ELSEVIER

S0032-3861(96)00507-1

Polymer Vol. 38 No. 2, pp. 247–253, 1997
Copyright © 1996 Elsevier Science Ltd
Printed in Great Britain. All rights reserved
0032-3861/97/\$17.00 + 0.00

polymer papers

Direct determination of polymer crystal structures from fibre and powder X-ray data

Douglas L. Dorset
Electron Diffraction Department, Hauptman-Woodward Institute, 73 High Street, Buffalo, NY 14203-1196, USA
(Received 12 December 1995; revised 23 February 1996)

Powder and fibre X-ray diffraction intensities, respectively from two synthetic polymers, polyethylene and poly(ϵ-caprolactone), were used to evaluate the prospects for structure analysis via direct methods. For the former, centrosymmetric, example, an entirely correct phase set was found when the origin-defining reflections were extended by the Sayre equation. Prediction of structure factor magnitudes by this equation also successfully deconvoluted clusters of overlapping reflections into their component values. In the analysis of the latter, non-centrosymmetric, problem with the Sayre equation, a very accurate phase set could also be derived, particularly when the initial model was improved by Fourier refinement (final mean error: 19°). Again, direct prediction of structure factor magnitudes was helpful for the separation of overlapping reflections during the course of the analysis. Copyright © 1996 Elsevier Science Ltd.

(Keywords: direct methods; crystal structure analysis; X-ray diffraction)

INTRODUCTION

Although electron diffraction techniques have shown considerable promise recently as a means for the determination of polymer structures[1], it must be recognized, nevertheless, that most of our understanding of polymer crystal packing has come from the analysis of fibre or powder X-ray diffraction patterns[2,3]. Because of the enhanced scattering cross-section of matter for electrons, individual chain-folded lamellae (and/or epitaxially oriented samples) can be used for collection of truly single crystal diffraction data. For oriented polymer fibres, on the other hand, the data set is cylindrically averaged, resulting in the overlap of reflections from reciprocal lattice vectors with nearly the same length. This reflection overlap can be more extensive for non-oriented powder samples. Given this problem, electron diffraction information has facilitated the identification of space group symmetry and the indexing of fibre patterns[3] so that quantitative structure analyses could be carried out with strictly kinematical X-ray intensities.

Structure determination with fibre X-ray intensity sets is still mostly model-based[2,3], even though Patterson functions have been useful for finding key intermolecular vectors[4]. In this approach, the conformation of stable regions of the polymer repeat are deduced from single crystal structures of monomers and oligomers, and the structure search around defined 'linkage-bonds' seeks simultaneous minima of the crystallographic residual and an atom–atom non-bonded potential energy[2,5].

Despite a tradition of relying on initial structural models, the idea of solving crystal structures from overlapped intensity sets by direct methods is no longer regarded to be a far-fetched concept. Interest has grown in the solution of small structures from powder X-ray diffraction data[6], for example, and methods have been proposed for the direct deconvolution of the overlapped reflections. The greatest difficulty envisioned in such studies is the one imposed by poor data resolution, a barrier that increasingly can be overcome experimentally by the use of synchrotron X-ray sources[7].

Probably because of the convenience of established methods for structure determination[8], there seems to be almost no effort currently in applying direct methods to the analysis of X-ray data from disordered polymer preparations. Although low-resolution or extensively-overlapped data sets will continue to be best approached by the model refinement method, there are numerous other polymer data sets, especially from synthetic materials with rather simple monomer units, where the relative fraction of overlapped reflections is small and where the resolution of the data set is large enough, in principle, to permit the visualization of atomic positions in the electron density maps, after observed structure factor amplitudes are assigned phases by techniques commonly employed in small molecule crystallography. This paper reports an initial effort in this direction, based on two published X-ray data sets, respectively, a powder pattern from polyethylene[9] and a fibre pattern from poly(ϵ-caprolactone)[10].

DATA AND METHODS

Data

X-ray intensities from linear polyethylene, isothermally crystallized from the melt, were obtained originally by Kawaguchi *et al.*[9]. (Although no photograph of the diffraction pattern was published, it is apparent that this is a powder data set.) The centrosymmetric orthorhombic space group is *Pnma* with cell constants: $a = 7.40$, $b = 4.93$, $c = 2.54$ Å. For the 28 possible unique *hkl* data in the set observed to 1.3 Å resolution, 15 reflections were observed independently, while the rest were found in six overlapped clusters. The overall temperature factor was assumed to be $B = 5.8\,\text{Å}^2$, as stated in the original paper.

