselves, do not give a satisfactory match of the calculated and the observed pattern. Lattice disorder is needed to remove excess Bragg intensity at high resolution. Repeating the procedure used to determine the lattice disorder parameters in the case of random screw disorder yielded the values $\sigma_{\text{axial}} = 0.7 \text{ \AA}$ and $\sigma_{\text{lat}} = 0.6 \text{ \AA}$ for $\sigma_\varphi = 5^\circ$. The corresponding diffraction pattern is shown in Fig. 3(d). The main differences between Figs. 3(c) and (d) are that, in the latter, the Bragg structure on layer lines 6 and 7 is suppressed at $R \simeq 0.13 \text{ \AA}^{-1}$, and is slightly enhanced on layer lines 6 and 7 for $R \simeq 0.04\text{--}0.09 \text{ \AA}^{-1}$ and throughout layer line 8, relative to the former. This appears to be at odds with the data, although the differences are rather small. Increasing the lateral disorder to reduce the intensity of these reflections reduces the intensity of reflections at smaller R and introduces excess diffuse intensity in this region and on layer lines 0 to 4 for $R \lesssim 0.1 \text{ \AA}^{-1}$. Small rotations and translations do not, therefore, appear to fit the data quite as well as random screw disorder.

The final alternative model is one that includes directional disorder (Stroud & Millane, 1995). Directional disorder was introduced into the models with random screw disorder, and small rotations and translations already considered, and diffraction patterns calculated for the full range of the parameters (φ_0, z_0) that describe the relationship between the ‘up’ and ‘down’ molecules (Stroud & Millane, 1995). For most values of (φ_0, z_0) , the calculated pattern changed considerably and no longer matched the data. However, for a narrow range of values of these parameters, there were only small changes to the resulting diffraction patterns. These values were $(\varphi_0, z_0) = (0^\circ, 0.01c)$ for the model with random screw disorder and $(\varphi_0, z_0) = (20^\circ, 0.04c)$ for the model with small independent rotations and translations. The effect of directional disorder is small in both cases, however, being only to suppress some of the weak Bragg intensity on layer line 1 for $R \simeq 0.15 \text{ \AA}^{-1}$. This is not significant relative to the observed pattern. There is therefore no evidence for or against the presence

of directional disorder in the specimen. However, the simulations do show what the geometric relationship between up and down molecules must be if directional disorder is present.

In conclusion, the model incorporating random screw disorder with axial and lateral disorder gives a diffraction pattern (Fig. 3c) that best matches the observed data. The model with independent small rotations and axial translations (Fig. 3d) does not match the data quite as well. The calculated two-dimensional diffraction pattern corresponding to Fig. 3(c), displayed in one quadrant of reciprocal space, is shown in Fig. 4(b), where it can

