width is rationalized in terms of the relative insensitivity to the effects of shearing and stacking faults along this plane. Fig. 2(*b*) illustrates the low-temperature structure sectioned parallel to symmetry equivalent (401) planes. It is evident that not only is the (401) plane effectively perpendicular to the molecular axis, it cuts the axis at a near optimum point so as to be relatively insensitive to any along-axis shearing of the structure. Moving away from this plane invokes directions that intersect or are in close proximity to the methyl groups, hence they are likely to be highly disrupted by the suggested stacking fault mechanism and so exhibit line broadening effects.

An initial analysis of the high-temperature rhombohedral phase also shows line-broadening effects and the structural shearing described above may also be invoked to explain the phase transition mechanism. In this model, the highly reconstructive phase transition involves the *a*-axis of the monoclinic phase most likely becoming *c* in the rhombohedral phase.

Preliminary experiments on PDMA were performed at L.L.B., Saclay. One of us (MP) thanks Dr J. Rodriguez for help with this experiment.

References

ALEFELD, B. & KOLLMAR, A. (1976). *Phys. Lett. A*, **57**, 289–290.
DAVID, W. I. F. (1990). *Mater. Res. Soc. Symp. Proc.* **166**, 203–208.
DAVID, W. I. F., IBBERSON, R. M. & MATTHEWMAN, J. C. (1992). Report RAL 92-032. Rutherford Appleton Laboratory.
DAVID, W. I. F. & MATTHEWMAN, J. C. (1988). Report RAL 88-102. Rutherford Appleton Laboratory.
GILMORE, C. J. (1983). *MITHRIL. Computer Program for the Automatic Solution of Crystal Structures from X-ray Data.* Version 1.0. Department of Chemistry, University of Glasgow, Scotland.
IBBERSON, R. M., DAVID, W. I. F. & KNIGHT, K. S. (1992). Report RAL 92-031. Rutherford Appleton Laboratory.
JOHNSON, C. K. (1971). *ORTEP*II. Report ORNL-3794, revised. Oak Ridge National Laboratory, Tennessee, USA.
KITAIGORODSKII, A. I. (1965). *Acta Cryst.* **18**, 585–590.
MIKSIC, M. G., SEGERMAN, E. & POST, B. (1959). *Acta Cryst.* **12**, 390–393.
PAULING, L., SPRINGHALL, H. D. & PALMER, D. H. (1939). *J. Am. Chem. Soc.* **61**, 927–937.
PAWLEY, G. S. (1981). *J. Appl. Cryst.* **14**, 357–361.
PRAGER, M., DAVID, W. I. F. & IBBERSON, R. M. (1991). *J. Chem. Phys.* **90**, 2473–2480.
PRAGER, M., MONKENBUSCH, M., IBBERSON, R. M., DAVID, W. I. F. & CAVAGNAT, D. (1993). *J. Chem. Phys.* **98**, 5653–5661.
SIVIA, D. S. & DAVID, W. I. F. (1994). *Acta Cryst.* A**50**, 703–714.
VISSER, J. W. (1988). *ITO. Autoindexing Program.* Technisch Physische Dienst, PO Box 155, Delft, The Netherlands.
YOST, D. M., OSBORNE, D. W. & GARNER, C. S. (1941). *J. Am. Chem. Soc.* **63**, 3492–3496.

