Acta Crystallographica Section A
**Foundations of Crystallography**
ISSN 0108-7673

Received 5 February 2001
Accepted 28 June 2001



# A modified $X - Y$ method

**Yan Chen and Wu-Pei Su***

Department of Physics and Texas Center for Superconductivity, University of Houston, Houston, TX 77204, USA. Correspondence e-mail: wpsu@uh.edu

The $X - Y$ method seeks to solve the unknown phases of the X-ray reflections by minimizing a function (the $X - Y$ function) of the phases. This cost function has been supplemented with a residual term. The total cost function is minimized by varying the positions of atoms. Simulated annealing is used to implement the minimization. Trial calculations for structures containing up to 176 non-H equal atoms have been carried out successfully.

## 1. Introduction

Many of the powerful direct methods involve filtering in both real and reciprocal spaces. A premiere example is the shake-and-bake method (Miller *et al.*, 1993). In a recent work (Liu & Su, 2000), we have proposed a variant of the method. We bypass the laborious procedure of Fourier transform and peak-picking by focusing directly on atomic positions. Starting from a completely random structure, the atoms are moved around in real space to minimize the crystallographic residual and the Hauptman minimal function (DeTitta *et al.*, 1994). Simulated annealing is employed in the minimization procedure. The results are encouraging.

To further the understanding of this dual filtering, in this paper we explore a different phase-filtering technique, the $X - Y$ method (Debaerdemaeker & Woolfson, 1983, 1989). This method was accidentally discovered to be a more efficient variant of the $X + Y$ method, which is a variational counterpart of the classical tangent formula. It is of interest to see how efficient this technique is when combined with an automated real-space filtering technique.

To illustrate the points, we have carried out trial calculations involving two space groups, $P2_1$ and $P2_12_12_1$. We have used known structures and fabricated data. For $P2_1$, we have been able to solve a structure containing 176 equal atoms. For $P2_12_12_1$, we have been successful with hexadecaisoleucinomycin (HEXIL). The method is very straightforward and can be used as an alternate direct method.

## 2. Methodology

The $X - Y$ function refers to the following summation,

$$R_1 = \sum_{hk} A_{hk}[\cos(\varphi_{hk}t) - \sin(\varphi_{hk})], \qquad (1)$$

where the triple-phase invariants, $\varphi_{hk} = \varphi_h + \varphi_k + \varphi_{-h-k}$, and the concentration parameter, $A_{hk} = 2N^{-1/2}|E_hE_kE_{h+k}|$, are defined in the usual way. $N$ is the number of atoms in the unit cell. Maximization of $R_1$ with respect to the phases leads to the tangent formula. It should be noted that the concentration parameter is calculated from observed $|E|$ values.

In contrast, in a pure real-space approach (Su, 1995; Chen *et al.*, 1997), the following crystallographic residual $R_2$ is minimized with respect to atomic coordinates,

$$R_2(\{\mathbf{r}_i\}) = \sum_{\mathbf{k}} w(E_{\mathbf{k}})(|E_{\mathbf{k}}| - |E_{\mathbf{k}}|_{\rm obs})^2. \qquad (2)$$

In (2), $E_{\mathbf{k}}$ is a normalized structure factor calculated from given atomic coordinates $\mathbf{r}_i$. $w$ is a weight function, which is chosen to be proportional to $E^2$ below.

What we propose to consider in this work is to combine the two cost functions $R_1$ and $R_2$ into a single cost function $R = R_2 - R_1$, and to anneal the atomic coordinates to minimize the total cost $R$. Hereon the methodology closely resembles that of the previous work on the hybrid minimal function (Liu & Su, 2000). In particular, a simulated annealing algorithm (Kirkpatrick *et al.*, 1983) is employed for the minimization procedure. Starting from a random real-space molecular configuration at a high fictitious temperature (the annealing temperature $T$), the atoms are moved around in the unit cell one by one to gradually lower the cost function (Giacovazzo, 1998). If a sensible structure does not emerge at low $T$, the procedure is repeated until it does.

In the following, we report on several trial calculations.