The second data set, collected by Chatani *et al.*[10] from uniaxially oriented fibres grown from a quenched melt, was from poly(ϵ-caprolactone). This polymer crystallizes in the non-centrosymmetric space group $P2_12_12_1$, with cell constants: $a = 7.47$, $b = 4.98$, $c = 17.05\,\text{Å}$. Of the 108 unique reflections reported to $1.2\,\text{Å}$ resolution, 20 were overlapped doublets of reflections with nearly the same reciprocal spacing. In initial work, the overall temperature factor was assumed to be $B = 8.0\,\text{Å}^2$, the value found for most of the atoms in the original analysis.

For the structure determination, it was important that the observed data should be resolved as individual reflections, instead of the clustered intensities carried through many typical model-based analyses. Encouraged by the success of direct methods when electron diffraction intensity data are perturbed somewhat by multiple scattering[1], it was felt that some error in the observed intensities would not necessarily comprise the outcome of the structure analysis. Thus, an equal distribution of known contributors was assumed initially for any overlapped intensity cluster to provide an intial estimate of observed structure factors. In this work, it was assumed that $I_{\text{obs}} = \Sigma m_i F_i^2$, where $|F_i|$ are the structure factor moduli that would be observed if the individual reflections could be resolved[2] and m_i are their multiplicities. To generate the initial data set for direct phase determination, the observed intensities were split into i equal components according to the values of m_i, that can be obtained from standard tables for any unit cell type[2]. A more accurate segregation of these overlapped reflections into their intensity components was sought as the analysis progressed, as will be described below. Given estimates for all the F_i^2, normalized structure factor magnitudes[11] were calculated from them with the usual formula: $|E_i|^2 = F_i^2/\epsilon\Sigma f_i^2$. Here, f_i are the atomic scattering factors corrected for thermal vibration and ϵ is a correction factor to compensate for certain classes of reflections, e.g. those where systematic absences occur.

Direct phase determination

Before the analysis began, it was important to map out the most efficient way of finding unknown phase values from the known values obtained in the process of defining the unit cell origin (and, when permitted, by setting an enantiomorph for a non-centrosymmetric structure, or, additionally, from the most probable values provided from estimates of Σ_1-phase invariant sums: $\psi = \phi_{h_1} + \phi_{h_1} + \phi_{-2h_1})$[11]. This procedure, involving the construction of a so-called convergence map[7], was started after the generation of all Σ_2-phase invariant sums[11]: $\psi = \phi_{h_1} + \phi_{h_2} + \phi_{h_3}$, where the sum of Miller indices $\Sigma h_i = \Sigma h_i k_i l_i = 0$. Simply stated, possible basis phase sets, made up of the origin-defining reflections, as well as phased amplitudes found by other means, were tested to find which combination yielded phase values for the most number of unknown terms in the data set. The convergence procedure actually depicts the sequence of reflections phased from any given starting point.

While the convergence map would be of direct use if the phase expansion were to be carried out with the tangent formula or symbolic addition, it was only used in this study as a method for finding the most useful starting point for phase extension with the Sayre equation[12]:

$$F_h = \frac{\theta}{V}\sum_k F_k F_{h-k},$$

a simple convolution of phased structure factors. (The term θ is a known function of the atomic scattering factors.) Defined values of F_k (i.e. amplitudes and phases given by origin definition, etc.) were extended into the unknown part of the data set by repeated convolution cycles. The additional inclusion an algebraic unknowns, $\phi_h = a, b, \ldots, n$, generated a number of possible solutions by this convolution operation (after permuting these terms through test values). Initially, undefined reflections were assigned zero amplitude and phase values. When a new phase value was assigned, it was then associated with the observed structure factor magnitude in the next cycle. Also, in the convolutions, which were carried out with F_h values (E_h was also possible), it was necessary to include the F_{000} term to ensure that the starting phases and would remain near their defined values.

Deconvolution of reflection intensities

Recently, several methods have been suggested for deconvoluting the overlapped intensities into their components, based on relationships often employed for direct phase determination[13]. In this paper, the obvious utility of the Sayre equation to predict structure factor amplitudes will be demonstrated to be an effective way of deconvoluting these maxima.