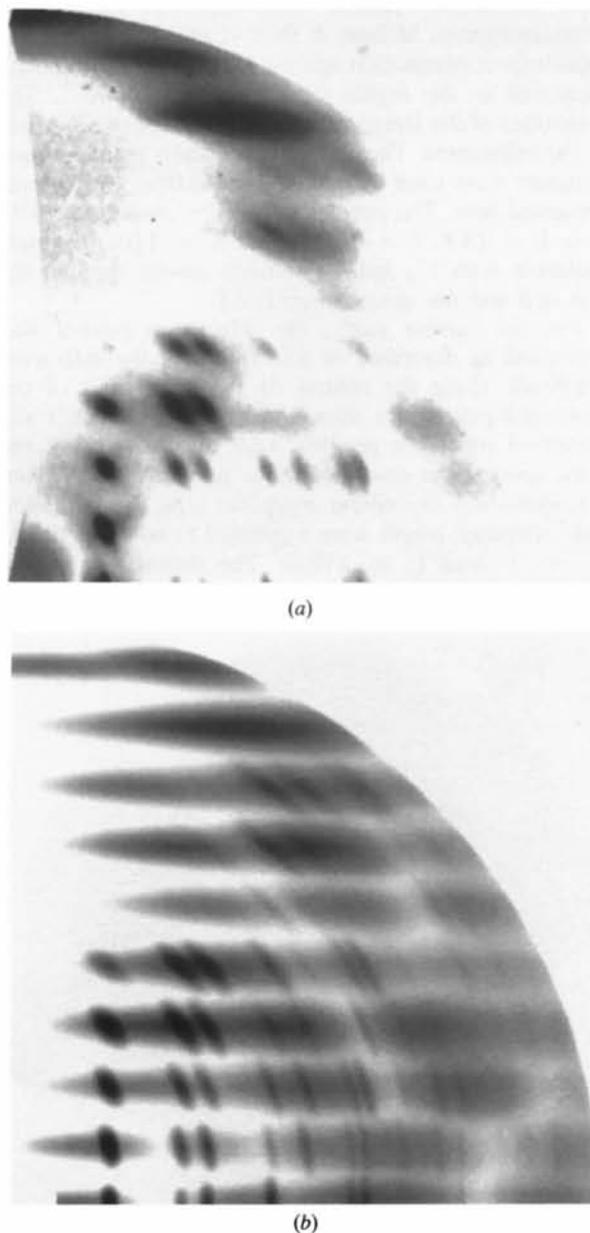


Fig. 4. Diffraction patterns for α -poly(dA)·poly(dT) in one quadrant of reciprocal space. (a) Measured pattern; (b) calculated pattern as in Fig. 3(c).

be compared with the actual diffraction pattern mapped into reciprocal space (Fig. 4a). The overall agreement between the two is seen to be rather good.

4. A DNA-RNA duplex

A fiber diffraction pattern from the DNA-RNA hybrid polynucleotide duplex poly(dA)·poly(rU) is shown in Fig. 5 (Arnott, Chandrasekaran, Banerjee, He & Walker, 1983). The pattern is dominated by continuous diffraction, with only seven Bragg reflections discernible on the first three layer lines at less than 7 Å resolution. This diffraction pattern was analyzed by Arnott, Chandrasekaran, Millane & Park (1986) by refining the molecular conformation against the continuous intensity measured in the region $0.14 \leq \rho \leq 0.30 \text{ \AA}^{-1}$. The intensities of the Bragg reflections were not used as data in the refinement. The atomic coordinates of the refined structure were used to calculate the diffraction patterns presented here. The average unit cell is monoclinic with $a = b = 24.8$, $c = 33.7 \text{ \AA}$ and $\gamma = 120^\circ$, a single molecule with 11_1 helix symmetry passes through the unit cell and the space group is $P1$.

For the current study, the diffraction pattern was processed as described in §2. Traces of the diffracted amplitude along the centers of the layer lines of the processed pattern are shown in Fig. 6. Calculated and measured reflection profiles were matched (using the same approach as described in the previous section) and the parameters describing crystallite size, disorientation and coherence length were estimated to be $r_c = 100 \text{ \AA}$, $\alpha_0 = 3^\circ$ and $l_c = 175 \text{ \AA}$. The diffraction pattern calculated for an ideally polycrystalline fiber with these

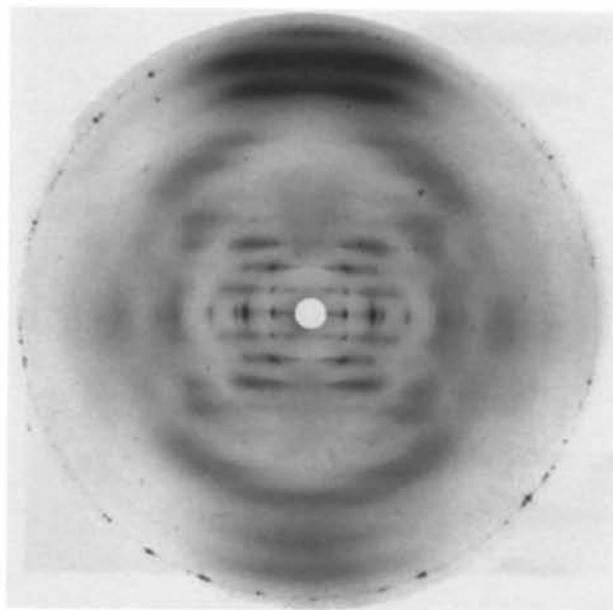


Fig. 5. Fiber diffraction pattern from poly(dA)·poly(rU) (Arnott *et al.*, 1983).

parameters is shown in Fig. 7(a). The upper layer lines in Fig. 6 show no Bragg interference, so that in this case there is no evidence for screw disorder. However, the predominance of continuous intensity indicates the presence of substantial lattice disorder (Stroud & Millane, 1995). Therefore, lateral disorder was introduced into the model in order to remove sharp reflections from the equator for $R \gtrsim 0.14 \text{ \AA}^{-1}$; this is achieved for $\sigma_{\text{lat}} = 1.2 \text{ \AA}$. The resulting diffraction pattern is shown in Fig. 7(b). Considerable Bragg intensity remains on the upper layer lines of this pattern but can be eliminated by introducing axial and rotational disorder into the model.

