typically between about one-third and one-half of the corresponding largest likely R factor, depending on the resolution. It is therefore reasonable to expect the R factor for a well determined fibre structure to be between one-third and one-half of the largest likely R factor calculated for the structure. R factors should, therefore, generally be less than 0.15 to 0.25, depending on the particular structure and the resolution as illustrated by the examples presented in Millane & Stubbs (1992).

The free R factor (Brünger, 1997) has become popular in single-crystal crystallography as a tool for validation of refinements. The free R factor is more difficult to implement (but is probably even more important) in fibre diffraction studies because of the smaller data sets, but has been used to advantage in recent studies (Hudson *et al.*, 1997; Welsh *et al.*, 1998, 2000).

4.5.3. Electron crystallography of polymers

By D. L. Dorset

4.5.3.1. Is polymer electron crystallography possible?

As a crystallographic tool, the electron microscope has also made an important impact in polymer science. Historically, single-crystal electron diffraction information has been very useful for the interpretation of cylindrically averaged fibre X-ray patterns (Atkins, 1989), particularly when there is an extensive overlap of diffracted intensities. An electron diffraction pattern aids indexing of the fibre pattern and facilitates measurement of unit-cell constants, and the observation of undistorted plane-group symmetry similarly places important constraints on the identification of the space group (Geil, 1963; Wunderlich, 1973).

The concept of using electron diffraction intensities by themselves for the quantitative determination of crystal structures of polymers or other organics often has been met with scepticism (Lipson & Cochran, 1966). Difficulties experienced in the quantitative interpretation of images and diffraction intensities from 'hard' materials composed of heavy atoms (Hirsch *et al.*, 1965; Cowley, 1981), for example, has adversely affected the outlook for polymer structure analysis, irrespective of whether these reservations are important or not for 'soft' materials comprising light atoms. Despite the still commonly held opinion that no new crystal structures will be determined that are solely based on data collected in the electron microscope, it can be shown that this extremely pessimistic outlook is unwarranted. With proper control of crystallization (*i.e.* crystal thickness) and data collection, the electron microscope can be used quite productively for the direct determination of macromolecular structures at atomic resolution, not only to verify some of the previous findings of fibre X-ray diffraction analysis, but, more importantly, to determine new structures, even of crystalline forms that cannot be studied conveniently by X-rays as drawn fibres (Dorset, 1995*b*). The potential advantages of electron crystallography are therefore clear. The great advantage in scattering cross section of matter for electrons over X-rays permits much smaller samples to be examined by electron diffraction as single-crystalline preparations (Vainshtein, 1964). (Typical dimensions are given below.)

Electron crystallography can be defined as the quantitative use of electron micrographs and electron diffraction intensities for the determination of crystal structures. In the electron microscope, an electron beam illuminates a semitransparent object and the microscope objective lens produces an enlarged representation of the object as an image. If the specimen is thin enough and/or the electron energy is high enough, the weak-phase-object or 'kinematical' approximation is valid (Cowley, 1981), see Chapter 2.5. That is to say, there is an approximate one-to-one mapping of density points between the object mass distribution and the image, within the resolution limits of the instrument (as set by the objective lens aberrations and electron wavelength). The spatial relationships between diffraction and image planes of an electron microscope objective lens are reciprocal and related by Fourier transform operations (Cowley, 1988). While it is easy to transform from the image to the diffraction pattern, the reverse Fourier transform of the diffraction pattern to a high-resolution image requires solution of the famous crystallographic phase problem (as discussed for electron diffraction in Section 2.5.8).

Certainly, in electron diffraction studies, one must still be cognizant of the limitations imposed by the underlying scattering theory. An approximate 'quasi-kinematical' data set is often sufficient for the analysis (Dorset, 1995*a*). However (Dorset, 1995*b*), there are other important perturbations to diffraction intensities which should be minimized. For example, the effects of radiation damage while recording a high-resolution image are minimized by so-called 'low-dose' procedures (Tsuji, 1989).