STRUCTURE ANALYSES

Polyethylene

From the list of largest magnitude $|E_h|$ terms, it was shown, via the convergence map, that the origin-defining combination of indices ϕ_{520}, ϕ_{110}, and ϕ_{231} would access 24 of 28 reflections. In keeping with an often-used origin for this structure[14], the values π, 0, π were given (although they each could equally well have been assigned any value within 0, π). In addition, four phases were defined by Σ_1-triples with a high probability, i.e.: $\phi_{200} = 0$; $\phi_{220} = 0$, $\phi_{002} = \pi$, and $\phi_{020} = 0$. Starting with these seven unique reflections as the starting set in the Sayre equation, 13 additional values were found after the first convolution cycle and then all assignments were made after the second cycle. As shown in *Table 1*, there were no phase errors (comparing to phases from a structure factor calculation based on the known methylene subcell model[14]).

In order to estimate the magnitudes of the individual contributors to overlapped reflections, two more cycles of the Sayre equation were carried out. Initially an average value was given to $|F_h|$ for reflections in each cluster, but, after the first cycle with complete phases, new estimates, locally scaled to the weighted sum of intensities for the cluster, were obtained for refinement in the second cycle. As shown in *Table 2*, the predicted

Table 1 Phase estimates for polyethylene

hkl	$\|F_0\|$	ϕ^a	Cycle 1	Cycle 2
110	17.2	0	0	
200	14.1	0	0	
210	3.3	π		π
020	7.9	0	0	
120	5.0	π		π
011	5.7	π	π	
310	4.0	0	0	
111	3.2	π		π
210	5.1	π		π
220	3.5	0	0	
211	3.0	π	π	
400	3.3	0	0	
320	2.2	π	π	
410	3.2	π	π	
121	3.2	π	π	
311	3.3	π	π	
130	2.5	0	0	
221	2.5	π		π
230	2.6	π		π
401	2.6	π		π
510	1.6	0		0
321	1.6	π	π	
411	1.6	π	π	
330	1.9	0	0	
031	1.9	π	π	
231	2.2	π	π	
520	3.4	π	π	
002	3.4	π	π	

[a] From structure factor calclation

Table 2 Estimates of individual structure factor magnitudes for overlapped polyethylene reflections

hkl	$\|F_k\|^a$	Cycle 1	Cycle 2
410	1.7	1.8	2.0
121	4.1	4.1	4.0
130	2.6	3.4	2.8
221	2.3	0.8	2.0
230	2.0	1.7	2.1
401	3.5	3.8	3.4
510	1.2	1.3	1.1
321	2.0	1.9	2.2
411	1.1	1.1	0.9
330	0.8	1.2	0.9
031	3.1	2.8	3.1
520	2.6	3.4	3.3
002	4.6	3.4	3.5

[a] Values predicted by Kawaguchi *et al.*[9]

values agreed well with the ones estimated by Kawaguchi *et al.*[9] from a model-based weighting scheme

$$|F_{\text{obs}}(hkl)| = (KI_{\text{obs}} m_{hkl} |F_{\text{calc}}(hkl)|^2 / \sum_{h'k'l'} m_{h'k'l'} \times |F_{\text{calc}}(h'k'l')|^2)^{1/2}$$

Electron density maps calculated from these phased and deconvoluted structure factor amplitudes clearly revealed the polyethylene structure at atomic resolution (*Figure 1*). Although the hydrogen atom positions could not be discerned, the carbon atom coordinate at $x/a = 0.043$, $y/b = 0.072$, $z/c = 0.25$, corresponded to a C–C bond distance of 1.59 Å and a C–C–C bond angle of 106°, i.e. values that were slightly distorted[15]. The observed chain setting angle, calculated from this coordinate, was 42°. (Slightly higher setting angle values were estimated in the initial study[9].)

