PII: S0032-3861(96)00747-1

Polymer Vol. 38 No. 8, pp. 1985–1988, 1997
© 1997 Elsevier Science Ltd
Printed in Great Britain. All rights reserved
0032-3861/97/$17.00+0.00

Direct phase determination for polymer fibre X-ray data—the structure of poly(tetramethyl-*p*-silphenylene siloxane)

Douglas L. Dorset* and Mary P. McCourt
Electron Diffraction Department, Hauptman-Woodward Institute, 73 High Street, Buffalo, NY 14203-1196, USA

(Received 1 March 1996; revised 11 July 1996)

The X-ray crystal structure of poly(tetramethyl-*p*-silphenylene siloxane) has been re-determined from fibre diffraction data by direct methods (Sayre equation via a multisolution approach). Despite the somewhat low resolution of the observed data set, this approach is sufficient for distinguishing the correct space group as well as producing an electron density envelope that fits the outlines of the monomer repeat. The resultant structural solution is virtually identical to the one found earlier by a model refinement procedure. © 1997 Elsevier Science Ltd. All rights reserved.

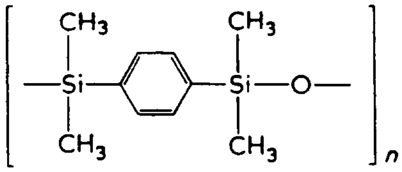
(Keywords: X-ray diffraction; crystal structure analysis; direct methods)

INTRODUCTION

Fibre X-ray diffraction intensity data are frequently used for the determination of polymer crystal structures, although, in some cases, the overlap of reflections caused by the cylindrical distribution of crystallites around the fibre axis may require additional information, e.g. from electron diffraction patterns, to facilitate the identification of space group symmetry and determination of unit cell constants[1]. Crystal structure analyses based on such data sets often are, in reality, refinements, where a constrained chain model, built up of conformationally-fixed monomer units, is flexed around defined linkage bonds to minimize simultaneously the crystallographic residual between calculated and observed data and an atom–atom nonbonded potential energy, corresponding to the chain internal geometry and the packing of these individual chains in the unit cell[1,2].

Although progress has been made recently in the application of direct methods for structure analysis to powder diffraction data from small structures[3], there seems to be almost no comparable effort in the phasing of fibre data for elucidation of polymer structures. Recently a study was carried out[4] with powder and fibre data, respectively, from two relatively simple polymers, polyethylene and poly(ϵ-caprolactone), demonstrating that their crystal structures could be determined directly by conventional approaches via the Sayre equation and also that a good estimate of the separate contributors to overlapped intensity maxima could be made at the same time. For each polymer set the diffraction resolution (e.g. 1.3 Å) was sufficient (via the Rayleigh criterion) to resolve interatomic distances and thus the positions of individual atoms were clearly noted in the ensuing electron density maps.

If the observed fibre diffraction data are not recorded to atomic resolution, it might be imagined that such a direct structure analysis may be frustrated. On the other hand, experience with electron diffraction intensities, where reciprocal space along the direction corresponding to the chain axis is undersampled, has shown that atomic resolution in itself may not be an absolute requirement for the success of direct phasing methods[5]. In an effort to resolve this issue, we report the analysis of the 2.0 Å (diffraction) resolution data set from poly(tetramethyl-*p*-silphenylene siloxane) that had been solved earlier with conventional model-based refinement[6]. The monomer repeat is represented as:



MATERIALS AND METHODS

Diffraction data

The preparation of the polymer and collection of fibre X-ray diffraction data from oriented specimens was described earlier[6]. The experimental patterns actually contained reflections that were well-separated with virtually no overlap. Intensities on film were measured with a flat-bed densitometer with values measured for 34 of 41 unique reflections. For this analysis, an estimated value ($|F_h| = 1000$) was given for a very strong (0 0 4) reflection and default values given to weak maxima ($|F_h| = 20$) that were not originally measured quantitatively (see also original paper[6]). Based on a calcite standard, the tetragonal unit cell dimensions were