The amplitudes of the Bragg reflections for $R \leq 0.14 \text{ \AA}^{-1}$ on layer lines $l = 1$ and 2 are determined only by the Fourier-Bessel structure factors of order $n = l$. For uncorrelated rotations and axial translations, these reflections are weighted by (Stroud & Millane, 1995)

$$w = \exp(-l^2\sigma_\varphi^2 - 4\pi^2\sigma_{\text{axial}}^2l^2/c^2). \quad (2)$$

If σ_φ and σ_{axial} are varied such that

$$\kappa = \sigma_\varphi^2 + 4\pi^2\sigma_{\text{axial}}^2/c^2 \quad (3)$$

is constant, the low-resolution region of the diffraction pattern does not change. However, as σ_φ is increased, the Bragg reflections at large R on the lower layer lines are attenuated since higher-order Fourier-Bessel structure factors contribute to their amplitudes. At large R then, the effect of small rotations is to assist lattice disorder in removing Bragg reflections from the pattern. Calculations showed that combinations of σ_φ and σ_{axial} for which $\kappa \simeq 0.3$ give satisfactory suppression of

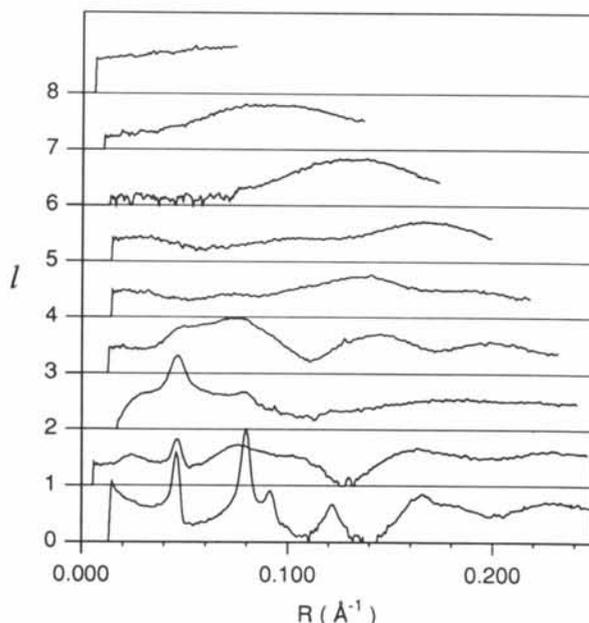


Fig. 6. Amplitudes along the centers of the layer lines of the diffraction pattern from poly(dA)·poly(rU).

reflections on the upper layer lines. Varying both σ_φ and σ_{axial} showed that the combination $\sigma_\varphi = 30^\circ$ and $\sigma_{\text{axial}} = 1.2 \text{ \AA}$ gives the optimum suppression of Bragg reflections at high resolution (Fig. 7c) and the best agreement between the calculated and observed diffraction patterns.

As a result of helix symmetry, a rotation of $\Delta\varphi = 32.7^\circ$ and a translation of $\Delta z = 3.06 \text{ \AA}$ maps one helix repeat unit of poly(dA)·poly(rU) onto the next. The value of σ_φ determined above is comparable to $\Delta\varphi$, suggesting that random screw disorder could be present in the specimen. However, calculations for models with random screw disorder showed that, in order to remove Bragg reflections on the upper layer lines, axial disorder,

with $\sigma_{\text{axial}} = 2.0 \text{ \AA}$, must also be present in order to remove Bragg reflections from the upper layer lines. The correlation between the rotations of the molecules and their axial translations is then extremely weak and, given the aliasing effect of molecular symmetry (Stroud & Millane, 1995), the resulting description of the disorder in the specimen is virtually identical to one in terms of uncorrelated rotations and axial translations.