Poly(ϵ-caprolactone)

Various origin and enantiomorph combinations were considered for the data set from poly(ϵ-caprolactone)[10]. The origin defining reflections: $\phi_{110} = \pi/2$, $\phi_{037} = \pi/2$, and $\phi_{207} = -\pi/2$ were chosen (again, only to preserve the origin of the previous determination) with the enantiomorph reflection $\phi_{307} = -\pi/2$. From a highly probable Σ_1-triple, $\phi_{200} = \pi$ was accepted. An algebraic phase term (correct value $-\pi/2$) was also assigned to ϕ_{201}. (As will be shown in a new analysis of electron diffraction data from this polymer, after permuting the algebraic term through its possible values, the correct value can be found in the one of two maps with the greatest 'peakiness', e.g. as defined by Stanley[16].) After three cycles of the Sayre equation a phase set was obtained with an overall mean error of $\langle|\Delta\phi|\rangle = 55°$. (If, instead of the algebraic term used for this expansion, alternative test values were given to ϕ_{111} and ϕ_{117}, necessitating the interpretation of 16 maps after their permutation, the Sayre expansion was slightly more accurate, leading to a mean phase error of 47°.) For the expansion of only centrosymmetric starting phases, predicted non-zonal reflections, which were still assigned centrosymmetric values after the first cycle, were off-set by 1° after the first cycle, to permit general phase values to be sought in successive convolutions. (This was not necessary when two non-zonal reflections mentioned above were assigned algebraic terms, that were then each cycled through $n = 4$ steps of $\pi/4 + n\pi/2$.) Also, it may have been profitable to carry out a 'block cycling' of phases, especially for zonal reflections, i.e. to keep the initial predicted value for such reflections, while the general phase terms were allowed to refine. Acceptance of all changes to centrosymmetric reflections accounted, in part, for the error in the initial phase set (generated via one algebraic unknown). (Block cycling had been used in the above analysis of polyethylene.)

Electron density maps were calculated from the first phase set, in which individual peak positions could be discerned. Initially, the [100] projection was calculated to determine the z/c coordinates of the chain. With calculated electron density maps based on the three-dimensional phases to visualize density in slices at levels indicated by the previous projection, the atomic positions of the chain were readily discerned. An approximate position for the carbonyl oxygen was found initially from elongation of one of the atomic sites. The initial set of atomic coordinates led to a model with a mean phase error of 29° for all 108 reflections. At

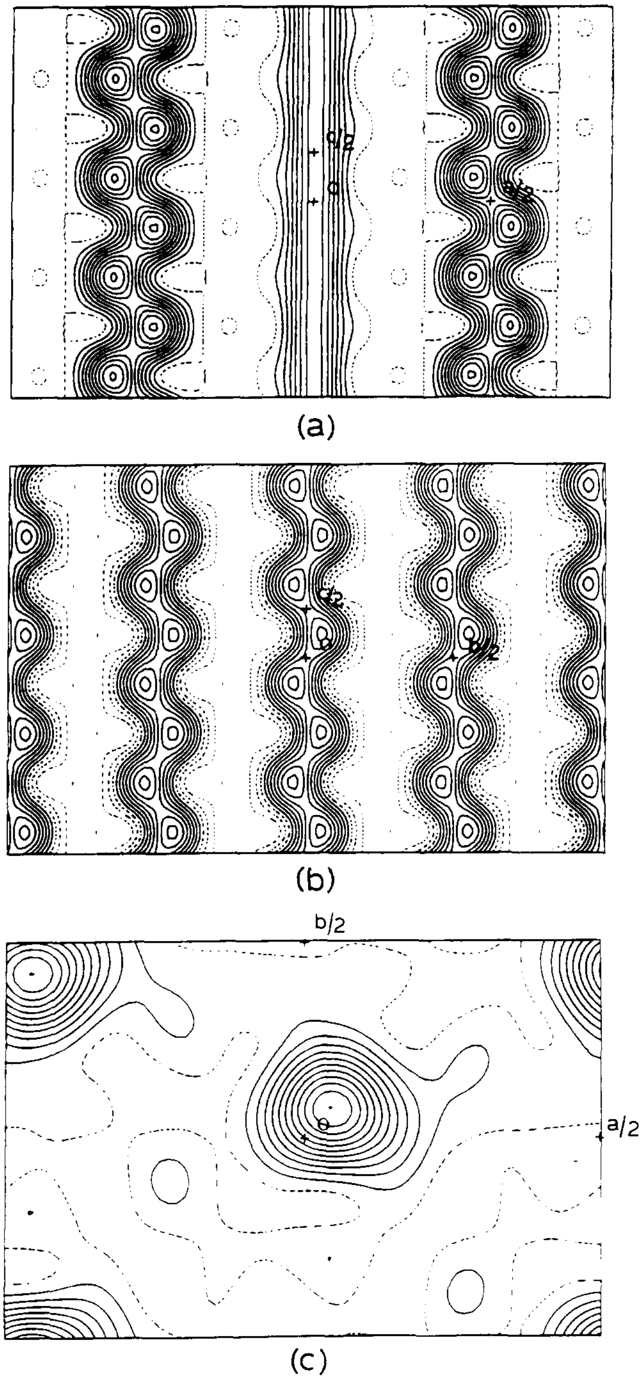


Figure 1 Electron density maps for polyethylene after direct phasing of experimental fibre diffraction data. (a) [110] projection; (b) [010] projection; (c) projection down the chain axis, section at $z/c = 0.25$. (Contour intervals of electron density are arbitrary)