4.5.3.2. Crystallization and data collection

The success of electron crystallographic determinations relies on the possibility of collecting data from *thin* single microcrystals. These can be grown by several methods, including self-seeding, epitaxic orientation, *in situ* polymerization on a substrate, in a Langmuir–Blodgett layer, *in situ* polymerization within a thin layer and polymerization in dilute solution. If these preparations do not provide sufficient information, then data can also be collected from microfibres. Thin cast films have also been examined after stretching.

Self-seeding (Blundell *et al.*, 1966) has been one of the most important techniques for growing single chain-folded lamellae. The technique is very simple. A dilute suspension of the polymer is made in a poor solvent. The temperature is raised to cause total solubilization of the macromolecule and then lowered to room temperature to crystallize ill-formed particles (mostly dendrites). The temperature is then elevated again until the suspension just clears, leaving small seeds of the polymer crystals behind. Upon lowering to a suitable temperature above ambience, which is then fixed, isothermal crystallization of well formed lamellae is allowed to occur over time. When the crystallization procedure is complete, the suspension can be cooled again to room temperature and the lamellae harvested. These lamellae are typically less than 10 nm thick, with lateral dimensions between 1.0 and 10.0 μm.

Epitaxic orientation techniques, to give alternative projections of the chain packing, have become increasingly important in recent years. While inorganic substrates have been described (Mauritz *et al.*, 1978), the use of organic layers for this purpose (Wittmann & Lotz, 1990; Lotz & Wittmann, 1993) has been more promising because these substrates are less easily contaminated by adsorbed gases and water vapour, and because the nucleation is anisotropic. Often the crystallization can be carried out from a cooled co-melt, *i.e.* a dilute solution of the polymer in the organic small molecule. When the liquidus curve of the eutectic phase diagram is crossed, the diluent crystals form first. Since these have a surface lattice spacing closely resembling that of the polymer-chain packing, the polymer chains can be directed to lie *along* the substrate surface, rather than normal to it, as the solidus line of the phase diagram is crossed. The substrate can then be removed by some suitable technique (sublimation, selective solvation) to permit the investigation of the oriented film. Variations of this procedure include crystallization of polymer-chain segments from the vapour phase onto a substrate (Wittmann & Lotz, 1985) and *in situ* crystallization of monomers that have first been epitaxically oriented on a suitable substrate (Rickert *et al.*, 1979).

A number of other possibilities for crystal growth also exist. Langmuir troughs have been used to orient monomers that may have hydrophilic moieties. If the monomers contain triple bonds that can be cross-linked, then a polymer film can be formed, *e.g.*,

if the condensed monomer film is exposed to ultraviolet light (Day & Lando, 1980). It may be possible to carry out the polymerization within a confined space (Rybnikar *et al.*, 1994) or in dilute solution (Liu & Geil, 1993) to form crystals suitable for electron diffraction data collection. In the latter case, whisker formation with the chain axis parallel to the lath plane has been observed. Films can be cast on a water surface by evaporation of an organic solvent from a polymer solution. The film can then be stretched to give a suitably oriented specimen for data collection (Vainshtein & Tatarinova, 1967). In addition, it may just be possible to obtain suitable data from drawn microfibres to supplement the single-crystal diffraction information from other preparations.

Data collection from these thin microcrystals often employs the selected-area diffraction technique in the electron microscope that is described in detail elsewhere (Dorset, 1995*b*). Using an approximately eucentric goniometric tilting device in the electron microscope, the sampling of three-dimensional reciprocal space is tomographic, *i.e.* the tilts of a nearly planar Ewald sphere surface (owing to the very small electron wavelength) are always referred to a set of reciprocal axes that intersect (0, 0, 0). For any given crystal habit, there is always a missing set of data owing to the physical limitation to the tilt imposed by the finite thickness of the specimen holder within the pole-piece gap of the electron microscope objective lens (Vainshtein, 1964). For this reason, it is desirable to crystallize two orthogonal orientations of the chain packing (using the above-mentioned approaches), if possible, so that all of the reciprocal lattice can be sampled. If electron micrographs are to be used as an additional source of crystallographic phases then 'low-dose' techniques for recording such images should be employed to reduce the deleterious effects of radiation damage caused by the inelastic interactions of the electron beam with the crystalline sample (Tsuji, 1989).