---

* To whom correspondence should be addressed

Table 1 Phases determined for poly(tetramethyl-*p*-silphenylene siloxane) by direct methods (values in degrees)

hkl	$\|F\|$	ϕ (model)	ϕ (Sayre)	*hkl*	$\|F\|$	ϕ (model)	ϕ (Sayre)
1 1 0[a]	1000	0	0	3 2 2	50	−129	210
2 0 0	150	0	0	1 0 3	130	−45	135
2 1 0[a]	940	0	0	1 1 3	60	90	−90
2 2 0	330	0	0	2 0 3	40	45	225
3 1 0	180	0	180	2 1 3	310	−112	−184
3 2 0[a]	380	0	0	2 2 3	80	90	90
1 0 1	180	45	45	3 0 3	140	−45	135
1 1 1[a]	440	−90	−90	3 1 3	180	2	149
2 0 1	440	−45	−45	0 0 4[a]	1000	0	0
2 1 1	360	−3	4	1 0 4	100	−90	−90
2 2 1	20	−90	−90	1 1 4	180	0	0
3 0 1	20	45	45	2 0 4	20	0	0
3 1 1	170	−113	30	2 1 4	230	−46	0
3 2 1[a]	250	−30	−30	2 2 4	60	0	0
1 0 2	470	0	180	3 0 4	20	−90	90
1 1 2	170	0	180	3 1 4	70	−20	172
2 0 2	160	−90	−90	1 0 5	20	45	45
2 1 2	230	−158	60	1 1 5	170	−90	−90
2 2 2	160	0	180	2 0 5	170	−45	−45
3 0 2	20	0	0	2 1 5	50	60	15
3 1 2	260	−65	−144				

[a] Initial phase values

measured to be: $a = b = 9.02$, c (fibre axis) $= 15.43$ Å. From systematic absences, the choice of space group is ambiguous[7], i.e. $P4_32_12$, or $P4_12_12$, so the selection of the most appropriate symmetry is also part of testing the efficacy of the direct phasing technique. Each of these symmetry groups leads to eight equivalent positions within the unit cell. Atomic coordinates found in the previous analysis[6] were used to calculate crystallographic phases for comparison with subsequent models indicated by the direct structure determination.

Structure analysis

A Wilson plot[8], made from the observed intensity data, indicated that the overall temperature factor for the molecular model was near $B = 23\,\text{Å}^2$, a value that was used to adjust the scattering factor values for the calculation of normalized structure factors[9]

$$|E|^2 = I_{\mathrm{obs}}/\epsilon \sum_i f_i^2$$

Three-phase structure invariant relationships of the $\sum_2$ type were then generated for the observed data set and, from these, a convergence map was made, after testing how certain combinations of starting reflections would access new phase values in the data set. (The convergence procedure[10] simply establishes the best sequence for finding new phase information, given a small number of previously defined values in a 'basis set'.) After assignment of two origin-defining reflections (only one is allowed in index parity $ug0$ or $gu0$ for the $hk0$ set[11], where 'g' denotes even- and 'u' denotes odd-index values), and an enantiomorphic phase, as well as finding the values of two more reflections from highly-probable $\sum_1$- and $\sum_2$-invariants, the basis set of known phases was expanded to the whole data set via the Sayre equation[12], either stated in its usual form

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

utilizing structure factor values, or in the Hughes[13] form

$$E_h = N^{\frac{1}{2}} \langle E_k E_{h-k} \rangle_k$$

based on the normalized magnitudes. After finding a complete phase set, electron density maps were generated and tested against a geometric model of the polymer monomer unit.