The calculated diffraction pattern shown in Fig. 7(c) matches the observed pattern (Fig. 6) rather well in terms of the overall distribution of Bragg and continuous intensity. However, the calculated pattern differs from that observed in the following respects: (1) there are four sharp peaks on layer line $l = 1$, where just one is

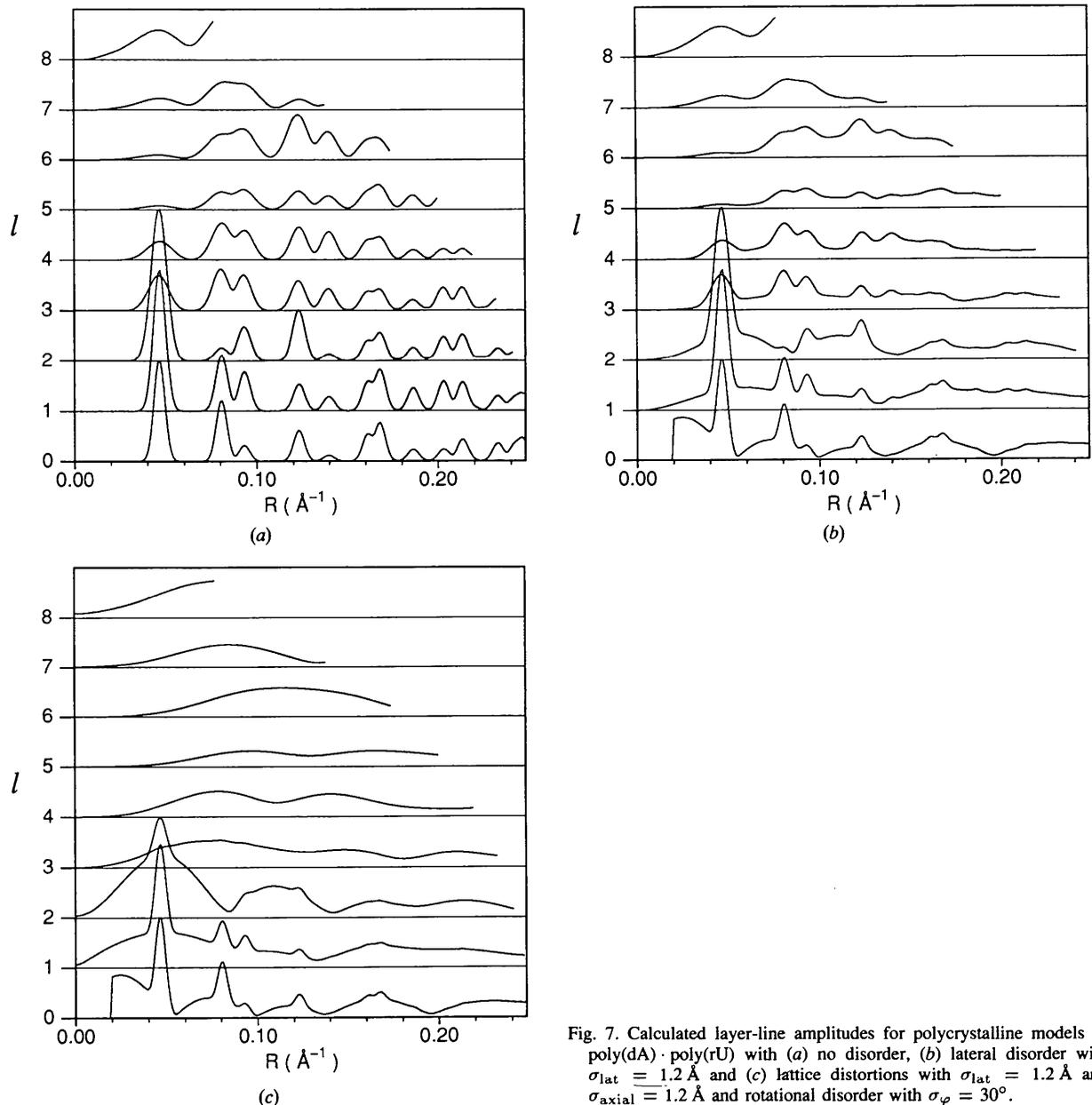


Fig. 7. Calculated layer-line amplitudes for polycrystalline models of poly(dA)·poly(rU) with (a) no disorder, (b) lateral disorder with $\sigma_{\text{lat}} = 1.2 \text{ \AA}$ and (c) lattice distortions with $\sigma_{\text{lat}} = 1.2 \text{ \AA}$ and $\sigma_{\text{axial}} = 1.2 \text{ \AA}$ and rotational disorder with $\sigma_\varphi = 30^\circ$.

present on the observed pattern; (2) on layer line $l = 2$, the continuous intensity at small R is slightly greater than observed and the weak reflection at $R \simeq 0.08 \text{ \AA}^{-1}$ on the observed pattern is missing; and (3) the first peak on layer line $l = 1$, at $R \simeq 0.05 \text{ \AA}^{-1}$, is stronger than the first peak on layer lines $l = 0$ and $l = 2$, whereas on the observed pattern it is weaker.