When the diffraction patterns are recorded on photographic film and these are then measured with a densitometer, relative reflection intensities can often be expressed simply as the integrated peak area without need for a Lorentz correction (Dorset, 1995*b*). Only if the diffraction maxima are extensively arced (*e.g.* in patterns from epitaxic films) is such a correction required. That is to say, $|\Phi_{\rm obs}| \propto KI_{\rm obs}^{1/2}$ where $|\Phi_{\rm obs}|$ is the observed structure-factor magnitude. Assuming the kinematical approximation holds, the calculated value is

$$\Phi_h^{\rm calc} = \sum_i f_i \exp 2\pi i(\mathbf{h}\cdot\mathbf{r}),$$

where f_i are the electron scattering factors (Doyle & Turner, 1968), *e.g.* as tabulated in Table 4.3.1.1 in *IT* C. By analogy with X-ray crystallography (see Chapter 2.2), normalized values can be found from

$$|E_h|^2 = I_h^{\rm obs}/\varepsilon\sum_i f_i^2,$$

with the usual scaling condition that $\langle E_h^2\rangle = 1.000$. [Note, however, that these intensities only describe the chain monomer packing in the 'stem' region of the lamellar microcrystal. Details owing to the surface chain folds are lost (even if they are strictly periodic), because of reasons similar to those described by Cowley (1961) for the electron scattering from elastically bent silicate crystals.]

4.5.3.3. Crystal structure analysis

Two approaches to crystal structure analysis are generally employed in polymer electron crystallography. As already mentioned, the procedure adapted from fibre X-ray crystallography relies on the construction of a model (Brisse, 1989; Perez & Chanzy, 1989). Conformational searches (Campbell Smith & Arnott, 1978) simultaneously minimize the fit of observed diffraction data to calculated values (the *R* factor based on structure factors computed *via* known atomic scattering factors) and a nonbonded atom–atom potential function (Tadokoro, 1979). Reviews of structures solved by this approach have been published (Dorset, 1989, 1995*b*).

Recently, direct phasing methods of the kind used in X-ray crystallography (Chapter 2.2 and, applied to electron diffraction, Section 2.5.8) have also been found to be particularly effective for electron crystallographic structure analyses (Dorset, 1995*b*). While the Fourier transform of an electron micrograph would be the most easily imagined direct method, yielding crystallographic phases after image analysis (see Section 2.5.5), this use of micrographs has been of less importance to polymer crystallography than it has been in the study of globular proteins, even though there is at least one notable example where it has been helpful (Isoda *et al.*, 1983*a*) for the determination of a structure from X-ray fibre data. On the other hand, high-resolution images of polymer crystals are of considerable use for the characterization of packing defects (Isoda *et al.*, 1983*b*).

In polymer electron crystallography, the sole reliance on the diffraction intensities for structure analysis has proven, in recent years, to be quite effective. Several direct-methods approaches have been pursued, including the use of probabilistic techniques, either in the symbolic addition procedure, or in more automated procedures involving the tangent formula (see Chapter 2.2). The Sayre (1952) equation has been found to be particularly effective, where the correct structure is identified *via* some figure of merit after algebraic phase values are used to generate multiple solutions (Stanley, 1986). More recently, maximum-entropy and likelihood methods (Gilmore *et al.*, 1993) have also been effective for solving such structures. After the initial atomic model is found, it can be improved by refinement, generally using Fourier techniques. Least-squares refinement can be carried out under most favourable circumstances (Dorset, 1995*a*), but requires the availability of a sufficient number of diffraction data. Even so, the refinement of thermal parameters must be uncoupled from that of the atomic positions. Also, positional shifts must be dampened (if X-ray crystallographic software is used) to prevent finding a false minimum, especially if the kinematical *R* factor is used as a figure of merit.

4.5.3.4. Examples of crystal structure analyses

At least four kinds of electron diffraction intensity data sets have been used as tests for direct phase determination *via* the approaches mentioned above.