Results

For either $P4_32_12$ or $P4_12_12$, there are two possible reflections that can be used for origin-definition[11]. To begin the analysis, the former space group was chosen. One $hk0$ phase, i.e. $\phi_{210} = 0$, was specified and a general hkl term $\phi_{321} = -30°$. (These could have been any permitted value, e.g. centrosymmetric for the former and a general phase term for the latter, but the assignments were chosen to maintain agreement to the previously found origin to facilitate comparison of results to the earlier determination.) An enantiomorph-defining term[11], permitted for a non-centrosymmetric unit cell[9], was also specified, i.e. $\phi_{111} = -\pi/2$. To complete a basis set to be expanded by the Sayre equation, the phase $\phi_{110} = 0$ was accepted from two $\sum_1$-triples with highest probability and the phase $\phi_{320} = 0$ was found from the $\sum_2$-triple with highest probability. Expanding these into 37 of the 41 reflections of the complete data set required two cycles of the Sayre convolution. After new phases were found in one cycle, their values were held fixed for the next cycle and not changed. The resultant mean phase error was not very promising, i.e. $\langle|\Delta\phi|\rangle = 75°$. Convolutions based on phased values of $|F|$ or $|E|$, moreover, produced equivalent results.

It was clear that another strong reflection was needed in order to find a more accurate phase set. For this purpose the ϕ_{004} was assigned an algebraic value $a = 0, \pi$ to generate two complete phase sets after two cycles of the Sayre equation. When $\phi_{004} = 0$, the mean phase error was 62°C, whereas when it was equal to π, the error was 76°. The match of the phases in the former case to the earlier model is given in *Table 1*.

The correct solution of the structure could be obtained by evaluation of the resultant electron density maps. For example, when the expanded phase set was generated from $a = 0$, a three-dimensional map could be observed (*Figure 1a*) where a very close fit could be made with the