Table 3 Estimates of individual structure factor magnitudes for overlapped reflections from poly(ϵ-caprolactone) fibres

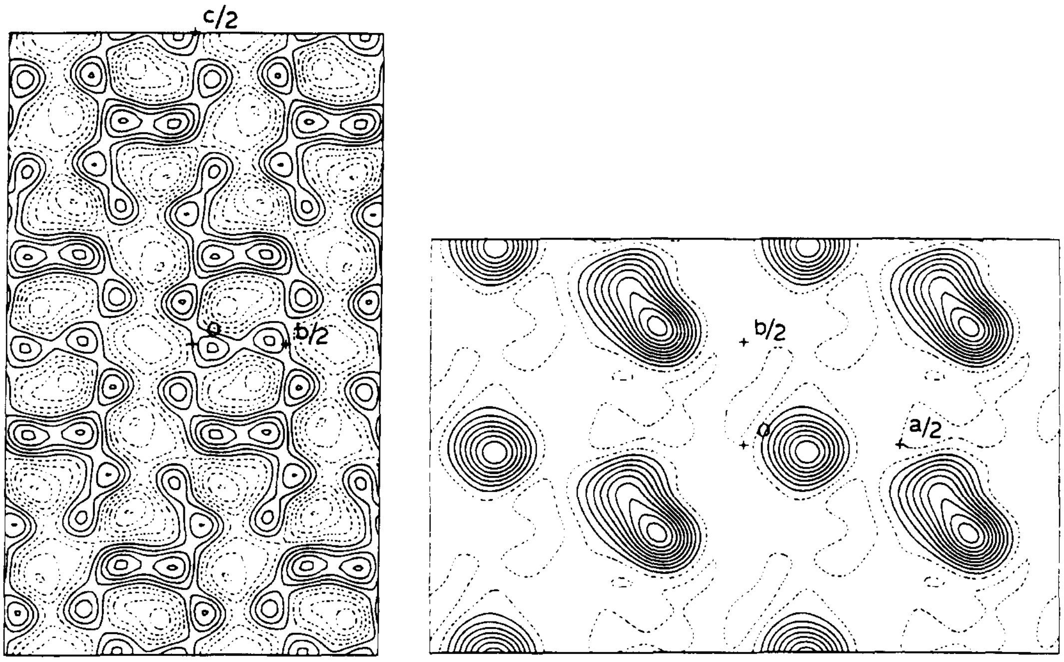
hkl	$\lvert F_c \rvert$[a]	Sample 1[b]	Sample 2[c]
320	0.80	0.73	0.73
410	0.45	0.52	0.57
301	0.40	0.60	0.29
021	0.92	0.72	1.13
321	0.87	0.12	1.04
411	0.81	1.56	0.64
322	0.84	0.72	0.83
412	0.49	0.61	0.55
303	0.55	0.19	0.73
023	0.25	0.61	0.13
304	0.25	0.72	0.38
024	0.68	0.21	0.64
324	0.45	0.25	0.42
414	0.24	0.44	0.30
306	0.55	0.53	0.85
026	0.70	0.72	0.41
307	0.90	0.81	0.61
027	0.90	0.99	1.19
327	1.10	1.12	1.55
417	0.80	0.78	0.37