Acta Cryst. (1995). B**51**, 76–80

A New Approach to Structure Resolution by Molecular Dynamics: Variable Force Constants Strategy

BY D. TRANQUI,* J. A. HERMOSO AND M. T. AVERBUCH

Laboratoire de Cristallographie, associé à l'Université Joseph Fourier, CNRS, 25 Avenue des Martyrs, BP 166, 38042 *Grenoble, CEDEX* 09, *France*

AND G. TAMAGNAN, G. LECLERC AND M. CUSSAC

Laboratoire de Chimie Organique, UFR de Pharmacie de Grenoble, Université Joseph Fourier, 5 Avenue de Verdun, BP 138, 38243 *Meylan, CEDEX, France*

(*Received* 22 *March* 1994; *accepted* 14 *July* 1994)

Abstract

Refinement using the 'object-oriented' variation of force constants (VFC) is proven to be an efficient and safe method of refining a molecular structure from a very crude initial model. We illustrate this method by a reinvestigation of the structure of a poorly crystallized compound, *trans*-1,2,3-tris(4-quinolyl)cyclopropane (TQP), which has been recently determined by conventional methods. In the case of TQP, the search model for the VFC method has been obtained by a two-step procedure: (*a*) determination of the most probable conformation of the isolated molecular model by energy-minimization; (*b*) application of the molecular replacement method using the previously computer-designed model as a starting point. The orientation and translation of the molecule were determined, leading to an approximate packing molecular structure. The *R*-factor at

* Author to whom correspondence should be addressed.

this step was 65%. Dynamic simulated annealing refinements from this very crude model using the 'object-oriented' modifications of force constants converged the *R*-factor down to 24%. At this stage, the refinement of the structure was completed by least squares to 7.6%. As expected, the global r.m.s. difference between the two models, obtained by this procedure and by the conventional method, is less than 0.10 Å.

1. Introduction

The difficulty in refining macromolecules lies both in their enormous size and in their inherent flexibility, which results in the possibility of multiple conformations and in the fast intensity decrease at high resolution. Recent years have seen considerable progress in the refinement of macromolecules using restrained least-squares (Jack & Levitt, 1978; Konnert & Hendrickson, 1980; Hendrickson, 1985) or the dynamic simulated-annealing (SA) refinement technique (Brünger, Kuriyan & Karplus, 1987; Brünger, 1990*c*) and recently time-averaged molecular dynamic refinements (Gros, van Gunsteren & Hol, 1990; Clarage & Phillips, 1994). These approaches minimize an objective function, which contains the experimental constraint and stereochemical knowledge terms of the investigated molecule by the following relation

$$E_{\text{tot}} = E_{\text{stereo}} + \omega E_{\text{exp}}, \qquad (1)$$

where E_{stereo} is an empirical energy (or only the squares of the residuals from ideal bond lengths and angles in Hendrickson's procedure) that provides *a priori* information about intra- and intermolecular bonding schemes, and the experimental constraint term E_{exp} is the residual between observed and model-based calculated structure factors, similar to the conventional *R*-factor. The contributions of E_{stereo} and E_{exp} to the total energy are balanced by the weighting factor ω. Thus, in the course of minimization of (1), the E_{exp} term attracts the molecular model toward agreement with the experimental data, while the E_{stereo} empirical energy term may prevent it from adopting unlikely geometries.

Furthermore, from the numerical point of view, SA refinement of (1) presents two major advantages. Firstly, the fact that it incorporates the soft-restrained stereochemical term in the objective function improves the ratio of observations to parameters, which is usually small in most macromolecular structure refinement. Secondly, SA exhibits a larger radius of convergence than that of least squares; it is designed to go uphill and hence, to jump out of the local minimum trap. In other words, the system can adopts a wide range of conformations. The SA algorithm has been intensively applied to NMR and X-ray refinement for various macromolecules such as proteins and nucleic acids (Brünger, Karplus & Petsko, 1989; Fujinaga, Gros & van Gunsteren, 1989).

The success of the SA procedure is critically dependent on the choice of parameters driving the annealing protocol, such as weighting scheme and temperature (Bounds, 1987). Recently, Brünger *et al.* (1990*c*) classify three types of schemes for the control parameters, pertinent to the minimization of (1), in SA refinement: one consists of rescaling the empirical energy term, whilst maintaining constant temperature (Brünger, Krukowski & Erickson, 1990).