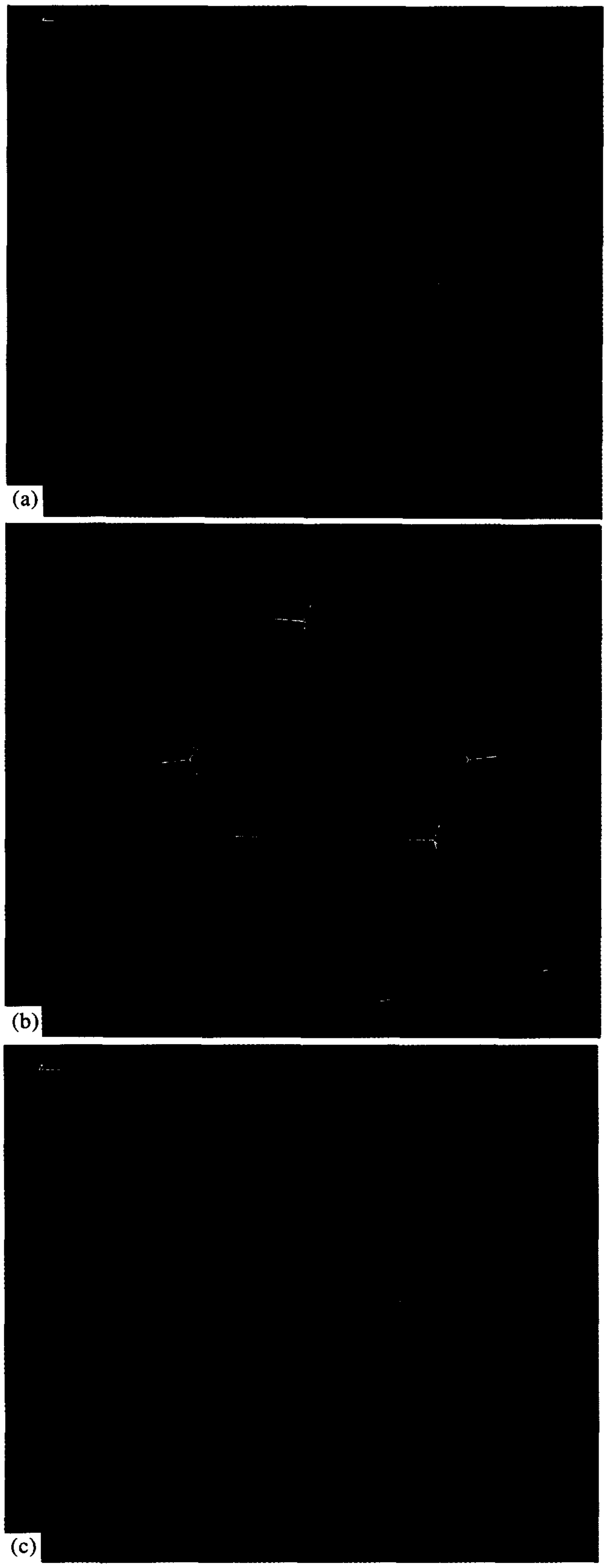


Figure 1 Direct phase determination for poly(tetramethyl-*p*-silphenylene siloxane) in space group $P4_32_12$. (In this and the other figure, three-dimensional electron density profiles generated from the phase determinations are compared to the monomer skeletal model, both in terms of fit and chain connectivity.) (a) Sayre expansion when $\phi_{004} = 0$; (b) Sayre expansion when $\phi_{004} = \pi$; (c) phase set from previous model[6] (see *Table 1*). Note that the continuous electron density in (a) is a good envelope for the previously determined structure model

Figure 2 Direct phase determination for poly(tetramethyl-*p*-silphenylene siloxane) in space group $P4_12_12$. (a) Sayre expansion when $\phi_{004} = 0$; (b) Sayre expansion when $\phi_{004} = \pi$. Note that neither solution results in an electron density profile that closely matches the previously determined structure

polymer monomer unit with reasonable connectivity, even though there was some extraneous density to one side of the phenylene ring. All of the linkage atoms were well fit to density and only the methyl group positions on the silicon atoms lay outside of the density profile. On the other hand, when, $a = \pi$, no connectivity could be established for the polymer chain (*Figure 1b*), so that the solution was rejected.

Despite the low resolution of the original data set, the fit of the monomer model to the electron density map was sufficiently constrained so that the model atomic coordinates were very close to those found in the earlier determination. When these were used in a structure factor calculation, the electron density map generated from the resultant phase set gave a closer match to the model structure (*Figure 1c*). The resultant crystallographic residual for all 41 reflections was $R = 0.29$, when an isotropic thermal parameter $B = 15\,\text{Å}^2$ was used for all atoms.

Another determination was attempted in space group $P4_12_12$, using the initial phase definitions outlined above. Since there was again an ambiguous phase choice for the

(0 0 4) reflection, two solutions were generated with the Sayre equation. Neither of the resultant electron density maps (*Figure 2*) served as a good density envelope for the polymer monomer unit, especially when the connectivity of the molecule along the chain axis was considered. By this criterion, the alternative choice of space group could be rejected.

DISCUSSION

Based on a growing wealth of experience with electron and X-ray diffraction intensity data from linear polymers[4,14], it is clear that the application of direct phasing methods for structure determination is not an unrealistic endeavour, despite the small number of observed data, the undersampling of the reciprocal lattice, or the overlapping of some intensities. For this example, we have successfully tested a data set where resolved atomic positions could not be observed in the final electron density map, even though the molecular shape would still be well-defined. Although the initial phase accuracy was not very good, it was still sufficient to permit calculation of a density map that gave a highly constrained envelope to the monomer model. This is because much of the error was associated with the weakest reflections. (For example, in the above example, the value of $\langle|\Delta\phi|\rangle$ was only 38° for the 14 most intense reflections when $\phi_{004} = 0$, whereas the error for this set increased to 56°, when this algebraic phase estimate was changed to π.)