None of the differences noted above can be reduced or eliminated by further manipulating the parameters describing lattice or rotational disorder. Increasing σ_{lat} , σ_{axial} or σ_{φ} in order to suppress the three Bragg reflections on layer line $l = 1$ reduces the intensity of Bragg reflections elsewhere; decreasing these disorder components, in order to decrease the continuous intensity on layer line $l = 2$, introduces unwanted Bragg intensity at low resolution. The absence of the small reflection at $R \simeq 0.08 \text{ \AA}^{-1}$ on layer line $l = 2$ cannot be corrected by manipulating disorder parameters since it results mainly from the small amplitude of the calculated molecular transform at this position. The amplitude of the first reflection on layer line $l = 1$ cannot be reduced so that it is less than that of the corresponding reflection on layer line $l = 2$ simply by varying the axial and rotation disorder in the model since, for small R , the weights corresponding to both these disorder components decrease with increasing l . The discrepancies noted above are therefore likely to be due either to errors in the molecular transform or to the presence of other types of disorder. Other types of disorder that we considered were discrete rotations and translations and directional disorder.

Discrete rotations and/or translations of the molecules (either alone or in conjunction with directional disorder) affect the Bessel orders on different layer lines in a manner that does not vary smoothly with layer-line index (Stroud & Millane, 1995) and could, therefore, account for discrepancy 3 noted above. To explore this possibility, we assumed that the molecules adopt one of two equally probable positions and, holding all of the other model parameters fixed at the values as for Fig. 7(c), we calculated a series of diffraction patterns that covered the full range of distinguishable relative rotations and translations (φ', z') between these two positions (Stroud & Millane, 1995). The desired effect of reducing the intensity of the first reflection on layer line $l = 1$ and eliminating the other reflections on this layer line, without adversely affecting the amplitude of the first reflection on the layer line $l = 2$, is obtained for $(\varphi', z') = (0^\circ, 0.38c)$ (Fig. 8a). The calculated pattern is then improved over the diffraction pattern shown in Fig. 7(c) with regard to these reflections, however it is worse in the sense that it now shows too much continuous intensity on layer line $l = 1$ for $R \lesssim 0.07 \text{ \AA}^{-1}$.

Discrete axial translations and rotations can occur in combination with directional disorder (random up/down chain direction). A survey of the parameters φ_0 and z_0 that define the rotation axis relating the up and down

molecules (Stroud & Millane, 1995), while σ_{lat} , σ_{axial} and σ_{φ} were held at the values previously determined, showed that the best match to the observed pattern is obtained for $(\varphi_0, z_0) = (0^\circ, 0.22c)$ or $(45^\circ, 0.35c)$. Fig. 8(b) shows the diffraction pattern for the first case and the pattern for the second case is essentially identical. Regardless of the values of φ_0 and z_0 , however, directional disorder reinforces a peak on layer line $l = 1$ at $R \simeq 0.12 \text{ \AA}^{-1}$ and introduces some weak additional structure at higher resolution (Fig. 8b). Neither of these features are seen in the observed diffraction pattern and they cannot be counteracted by increasing σ_{φ} .