In the case of the refinement of a metalloprotein, the ferredoxin from *Clostridium acidurici* (FdCA), using the force constants taken from the literature, led to a highly deteriorated refined model. The starting molecular-replacement *R*-factor of 40% converged to *R*-values only in the vicinity of 25% (r.m.s. difference ≈ 0.56 Å). Finally, the structure of FdCA was successfully refined by modifying the force constants. The modification of force constants was proceeded in such a way it reflects the hierarchy of *a priori* information, *i.e.* the geometries of the iron–sulfur clusters and that of their immediate environments in the case of FdCA protein (Tranqui & Jesior, 1995). Modifying the force constant of pair atoms is, in some extents, equivalent to the modification of the weighting scheme, but it is more 'selective' in the sense that it allows to privilege and/or to preserve the structural information on specific regions rather than other areas in the molecule. This success incites us to develop the VFC refinement method.

In the present paper, the efficiency of this approach is illustrated by the refinement of the molecular structure of *trans*-1,2,3-tris(4-quinolyl)-cyclopropane using a very coarse starting atomic set of coordinates derived from molecular modeling.

2. Obtaining the approximate molecular structure

The structure of TQP was partially determined by the direct-methods *MULTAN*77 program (Main, Lessinger, Woolfson, Germain & Declercq, 1977). This crystal is monoclinic, space group *C*2/*c*, with $a = 26.197$ (16), $b = 7.910$ (5), $c = 23.827$ (10) Å, $\beta = 96.11$ (3)°, $V = 4911$ (8) Å^3, $Z = 8$. Owing to the poor quality and incompleteness of its X-ray data and the presence of disordered solvent in the crystal, many manual interventions were necessary to obtain the complete solution of this structure. The *R*-factor was subsequently refined down to 6.5% for 2810 reflections with $F^2 > 4\sigma(F^2)$ (Tamagnan *et al.*, 1994).

In the following section, we refer to this structure as model **a** [Fig. 1(*a*)], which will be later referred for

comparison, but was not at all used as an indication for rotational and translational searches.

The initial model of the isolated molecule used for SA refinement was that derived from computer modeling by the application of the *SYBYL*6.03 package (TRIPOS Associates, 1993). Fig. 1(*b*) represents this energy minimization model (referred as model **b**). It shows a fair similarity with **a**, considering the general conformational aspect. However, a close inspection of their geometrical features reveals substantial disagreements between X-ray and energy minimization models: the orientations of the three quinoline planes in **a**, Q_1, Q_2 and Q_3 [where Q_1, Q_2, Q_3 are the atom groups defined by N(1)—C(1)···C(9), N(2)—C(10)···C(18) and N(3)—C(19)···C(27), respectively], with respect to that of the cyclopropane [C(28), C(29), C(30)], are 54.5, 66 and 57.9°, respectively, while those in **b** are 85.8, 36.8 and 30.2°. Clearly, the angle formed by Q_1 and Q_2 is more open in **a** than in **b** (see Fig. 1). However, at this stage, it is still unclear whether these different values are due to the inappropriate force constants used in the energy calculation and/or to the fact that the intermolecular packing energy was not taken into account in the modeling of the isolated molecule.