## 3. Examples

### 3.1. Isoleucinomycin ($C_{60}H_{102}N_6O_{18}$)

We have chosen this structure (Pletenev *et al.*, 1980) and HEXIL (Pletenev *et al.*, 1992) for comparison with our previous calculations

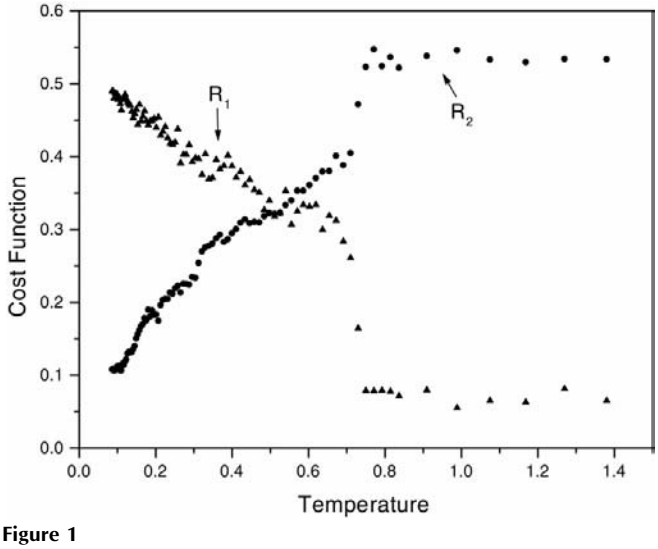


**Figure 1**
Annealing curves for isoleucinomycin.

of the hybrid method (Liu & Su, 2000). The space group is $P2_12_12_1$ with $Z = 4$. H atoms are ignored and the remainder are treated as C atoms. 746 reflections are singled out from the synthetic data of 1.0 Å to compose 7083 triplets. We adjust the scale of $R_1$ so that it contributes by about the same amount to the total cost as $R_2$.

The annealing curves of a successful run displayed in Fig. 1 are characterized by a phase-transition-like feature near $T = 0.8$, *i.e.* there is a sudden change in the magnitude of both $R_1$ and $R_2$. For a typical unsuccessful run, at low $T$, $R_1$ can reach as high as it does in Fig. 1. $R_2$, however, levels off at around 0.25. The fact that $R_2$ continues to dive below 0.25 for $T > 0.3$ is a sure indication of a correct structure.

Other relevant data which pertain to this calculation are (i) each atom is moved 2000 times at each temperature, (ii) 15 runs preceded the first successful run presented above. The entire 16 runs take 8 CPU days on a Digital XP1000 Alpha Station.

The original unmodified $X - Y$ method is powerful enough to solve this structure in about the same amount of computer time. Our previous hybrid method using the Hauptman minimal function seems to be somewhat more efficient than the modified $X - Y$ method for the same structure.

We have also been able to solve this structure utilizing actual diffraction data (Pletenev *et al.*, 1980) in 13 runs (6 CPU days). The corresponding annealing curves are very similar to those in Fig. 1, except that $R_1$ and $R_2$ level off at 0.45 and 0.2, respectively, at low temperatures.

### 3.2. Hexadecaisoleucinomycin (HEXIL) ($C_{80}H_{136}N_8O_{24}$)

HEXIL shares the same symmetry group and $Z$ with isoleucinomycin. 1317 reflections are chosen out of the synthetic 1.0 Å data to compose 10434 triplets. Each atom is moved 2500 times at a fixed temperature. Nine runs take ten days of CPU time on the same machine as described previously. The annealing curve of the last run (the successful one) is displayed in Fig. 2. Compared with Fig. 1, Fig. 2 is almost featureless. Nonetheless, the continuing decrease in $R_2$ for $T < 0.1$ is again a sure sign of a correct structure. An incorrect structure will yield an $R_2$ value above 0.25.

Solving the same structure utilizing real diffraction data (Pletenev *et al.*, 1992) takes a slightly longer time (13 CPU days). On the other hand, the original $X - Y$ method does not yield any sensible result even with synthetic data after two weeks of CPU time.