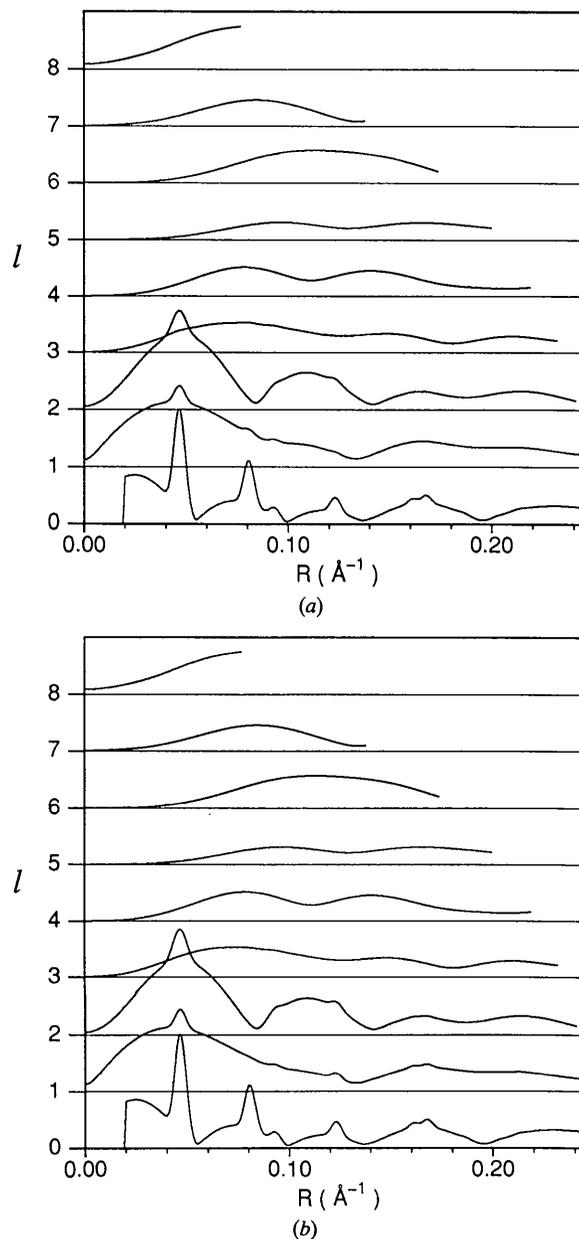


Fig. 8. Calculated layer-line amplitudes for polycrystalline models of poly(dA) · poly(rU) as in Fig. 7(c), but with (a) two discrete positions and (b) two discrete positions and directional disorder.

Models incorporating directional disorder are therefore marginally poorer than those that do not.

In conclusion, the above analysis allows the dominant forms of disorder in this specimen to be identified, although the solution is not quite as satisfactory as it is in the case of poly(dA)·poly(dT). Both models with (Fig. 8a) and without (Fig. 7c) discrete positions reproduce the restriction of the Bragg reflections to the central region of the pattern and the overall agreement of these patterns with the observed pattern (Fig. 6) is rather good. For one molecular position, the remaining discrepancies relate to the relative amplitudes of the Bragg reflections at low resolution and these could be due either to additional components of disorder or to errors in the molecular transform. Introducing discrete positions corrects these discrepancies but produces excessive continuous intensity in this region. One other disparity that is present in both Figs. 7(c) and 8(a) is that the second reflection on the equator is weaker than the first, whereas the reverse is true on the observed pattern. The disorder weight applied to the Bragg reflections in this region decreases monotonically with increasing R so that increased disorder would not affect this reversal in peak amplitudes. This discrepancy therefore points to an inconsistency between the molecular transform and the diffraction data. This could be due to small errors in the conformation of the molecule or to organized water molecules or ions on the periphery of the molecule, which would increase its effective radius and thus modify the molecular transform in this region.

5. Orientations of the molecules in the unit cell

In the above description of the disordered crystal structures of α -poly(dA)·poly(dT) and poly(dA)·poly(rU), we have omitted reference to the average position of the molecules in the unit cell. Given the space group of these structures and the symmetries of the molecules, this position is described by a single parameter, the orientation of the molecule about its long axis. Here we show that this parameter cannot be determined from the available diffraction data.

The cylindrical average of the intensity of the hkl reflection diffracted from a fiber with uncorrelated lattice and substitution disorder can be written as

$$I_{hkl}(R) = (1/A^2)w(R_{hk}, l/c) \left| \sum_n w_{nl} G_{nl}(R_{hk}) \times \exp[in(\psi_{hk} + \pi/2)] \right|^2 \mathcal{P}_{hk}(R), \quad (4)$$

where $(R_{hk}, \psi_{hk}, l/c)$ are the cylindrical polar coordinates of the reflection in reciprocal space, $w(R, Z)$ is the lattice disorder weight, w_{nl} is the substitution disorder weight, A is the area of the unit cell in the plane perpendicular to the Z axis, $\mathcal{P}_{hk}(R)$ is the cylindrically averaged profile of the hkl reflection and $G_{nl}(R)$ is

the Fourier–Bessel structure factor of order n (Stroud & Millane, 1995). The summation in (4) is over all Bessel orders n satisfying the helix selection rule (Cochran, Crick & Vand, 1952).