Case 1: Zonal data sets – view down the chain axis. Such data are from the least optimal projection of the polymer packing, because of extensive atomic overlap along the chain axis. Initially, symbolic addition was used to find phase values for *hk*0 data sets from six representative polymers, including three complicated saccharide structures (Dorset, 1992). Most of the determinations were strikingly successful. Later, an unknown data set from the polysaccharide chitosan was obtained from Grenoble (Mazeau *et al.*, 1994) and direct phase determination was able to find a correct model (Dorset, 1995*b*). More recently, other polymers have been tested [including one case where an electron micrograph provided many of the starting phase terms (Dorset, 1995*b*)] also comparing favourably with the solution found by energy minimization of a linkage model. For all examples considered so far, the projected symmetry was centrosymmetric.

Later, it was found that a partial phase set provided by symbolic addition could be expanded to the complete zone by the Sayre equation (Dorset *et al.*, 1995). In all of these tests (Dorset, 1995*b*), there were only one or two examples where there were serious deviations from the phase terms found by other methods. Even in these instances, the potential maps could still be used as

Table 4.5.3.1. *Structure analysis of poly-γ-methyl-L-glutamate in the β form*

h0l	$\|E_h\|$	$\|F_o\|$	$\|F_c\|$	φ (°) (previous)	φ (°) (this study)
002	0.48	0.72	0.57	−63	−51
004	0.43	0.38	0.31	49	73
006	3.01	1.47	0.88	1	−3
100	1.48	2.12	2.37	0	0
200	1.03	1.04	1.06	0	0
300	0.30	0.65	0.89	0	0
400	0.35	0.15	0.46	0	0
500	0.23	0.07	0.04	180	180
101	0.75	1.02	0.67	−169	−178
201	0.32	0.31	0.42	90	108
102	0.42	0.48	0.56	17	14
202	0.40	0.33	0.64	41	43
103	0.95	0.85	0.77	88	90
203	0.51	0.36	0.42	91	88
303	0.12	0.06	0.31	92	87
403	0.13	0.04	0.54	90	90
104	0.66	0.45	0.27	−22	−13
105	0.55	0.28	0.29	−26	−7
106	1.75	0.69	0.58	5	−5

Fractional coordinates

	This study		Vainshtein & Tatarinova (1967)	
	x	*z*	*x*	*z*
Cα, β	0.048	0.000	0.042	0.000
C′	0.067	0.331	0.092	0.330
O	0.281	0.335	0.300	0.330
N	0.000	0.161	−0.025	0.175

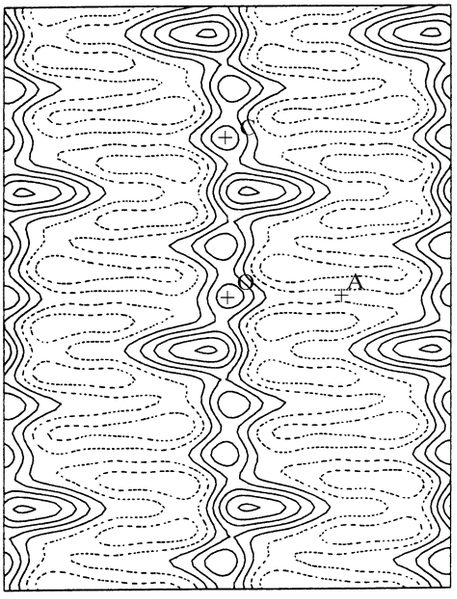

Fig. 4.5.3.1. Initial potential map for poly-γ-methyl-L-glutamate (plane group *pg*) found with phases generated by the Sayre equation.

envelopes for the actual projection of the chain structure (Dorset, 1992).

Case 2: Zonal data set – view onto the chain axes. Electron diffraction data from a projection onto the polymer chain axes would be more useful if individual atomic positions were to be resolved. An interesting example where such a view can be obtained is an *h0l* data set from the polypeptide poly-γ-methyl-L-glutamate. Electron diffraction data were collected from stretched films by Vainshtein & Tatarinova (1967). In projection, the cell constants are $a = 4.72$, $c = 6.83$ Å with plane-group symmetry *pg*. As shown in Table 4.5.3.1, there were 19 unique intensity data used for the analysis. After initial phase assignment by symbolic addition, a correct solution could be visualized which, after Fourier refinement (Dorset, 1995*b*), differed from the original one by a mean phase difference of only 6°.