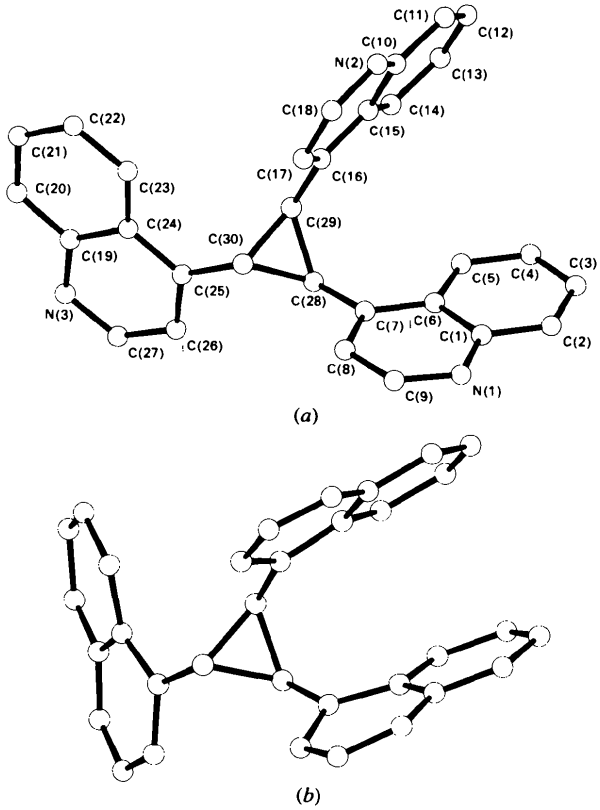

Fig. 1. Molecular conformation [*PLUTON*92; Spek (1992)] of *trans*-1,2,3-tris(4-quinolyl)cyclopropane: (*a*) model observed by X-ray, (*b*) model obtained by energy minimization.

Rotation search was applied [*X-PLOR*; Brünger (1990*b*)] using 1327 reflections included in the range 1.5–15 Å; the Patterson radius and angular increment were 15 Å and 3.0°, respectively. The rotation search in the domains $0 \le \theta 1 \le 360°$, $0 \le \theta 2 \le 90°$ and $0 \le \theta 3 \le 360°$ (Rao, Jih & Hartsuck, 1980) and subsequent PC refinement (Brünger, 1990*a*) was proceeded in three steps: first the molecule was kept rigid, then it was divided into three rigid groups and finally into four groups in order to gain more flexibility. This strategy produced a single value for the orientation of the molecule.

Transition searches (TS) were performed (using the data range 1.5–15 Å and the grid sizes were 1/24, 1/8 and 1/28 along x, y and z, respectively). Figs. 2(*a*) and (*b*) show the molecular packing of model **a** and the initial model **b**, respectively. The r.m.s. difference between the two models is quite large, 1.32 Å, reflecting the high value of the agreement factor to the X-ray data for model **b** (65%). This value is not very far from that estimated from a centrosymmetric random model [83%; Wilson (1950)] and is therefore generally considered as too crude a model for successful least squares of SA refinements using fixed force constants. A new protocol for AS refinement using variable force constant interaction (AS-VFC) is used in the following section to retrieve the molecular structure from model **b**.

3. SA–VFC refinements and results

Model **b** was placed in the unit cell according to the orientation and translation vector found in the previous step [Fig. 2(*b*)], which corresponds to an *R*-factor of 64.9%. The first refinement attempts of this model by least-squares [*SHELXL*93, Sheldrick (1993)] led, as expected, (1) to a highly deteriorated model and (2) to a high *R*-factor, indicating that the refined model **b** is trapped in a false minimum in which the least-squares iterations, due to their limited convergence radius, were unable to jump out over the barrier.

A second attempt consisted of rigid body refinements by conjugate gradient minimization (Powell, 1977) using *X-PLOR* (Brünger, 1990*b*): the procedure stopped at a very poor convergence level (*R*-factor of 64.2%). Additional flexibility was included in the molecule through constant forces taken from the parameter set *PARAM*19*X* used in *X-PLOR*, version 2.1. This is essentially the original *CHARMm PARAM*19 set (Brooks *et al.*, 1983). This latter attempt resulted in an irreversible splitting of the molecule into separate parts and the global target function was stuck in an unusually high *R* minima (57.6%). Clearly, the gap between the packing model (obtained by molecular modeling and molecular replacement) and the actual TOP molecular structure

was still too large to be efficiently refined by the above strategies.