[a] As separated by Chatani *et al.*[10] based on structural model
[b] After complete Sayre expansion of phases
[c] Sayre re-cycling of phases found after Fourier refinement

this point, phases from the Fourier-transformed map (*Figure 2*) were input to the Sayre equation to predict overlapped structure factor amplitudes (*Table 3*). These amplitude estimates were retained for the next two cycles of Fourier refinement. The first cycle of Fourier refinement based on $|F_0|$ lowered the mean phase error to 25° and this was continued through two more cycles, including a final $2|F_0| - |F_c|$ map, to lower the error to 19°. (An 18° mean phase error had been found after atomic parameters from the map, calculated from the alternate direct phase determination via the (111) and (117) reflections, were used for a structure factor calculation.) At this point, a new prediction of the overlapped structure factor amplitudes was made. The final *R*-factor, based on the atomic coordinates of *Table 4*, was 0.26, compared to the value 0.20, calculated from the originally reported structure[10]. These coordinates, however, corresponded to a structure with slightly distorted bond distances and angles. For example, an average C–O single bond distance was found to be 1.36 Å vs. 1.43 Å as a standard value[15]. Similarly the average C–C distance was 1.46 Å vs. 1.54 Å. The average tetrahedral bond angle was measured to be 115.2° vs. the ideal value of 109.5°. The model could not be improved further just by additional Fourier refinement cycles.

DISCUSSION

From these analyses, it is clear that conventional direct phasing techniques used in small molecule X-ray crystallography can also be quite useful for the determination of polymer structures from fibre or powder diffraction intensities. In these examples, assumption of an equal distribution of intensity over all possible contributors to a reciprocal lattice vector at d_{hkl}, to give initial values for the individual reflections, serves as a useful starting point for the analysis. It is also clear that this approximation is allowed as long as the overlapped reflections do not account for the most intense parts of the diffraction pattern or that they do not represent a large fraction of the observed data. As noted before in the use of direct methods (including Sayre equation) to solve crystal structures from electron diffraction intensities[1], phase estimates via the Σ_2-triple relationship often are useful with somewhat inaccurate structure factor amplitudes as long as the most probable phase invariants are still accessed by large experimental $|E_h|$. The reliance on invariants based on large normalized structure factor magnitudes (or, in this paper, large values of $|F_0|$ in the Sayre convolution, from which the Σ_2 invariant can be derived[17]) permits the analysis to be carried out successfully. On the other hand, similar success will not be expected when other kinds of invariants, such as the negative quartets[11], are used instead.



(a) (b)

Figure 2 Electron density maps for poly(ϵ-caprolactone) after direct phasing of experimental fibre diffraction data. (a) [100] projection; (b) projection down the chain axis, section at $z/c = 0.36$ showing C1 and O1 atoms of carbonyl group. (Maps obtained during Fourier refinement.) (Contour intervals of electron density are arbitrary)

Figure 3 Experimental three-dimensional electron density map for poly(ϵ-caprolactone) with superimposed skeletal model of Chantani *et al.*[10]

Table 4 Final atomic coordinates for poly(ϵ-caprolactone) after Fourier refinement

	Fourier refinement			Structure of Chatani *et al.*[10]		
Atom	x/a	y/b	z/c	x/a	y/b	z/c
C1	0.740	0.556	0.361	0.721	0.566	0.366
C2	0.749	0.606	0.222	0.732	0.615	0.228
C3	0.813	0.450	0.144	0.810	0.471	0.156
C4	0.763	0.576	0.073	0.733	0.604	0.082
C5	0.807	0.436	0.005	0.813	0.449	0.011
C6	0.776	0.446	0.434	0.761	0.408	0.437
O1	0.636	0.770	0.361	0.621	0.735	0.366
O2	0.787	0.466	0.288	0.786	0.481	0.298

Although the observed intensity data are not numerous and are somewhat limited in resolution, the structure derived from the initial phase set can be identified by individual atomic positions in the calculated electron density maps, and the phase accuracy can be improved further by Fourier refinement. Problems with insufficient data and resolution are apparent, nevertheless, since such a refinement will not, by itself, produce the most accurate bond distances and angles. For poly(ϵ-caprolactone), it was clear that the geometrically correct structure still corresponded to a crystallographic residual lower than the final value found in the Fourier refinement. Therefore, if a linked model with ideal bond parameters were fit to the final three-dimensional map, the match itself would produce the desired result. This is demonstrated in *Figure 3*, where the final model of Chatani *et al.*[10] is shown to fit the experimentally determined electron density map very well. The experimental three-dimensional map predicts the non-planar twist to the chain zig-zag in the vicinity of the ester linkage and would, therefore, dictate the conformational geometry for a model constructed with idealized bonding parameters.