As a result of cylindrical averaging, an observed Bragg reflection is a composite reflection and has an intensity, denoted here by $I'_{hkl}(R)$, equal to the sum of the intensities $I_{hkl}(R)$ of the reflections centered at R_{hk} on the plane $Z = l/c$. For a hexagonal lattice, $I'_{hkl}(R)$ is given by

$$I'_{hkl}(R) = (1/A^2)w(R_{hk}, l/c) \left| \sum_n w_{nl} G_{nl}(R_{hk}) \times \sum_{j=0}^5 \exp[in(\psi'_{hk} + j\pi/3 + \pi/2)] \right|^2 \times \mathcal{P}_{hk}(R), \quad (5)$$

where ψ'_{hk} is the value of ψ_{hk} for any one of the overlapping reflections. If the position of the molecule in the unit cell is changed by rotating it through an angle φ about its long axis, then from the expression for the Fourier–Bessel structure factors (Cochran, Crick & Vand, 1952) it follows that (5) is replaced by

$$I'_{hkl}(R) = (1/A^2)w(R_{hk}, l/c) \sum_n \sum_m w_{nl} w_{ml}^* \times G_{nl}(R_{hk}) G_{ml}^*(R_{hk}) \exp[i(n-m)(\psi'_{hk} - \varphi + \pi/2)] \Delta_{nm} \mathcal{P}_{hk}(R), \quad (6)$$

where

$$\Delta_{nm} = \sum_{j=0}^5 \exp[i(n-m)j\pi/3] = \begin{cases} 1 & \text{for } n-m \text{ divisible by } 6, \\ 0 & \text{otherwise.} \end{cases} \quad (7)$$

For a molecule with u_v helix symmetry, the difference $(n-m)$ between any two solutions of the helix selection rule for the same layer line is an integral multiple of u . Therefore, the only non-zero terms in (6) are those for which $n-m$ is divisible both by u and by 6. Of these, only those terms for which n and m are less than $n_{\max} = 2\pi r_{\max} R + 2$, where r_{\max} is the maximum radius of the molecule, are significant at radius R (Crowther, DeRosier & Klug, 1970). The maximum radius of α -poly(dA)·poly(dT) is approximately 11 Å and, for $R \leq 0.3 \text{ \AA}^{-1}$, $n_{\max} \simeq 22$ so that over the observed diffraction pattern the only significant terms in (6) are those for which $n = m$. These terms do not involve the angle φ and hence changing the average orientation of the molecule in the unit cell has no effect on the observed Bragg intensity. A similar argument leads to the same conclusion regarding the continuous intensity. Therefore, within the observed region of reciprocal space, the diffraction pattern from α -poly(dA)·poly(dT) does not depend on the orientation of the molecule in the unit cell. With the same reasoning, the same can be shown for the diffraction pattern from poly(dA)·poly(rU).

6. Discussion

The results presented here show that a detailed model of disorder in polycrystalline fibers and expressions for the effects on diffraction patterns (Stroud & Millane, 1995) can be used to analyze disorder in these kinds of specimen. With an approximate molecular model, the primary components of the disorder can be identified and quantified, by matching the calculated and measured distributions of Bragg and continuous amplitudes in different regions of a diffraction pattern. This appears to lead to unique descriptions of disorder, although elucidation of finer details would require that the molecular transform, and therefore the molecular model, be specified rather accurately. Once the disorder in a specimen has been characterized, the potential exists for including its effects in structure determination, leading to a more accurate structure.

The methods described here are semi-quantitative in the sense that, for the most part, the calculated and observed patterns were compared visually. The range of possible types and degrees of disorder that may occur in fibers is large and the semi-quantitative approach adopted here permits an initial survey of the disorder parameter space. The kinds of disorder and approximate values of disorder parameters inferred from such a survey give starting values that could be subjected to quantitative refinement. The approach we adopted for conducting this survey is based on matching the presence of key features on the diffraction patterns. The nature of these features varies with the type of disorder, and inspection of the overall distribution of Bragg and continuous intensity over the diffraction pattern should in many cases permit a large number of possible disorder models to be disqualified from consideration, leaving relatively few models to be considered in detail. Detailed analysis may then lead to a unique model of disorder. The issue of uniqueness has previously been discussed by us in detail (Stroud & Millane, 1985). In many cases, non-uniqueness in the description of disordered packing may only be illusory, with the underlying disorder actually being the same and the different descriptions merely parameterizing this disorder in different ways. In the case of poly(dA)·poly(dT), it appears that a unique description of the disorder can be derived but for poly(dA)·poly(rU) the disorder is more pronounced, and there is some ambiguity at the level of detail of the analysis presented here.