The progress of this structure analysis can be reviewed to give a representative example. Since the *h*00 reflections have centrosymmetric phases, the value $\varphi_{100} = 0$ was chosen as a single origin-defining point. From high-probability Σ_1 three-phase invariants (assessed after calculation of normalized structure factors $|E_h|$), one could assign $\varphi_{200} = \varphi_{400} = 0$. Symbolic values were then given to three other phases, *viz.* $\varphi_{106} = a$; $\varphi_{103} = b$; $\varphi_{101} = c$. From this entire basis, other values could be found from highly probable Σ_2 three-phase invariants, as follows:

$$\begin{aligned}
\varphi_{006} &= \varphi_{106} + \varphi_{\bar{1}00} \;\;\therefore\;\; \varphi_{006} = a \\
\varphi_{105} &= \varphi_{006} + \varphi_{10\bar{1}} \;\;\therefore\;\; \varphi_{105} = a - c + \pi \\
\varphi_{203} &= \varphi_{106} + \varphi_{10\bar{3}} \;\;\therefore\;\; \varphi_{203} = a - b + \pi \\
\varphi_{300} &= \varphi_{100} + \varphi_{200} \;\;\therefore\;\; \varphi_{300} = 0 \\
\varphi_{002} &= \varphi_{103} + \varphi_{\bar{1}01} \;\;\therefore\;\; \varphi_{002} = b - c \\
\varphi_{004} &= \varphi_{006} + \varphi_{00\bar{2}} \;\;\therefore\;\; \varphi_{004} = a - b + c.
\end{aligned}$$

(These invariant relationships include phase interactions among symmetry-related Miller indices characteristic of the plane group.) Additionally $c = \pi$ could be specified to complete origin definition for the zone. It was then possible to permute values of *a* and *b* to arrive at test phase values for this subset, *i.e.* to generate a multiple set of solutions. When $a = 0$, $b = \pi/2$, the map in Fig. 4.5.3.1 was observed. After finding trial atomic positions for Fourier refinement (assuming that two carbon-atom positions were eclipsed in this projection), the final phase set was found as shown in Table 4.5.3.1. Although the crystallographic residual to the observed data, calculated with the model coordinates, was rather large (0.32), there was a close agreement with the earlier determination.

More recently a similar data set, collected from oriented crystal 'whiskers' of poly(*p*-oxybenzoate) in plane group *pg* was analysed. Again the Sayre equation, *via* a multisolution approach, was used to produce a map that contained 13 of 18 possible atomic positions for the two subunits in the asymmetric unit. The complete structure was observed after the remaining five atom sites were identified in two subsequent cycles of Fourier refinement (Liu *et al.*, 1997) and the average atomic positions were found to be within 0.2 Å of the model derived from an energy minimization.

Case 3: Three-dimensional data – single crystal orientation. The first data set from a chain-folded lamella for a direct structure analysis was a centrosymmetric set (space group $P2_1/n$) from poly(1,4-*trans*-cyclohexanediyl dimethylene succinate), composed of 87 reflections (Brisse *et al.*, 1984). The phase determination was quite successful and atomic positions could be found as somewhat blurred density maxima in the three-dimensional maps (Dorset, 1991*a*). A model was constructed from these positions and the bonding parameters optimized to give the best fit to the data ($R = 0.29$).

Noncentrosymmetric three-dimensional intensity sets (orthorhombic space group $P2_12_12_1$) from the polysaccharides mannan (form I) (Chanzy *et al.*, 1987) and chitosan (Mazeau *et al.*, 1994) were also collected from tilted crystals. In both cases, direct phase determination by symbolic addition *via* an algebraic unknown was successful, even though the data were not sampled along the chain repeat. For the former polymer, a monomer model could be fitted to the blurred density profile, much as one would fit a polypeptide sequence to a continuous electron-density map (Dorset & McCourt, 1993; Dorset, 1995*c*). If the Sayre equation were used to predict phases and amplitudes within the 'missing cone' of unsampled data, then the fit of the monomer could be much more highly constrained.