If this example is typical of other low resolution X-ray data sets, there are precautions to be taken before carrying out a direct phase analysis, however. In the initial use of the Sayre equation on polymer electron diffraction data sets at near atomic resolution[15], it was found to be useful if the restricted, centrosymmetric reflections were assigned phase values first by symbolic addition, by evaluation of individual $\sum_2$-triple invariants. This could help to enlarge the basis set used for the convolutional expansion into the unknown reflections as demonstrated with a number of examples[15]. However the accuracy of such determinations decreases with lower resolution data, particularly when probabilistic estimates for phase invariant sums are consulted. While only one new phase value was assigned in this X-ray determination from the evaluation of a $\sum_2$-triple, based on the invariant sum with the largest value of

$$A = \frac{2\sigma_3}{\sigma_2^{3/2}}|E_h E_k E_{h+k}|$$

it was found that the triple with the next largest A value led to a false estimate for the unknown phase. Further, the correct, but contradictory value for this estimated ϕ_{130} term was seen at a lower A value. In other words, it would not have been possible to assign a phase value to this reflection from such an evaluation because of the reduced accuracy of individual invariants.

It is better, therefore, to rely on the phase estimates produced by a procedure that automatically averages over all possible invariant sum contributors to any reflection. The Sayre equation would be superior to the tangent formula[9], for example, since it is an algebraic approach, whereas the latter is probabilistic, and therefore less likely to be successful for small, low-resolution data sets. (While some inaccuracies are expected from the distribution of atomic numbers for the molecular components, this has not, so far, imposed a severe constraint for many relatively simple structures.) The generation of multiple solutions from a minimal basis set of strong reflections appears to be required, followed by some other criterion for identification of the structure. Since there the molecular profile is still clearly defined in the map, it is easy enough to identify the correct solution by the connectivity of the polymer chain and the fit to the monomer model.

ACKNOWLEDGEMENTS

Research was supported in part by a grant from the National Science Foundation (CHE94-17835) which is gratefully acknowledged. The authors are grateful to Dr K. H. Gardner for suggesting this interesting problem.

REFERENCES

1 Atkins, E. D. T. in 'Comprehensive Polymer Science, Volume 1, Polymer Characterization' (Eds C. Booth and C. Price), Pergamon, Oxford, 1989, p. 613
2 Tadokoro, H. 'Structure of Crystalline Polymers', Wiley, New York, 1979
3 Gilmore, C. J., Shankland, K. and Bricogne, G. *Proc. Roy. Soc. (London)* 1993, **A442,** 97
4 Dorset, D. L. *Polymer,* 1997, **38,** 247
5 Dorset, D. L. and McCourt, M. P. *J. Struct. Biol.* 1993, **111,** 118
6 Gardner, K. H., Magill, J. H. and Atkins, E. D. T. *Polymer* 1978, **19,** 370
7 Hahn, T. (Ed.) 'International Tables for Crystallography, Volume A, Space Group Symmetry', Third Revised Edition, Kluwer, Dordrecht, 1992
8 Wilson, A. J. C. *Nature* 1942, **150,** 151
9 Haputman, H. A. 'Crystal Structure Analysis. The Role of the Cosine Seminvariants', Plenum, New York, 1972
10 Gilmore, C. J. in 'Direct Methods of Solving Crystal Structures', (Ed. H. Schenk), Plenum, New York, 1991, p. 303
11 Rogers, D. in 'Theory and Practice of Direct Methods in Crystallography' (Eds M. F. C. Ladd and R. A. Palmer), Plenum, New York, 1980, p. 23
12 Sayre, D. *Acta Crystallogr.* 1952, **5,** 60
13 Hughes, E. W. *Acta Crystallogr.* 1953, **6,** 871
14 Dorset, D. L. 'Structural Electron Crystallography', Plenum, New York, 1995
15 Dorset, D. L., Kopp, S., Fryer, J. R. and Tivol, W. F. *Ultramicroscopy* 1995, **57,** 59