It is not possible to arrive at a single final packing model for the two molecules discussed here since the diffraction data are not sensitive to the average orientation of the molecules in the unit cell. This is a result of the particular combinations of molecular and lattice symmetry, not of disorder, and does not occur in general. Where it does occur, our analysis allows at least some aspects of the packing to be quantified. In many cases, molecular models may reveal steric

interactions between adjacent molecules that limit the range of possible orientations in the unit cell.

The fiber of poly(dA)·poly(dT) is the more ordered of the two specimens examined here and the diffraction pattern is best explained by random screw disorder accompanied by small axial and lateral disorder. Presumably, the looser lateral packing (relative to a fully polycrystalline specimen) allows the molecules to adopt a variety of axial and rotational positions relative to their neighbors but the lateral packing is still tight enough that such distortions are constrained by the helical nature of the molecule.

The fiber of poly(dA)·poly(rU) is more disordered and the diffraction pattern is better explained by rather larger lateral and axial disorder, together with significant rotations of the molecules. In this case, it seems reasonable that the larger degree of lateral disorder allows the molecules to move more independently of their neighbors so that they are not restricted to only screw rotations.

Supported by Purdue University and the US National Science Foundation (DMB-8916477 and MCB-9219736 to RPM). We thank the Electron Microscopy Center in Agriculture at Purdue University for use of their equipment to reproduce Fig. 4.

References

- ARNOTT, S. (1980). *Twenty Years Hard Labor as a Fiber Diffractionist. Fiber Diffraction Methods*, edited by A. D. FRENCH & K. H. GARDNER, pp. 1–30. *ACS Symposium Series*, Vol. 141. Washington: American Chemical Society.
- ARNOTT, S., CHANDRASEKARAN, R., BANERJEE, A. K., HE, R. & WALKER, J. K. (1983). *J. Biomol. Struct. Dynam.* **1**, 437–452.
- ARNOTT, S., CHANDRASEKARAN, R., MILLANE, R. P. & PARK, H. (1986). *J. Mol. Biol.* **188**, 631–640.
- ARNOTT, S. & SELSING, E. (1974). *J. Mol. Biol.* **88**, 509–521.
- COCHRAN, W., CRICK, F. H. C. & VAND, V. (1952). *Acta Cryst.* **5**, 581–586.
- CROWTHER, R. A., DEROSIER, D. J. & KLUG, A. (1970). *Proc. R. Soc. London Ser. A*, **317**, 319–340.
- FRASER, R. D. B., MACRAE, T. P., MILLER, A. & ROWLANDS, R. J. (1976). *J. Appl. Cryst.* **9**, 81–94.
- FRASER, R. D. B., MACRAE, T. P. & SUZUKI, E. (1978). *J. Appl. Cryst.* **11**, 693–694.
- HOLMES, K. C. & BARRINGTON LEIGH, J. (1974). *Acta Cryst.* **A30**, 635–638.
- MILLANE, R. P. (1988). *Structure Determination by X-ray Fiber Diffraction. Crystallographic Computing 4: Techniques and New Technologies*, edited by N. W. ISAACS & M. R. TAYLOR, pp. 169–186. Oxford Univ. Press.
- MILLANE, R. P. & ARNOTT, S. (1985). *J. Appl. Cryst.* **18**, 419–423.
- MILLANE, R. P. & ARNOTT, S. (1986). *J. Macromol. Sci. Phys.* **B24**, 193–227.
- MILLANE, R. P., CHANDRASEKARAN, R., ARNOTT, S. & DEA, I. C. M. (1988). *Carbohyd. Res.* **182**, 1–17.
- MILLER, A. & PARRY, D. A. D. (1974). *Polymer*, **15**, 706–712.
- NAMBA, K. & STUBBS, G. J. (1985). *Acta Cryst.* **A41**, 252–262.
- PARK, H., ARNOTT, S., CHANDRASEKARAN, R., MILLANE, R. P. & CAMPAGNARI, F. (1987). *J. Mol. Biol.* **197**, 513–523.
- STROUD, W. J. & MILLANE, R. P. (1995). *Acta Cryst.* **A51**, 000–000.
- STUBBS, G. J. (1974). *Acta Cryst.* **A30**, 639–645.