The basic idea we propose to extend the convergence radius of SA refinement, *i.e.* to overcome the false minimum barrier, is to apply a weighting scheme to the force constants involving bond distances and bond angles in the E_{stereo} empirical energy calculation. The 'constant' values were actually modulated in terms of the stereochemical knowledge of the part of the molecule in order to preserve the effect of over-refinement, for example, the force constants involving atoms of the cyclopropane group were artificially increased to prevent their well known geometry from large shifts, while the restraints on the orientations of the quinoline planes in respect to the cyclopropane group were loosened.

In conventional SA refinement protocol, the temperature of the system was increased in order to ensure the atoms overcome the local energy barriers, while in SA–VFC refinement, the gain of kinetic energy was replaced by varying the force constants in maintaining the stereochemical features which are assumed *a priori* to be correct. It is noted that the two procedures are numerically equivalent, since overall scaling can substitute the rising of the temperature (Brünger, Krukowski & Erickson, 1990). However, the SA–VFC method has the advantage over other protocols because it better controls the model changes in the course of refinement.

Figs. 3(*a*) and (*b*) show the variations of the *R*-factor and global r.m.s. differences *versus* refinement steps.

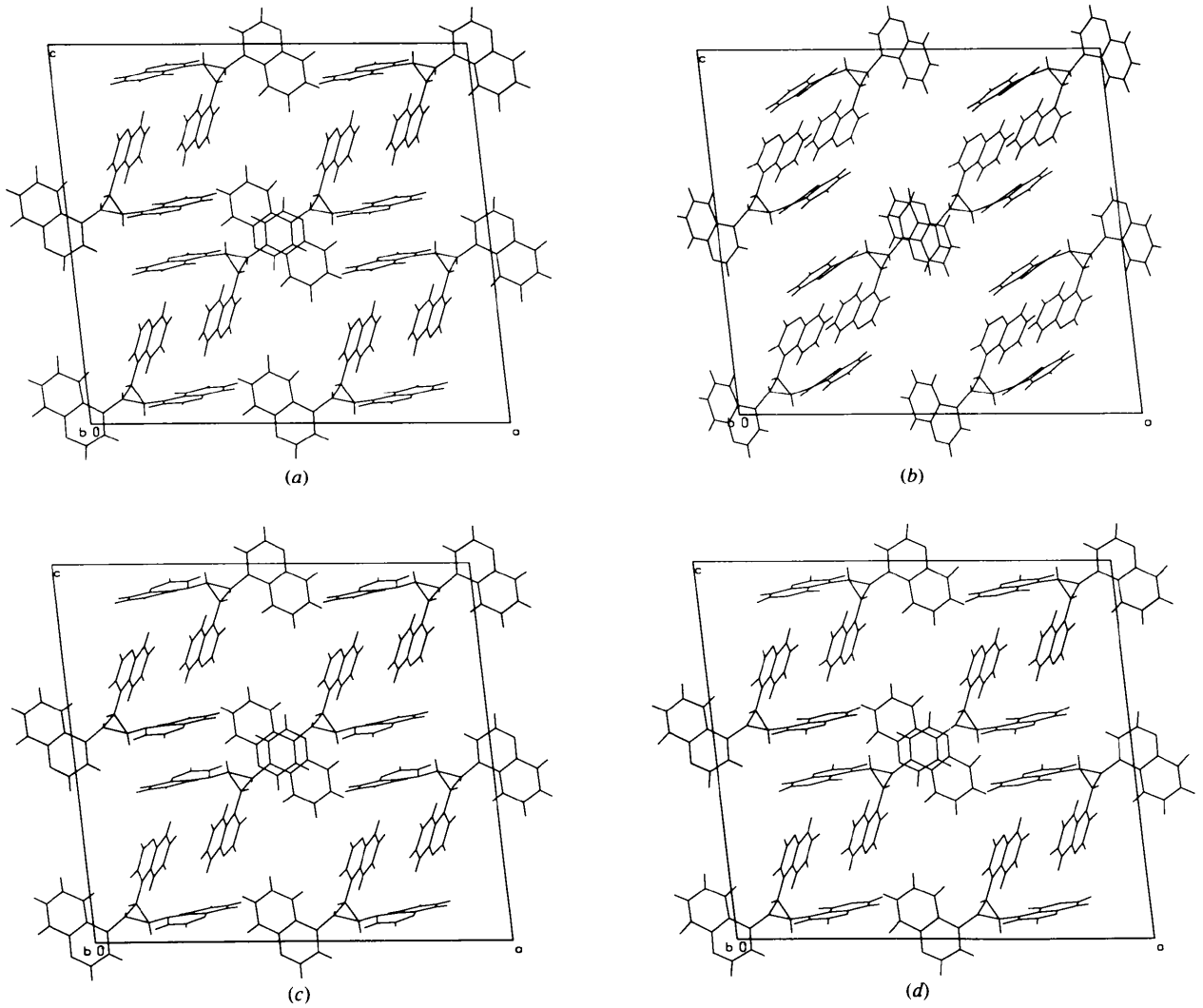

Fig. 2. Molecular packing [*PLUTON*92; Spek (1992)] of TQP obtained (*a*) by direct methods and a model obtained by computer modeling; (*b*) after rotational and translational searches; (*c*) after SA–VFC refinement; (*d*) final model.

In the case of TQP, the successful refinement strategy consisted of:

(i) Partitioning the molecule in three rigid groups: each group is formed by Q_i atoms and one adjacent C atom of cyclopropane (later referred as C_c). Force constants were taken from the literature (Brooks *et al.*, 1983). At this step, the refined model still gave high global r.m.s. differences (1.0 Å).