Case 4: Three-dimensional data – two crystal orientations. The optimal case for collection of diffraction data is when two

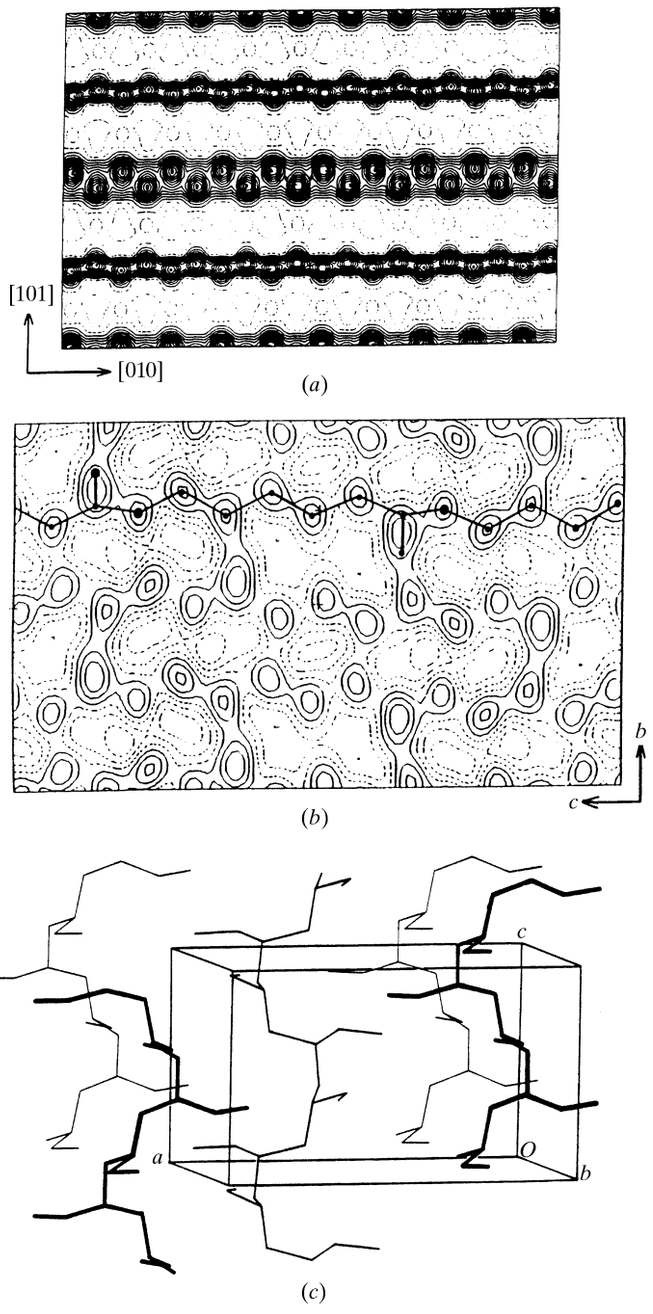

Fig. 4.5.3.2. Crystal structures of linear polymers determined from three-dimensional data. (*a*) Polyethylene; (*b*) poly(ε-caprolactone); (*c*) poly(1-butene), form (III).

orthogonal projections of the same polymer polymorph can be obtained, respectively, by self-seeding and epitaxic orientation. While tilting these specimens, all of reciprocal space can be sampled for intensity data collection.

Polyethylene crystals were used to collect 50 unique maxima (Hu & Dorset, 1989) and, *via* symbolic addition, the centrosymmetric phases of 40 reflections (space group *Pnma*) could be readily determined (Dorset, 1991*b*). The structural features were readily observed in the three-dimensional potential maps (Fig. 4.5.3.2*a*), and atomic coordinates (with estimated values for hydrogen-atom positions) could be refined by least squares (Dorset, 1995*b*) to give a final R value of 0.19.