It is also interesting to observe that the Sayre equation has performed well as a tool for deconvoluting overlapped reflections into their components. Conceptually this approach is superior to an earlier suggestion that a phaseless convolution of structure factors might be appropriate for the separation of powder diffraction intensities[13]. The phase terms, after all, should always be an important part of the amplitude prediction, as expressed in the Sayre equation. By contrast, a phaseless convolution would also predict large values for systematically absent reflections, e.g. by analogy to the secondary scattering correction used in electron crystallography[11], whereas these forbidden reflections would remain absent after the phased Sayre convolution. Since the polyethylene structure essentially contains only one atom type, which is the most optimal condition for the Sayre convolution to predict accurate new structure factors[12], it is no surprise that the separated amplitudes are quite close to their theoretical values. However, even for the polyester, which contains two components with nearly the same atomic number, there is also good general agreement for many of the reflection pairs.

During a structure determination, this convolution method for finding the separate contributors to an overlapped peak may be preferable to prejudicing this estimate with a preconceived structural model, as has been the usual technique for estimating the individual structure factor amplitudes. The estimate can be revised as the refinement progresses, based on atomic positions identified in experimental electron density maps. The major caveat to applying this method might be to limit it to structures where the components have similar atomic numbers. It is well-known that the Sayre equation loses its accuracy for a wide distribution of atomic scattering factors[18]. Fortunately, this constraint will not present a problem for many polymer applications.

In conclusion, even though the number of measured intensity data from fibre or powder samples is small relative to a small molecule crystal, there may still be enough information about the unit cell transform in these data to permit a direct phase determination and the beginnings of a refinement. At the very least, this will lead to a three-dimensional density envelope giving a highly-constrained fit to a chain model to permit the necessary conformational adjustment to be made *ab initio*. Future work will attempt to sample other representative experimental data sets to gain an accurate overview of the utility of this procedure for structure determination.

ACKNOWLEDGEMENTS

Research described in this paper, which can be regarded as an important offshoot of the direct phasing of electron diffraction data, was supported by a grant from the National Science Foundation (CHE-9417835) which is gratefully acknowledged. I am grateful to Dr Mary P. McCourt for assistance in preparing *Figure 3*.

REFERENCES

1. Dorset, D. L. 'Structural Electron Crystallography', Plenum, New York, 1995
2. Tadokoro, H. 'Structure of Crystalline Polymers', Wiley, New York, 1979
3. Atkins, E. D. T. in 'Comprehensive Polymer Science, Volume 1, Polymer Characterization' (Eds C. Booth and C. Price), Pergamon, Oxford, 1989, p. 613
4. Vainshtein, B. K. 'Diffraction of X-rays by Chain Molecules', Elsevier, Amsterdam, 1966, p. 181
5. Campbell Smith, P. J. and Arnott, S. *Acta Crystallogr.* 1978, **A34**, 3
6. Gilmore, C. J., Shankland, K. and Bricogne, G. *Proc. Roy. Soc. (London)* 1933, **A442**, 97
7. Gilmore, C. J. in: 'Direct Methods of Solving Crystal Structures' (Ed. H. Shenk), Plenum, New York, 1991, p. 303
8. Woolfson, M. and Fan, H. F. 'Physical and Non-physical Methods of Solving Crystal Structures', Cambridge Univ., Cambridge, 1995
9. Kawaguchi, A., Ohara, M. and Kobayashi, K. *J. Macromol. Sci. Phys.* 1979, **B16**, 193

10 Chatani, Y., Okita, Y., Tadokoro, H. and Yamashita, Y. *Polym. J.* 1970, **1**, 555

11 Hauptman, H. A. 'Crystal Structure Analysis. The Role of the Cosine Semivariants', Plenum, New York, 1972

12 Sayre, D. *Acta Crystallogr.* 1952, **5**, 60

13 Jansen, J., Peschar, R. and Schenk, H. *J. Appl. Cryst.* 1992, **25**, 237

14 Abrahamsson, S., Dahlén, B., Löfgren, H. and Pascher, I. *Progr. Chem. Fats Other Lipids* 1978, **16**, 125

15 Ladd, M. F. C. and Palmer, R. A. 'Structure Determination by X-ray Crystallography', Third Edition, Plenum, New York, 1993, pp. 434–435

16 Stanley, E. *Acta Crystallogr.* 1986, **A42**, 297

17 Fan, H. F., Hao, Q. and Woolfson, M. M. *Z. Krist.* 1991, **197**, 197

18 Woolfson, M. M. *Acta Crystallogr.* 1958, **11**, 177