(ii) Subsequently, dividing the molecule into four regions, highly restrained: three Q_1, Q_2, Q_3 and one cyclopropane group. Moreover, bond distances C_q—C_c (C_q = C atom of the quinoline group) and bond angles C_q—C_q—C_c were tightly restrained (increments of 44% and 40% for C_q—C_c and C_q—C_q—C_c, respectively), while the force constants of the bond angles C_q—C_c—C_c were down-weighted (diminution of 40%) in order to give the quinoline planes more flexibility to interact each other. It is worth noting that the modification of force constants is crucial for the success of SA–VFC refinement. However, it is still empirically based on refinement results and stereochemical knowledge. This process clearly improves the model in reducing the global r.m.s. difference to 0.38 Å. At this stage, the packing diagram was fairly close to that of model **a** [Fig. 2(*c*)]. The agreement factor was reduced from 60% to 28.9%. In the first 100 steps of region II, it is worth noting the fast reduction of the r.m.s. difference with respect to that of the *R*-factor (Fig. 3).

At this stage, the weight factor ω between E_{stereo} and E_{exp} was optimized and used in an unrestrained conjugate gradient minimization process (Powell, 1977). This procedure (Fig. 3, region III) improves the structure model to an *R*-factor of 24.3%. This value suggests that the current model was within the least-squares convergence radius. The final model was completed by the usual least-squares refinement technique (IV and V), yielding an *R*-factor of 7.6% [*SHELXL*93; Sheldrick (1993)]. Fig. 2 clearly shows a quasi-similarity between the refined model **b** and model **a**. The r.m.s. difference between the two models was 0.1 Å. This value is currently considered as within the error limits due to independently refined structures (Brünger, 1988).

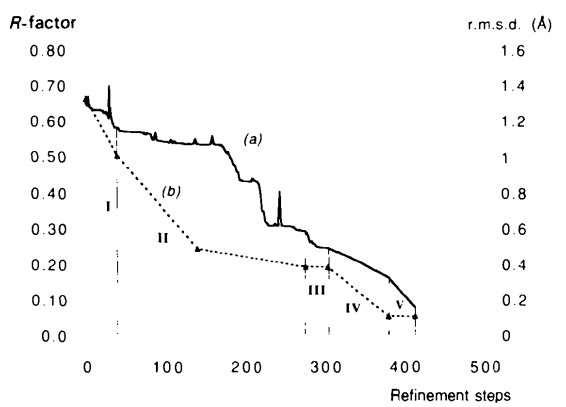

Fig. 3. Dependence of convergence of (*a*) the *R*-factor and (*b*) r.m.s. difference on the refinement strategy: (I) refinement with constants taken from Brooks *et al.* (1983); (II) SA–VFC refinement; (III) conjugate gradient minimization; (IV, V) least-squares refinement.