Poly(ε-caprolactone) was epitaxically crystallized on benzoic acid and, with *hk*0 data from solution-crystallized samples, a unique set of 47 intensities was collected for the noncentrosymmetric orthorhombic unit cell (space group $P2_12_12_1$) (Hu & Dorset, 1990). Direct phase determination was achieved *via* symbolic addition, using one algebraic unknown to assign values to 30 reflections (Dorset, 1991*c*). Atomic positions along the chain repeat, including the carbonyl position, were clearly discerned in the [100] projection (Fig. 4.5.3.2*b*) and the three-dimensional model was constructed to fit to the map calculated from all phased data, yielding a final crystallographic residual R = 0.21. This independent determination was able to distinguish between two rival fibre X-ray structures, in favour of the one that predicted a nonplanar chain conformation. Because of the methylene repeat, this is actually a difficult structure to solve by automated techniques. For example, the tangent formula and *SnB* (Miller *et al.*, 1993) could only find chain zigzag positions and not the position of the carbonyl oxygen atom (Dorset, 1995*b*).

The most complicated complete polymer crystal structure solved so far by direct methods using electron diffraction data (Dorset *et al.*, 1994) was based on 125 unique data (space group $P2_12_12_1$) from isotactic poly(1-butene), form (III), using orthogonal molecular orientations crystallized in Strasbourg (Kopp *et al.*, 1994). Initially, the standard NQEST figure of merit (FOM) (De Titta *et al.*, 1975) was not suitable for identifying the correct solution among the multiple sets generated with the tangent formula. A solution could only be found when a separate phase determination was carried out with the *hk*0 data to compare with the multiple solutions generated. More recently, the minimal principle (Hauptman, 1993), used as a FOM with the tangent formula or with a multiple random structure generator, *SnB*, correctly identified the structure on the first try (Dorset, 1995*b*). The maps clearly show individual carbon-atom positions in a 4_1 helix that parallels 2_1 helices of the space group (Fig. 4.5.3.2*c*). After Fourier refinement, the crystallographic residual was R = 0.26. The previous powder X-ray diffraction determination was based on only 21 diffraction maxima, some of which had as many as 15 individual contributors.

The work of both RPM and DLD was supported by the US National Science Foundation (grant DBI-9722862 for RPM).

References

Alexeev, D. G., Lipanov, A. A. & Skuratovskii, I. Y. (1992). *Patterson methods in fibre diffraction. Int. J. Biol. Macromol.* **14**, 139–144.

Arnott, S. (1980). *Twenty years hard labor as a fibre diffractionist*. In *Fibre Diffraction Methods*, ACS Symposium Series, Vol. 141, edited by A. D. French & K. H. Gardner, pp. 1–30. Washington: American Chemical Society.

Arnott, S., Chandrasekaran, R., Millane, R. P. & Park, H. (1986). *DNA–RNA hybrid secondary structures. J. Mol. Biol.* **188**, 631–640.

Arnott, S. & Mitra, A. K. (1984). *X-ray diffraction analyses of glycosamionoglycans*. In *Molecular Biophysics of the Extracellular Matrix*, edited by S. Arnott, D. A. Rees & E. R. Morris, pp. 41–67. Clifton: Humana Press.

Arnott, S., Wilkins, M. H. F., Fuller, W. & Langridge, R. (1967). *Molecular and crystal structures of double-helical RNA III. An 11-fold molecular model and comparison of the agreement between the observed and calculated three-dimensional diffraction data for 10- and 11-fold models. J. Mol. Biol.* **27**, 535–548.

Arnott, S. & Wonacott, A. J. (1966). *The refinement of the crystal and molecular structures of polymers using X-ray data and stereochemical constraints. Polymer*, **7**, 157–166.

Atkins, E. D. T. (1989). *Crystal structure by X-ray diffraction*. In *Comprehensive Polymer Science*, Vol. 1. *Polymer Characterization*, edited by G. A. Allen, pp. 613–650. Oxford: Pergamon Press.

Barham, P. J. (1993). *Crystallization and morphology of semicrystalline polymers*. In *Materials Science and Technology. A Comprehensive Treatment*, Vol. 12. *Structure and Properties of Polymers*, edited by E. L. Thomas, pp. 153–212. Weinheim: VCH.

Baskaran, S. & Millane, R. P. (1999*a*). *Bayesian image reconstruction from partial image and aliased spectral intensity data. IEEE Trans. Image Process.* **8**, 1420–1434.

Baskaran, S. & Millane, R. P. (1999*b*). *Model bias in Bayesian image reconstruction from X-ray fiber diffraction data. J. Opt. Soc. Am. A*, **16**, 236–245.