In conclusion, the proposed approach allows the efficient preservation of the geometrical knowledge of specific parts of the model in keeping the general properties of a simulated annealing refinement.

References

BOUNDS, D. G. (1987). *Nature*, **329**, 215–219.

BROOKS, B., BRUCCOLERI, R., OLAFSON, B., STATES, D., SWAMINATHAN, S. & KARPLUS, M. (1983). *J. Comp. Chem.* **4**, 187–217.

BRÜNGER, A. T. (1988). *J. Mol. Biol.* **203**, 803–816.

BRÜNGER, A. T. (1990*a*). *Acta Cryst.* A**46**, 46–57.

BRÜNGER, A. T. (1990*b*). *X-PLOR Manual.* Version 2.1. Yale Univ., New Haven, USA.

BRÜNGER, A. T. (1990*c*). *Crystallographic Computing School*, pp. 359–372. Bischenberg.

BRÜNGER, A. T., KARPLUS, M. & PETSKO, G. (1989). *Acta Cryst.* A**45**, 50–61.

BRÜNGER, A. T., KRUKOWSKI, A. & ERICKSON, J. W. (1990). *Acta Cryst.* A**46**, 585–593.

BRÜNGER, A. T., KURIYAN, J. & KARPLUS, M. (1987). *Science*, **235**, 458–460.

CLARAGE, J. B., PHILLIPS, G. N. JR (1994). *Acta Cryst.* D**50**, 24–36.

FUJINAGA, M., GROS, P. & VAN GUNSTEREN, W. F. (1989). *J. Appl. Cryst.* **22**, 1–8.

GROS, P., VAN GUNSTEREN, W. F. & HOL, G. J. (1990). *Science*, **249**, 1149–1152.

HENDRICKSON, W. A. (1985). *Methods Enzymol.* **115**, 252–270.

JACK, A. & LEVITT, M. (1978). *Acta Cryst.* A**34**, 931–935.

KONNERT, J. H. & HENDRICKSON, W. A. (1980). *Acta Cryst.* A**36**, 344–349.

MAIN, P., LESSINGER, L., WOOLFSON, M. M., GERMAIN, G. & DECLERCQ, J.-P. (1977). *MULTAN*77. *A System of Computer Programs for the Automatic Solution of Crystal Structures from X-ray Diffraction Data.* Univs. of York, England, and Louvain, Belgium.

POWELL, M. J. D. (1977). *Math. Program.* **12**, 241–254.

RAO, N. S., JIH, J. & HARTSUCK, J. A. (1980). *Acta Cryst.* A**36**, 878–884.

SHELDRICK, G. (1993). *SHELXL*93. *Program for the Refinement of Crystal Structures.* Univ. of Göttingen, Germany.

SPEK, A. L. (1992). *PLUTON*92. *Molecular Graphics Program.* Univ. of Utrecht, The Netherlands.

TAMAGNAN, G., DURET, J. A., LECLERC, G., HERMOSO, J. A., AVERBUCH, M. T., TRANQUI, D. & THOMASSON, F. (1994). *J. Chem. Soc. Perkin Trans.* Submitted.

TRANQUI, D. & JESIOR, J. C. (1995). *Acta Cryst.* D**50**. Submitted.

TRIPOS Associates Inc. (1993). *SYBYL Manual.* Version 6.03.

WILSON, A. J. C. (1950). *Acta Cryst.* **3**, 397.