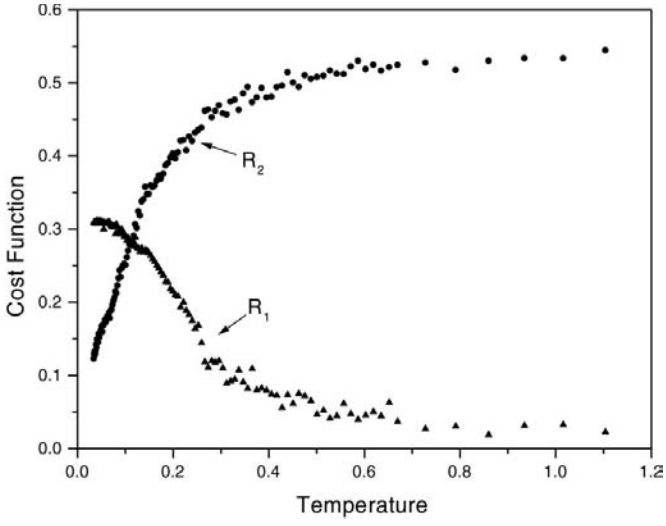


**Figure 2**
Annealing curves for HEXIL.

### 3.3. Valinomycin monohydrate cage complexes

This is a structure of a valinomycin monohydrate cage complex crystallized from dioxane (Langs *et al.*, 1992). There are two valinomycin ($C_{54}H_{90}N_6O_{18}$), two water and three dioxane ($C_4H_8O_2$) molecules in the asymmetric unit cell. The space group is $P2_1$ with $Z = 4$. Again the H atoms are ignored, and the remainder are treated as C atoms in fabricating the reflection data. 1756 reflections are chosen from the 1.0 Å data to make up 7733 triplets. Each atom is moved 3000 times at a given temperature.

After 24 runs, which took 21 CPU days, a successful (the last) one was recorded, whose annealing curves are displayed in Fig. 3. A phase-transition-like feature appears again at about $T = 0.25$. $R_2$ continues to fall below 0.25 for $T < 0.12$, indicating a correct structure (Fig. 4). To contrast with the unsuccessful runs, we superimpose the annealing curves from a typical unsuccessful run onto Fig. 3; the result is shown in Fig. 5, where the hollow symbols (△, ○) represent the unsuccessful run and the solid symbols (▲, ●) represent the successful run.

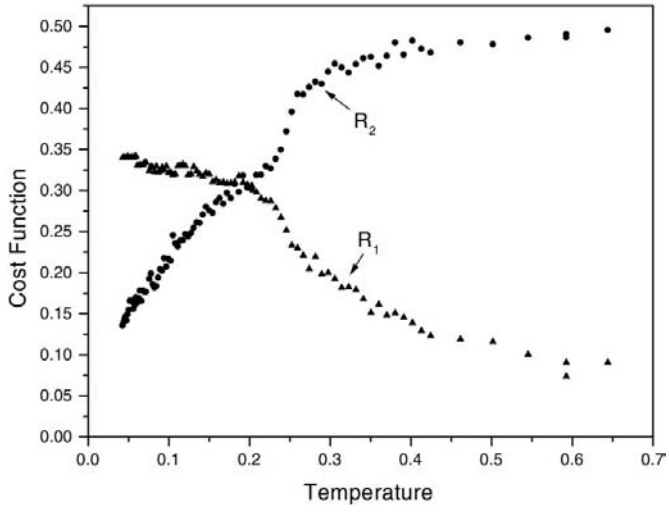


**Figure 3**
Annealing curves of a successful run for valinomycin complexes.

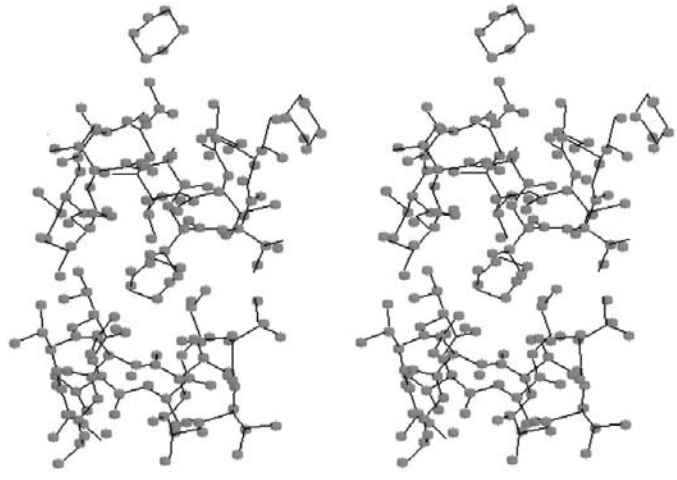

**Figure 4**
Stereoview of the calculated structure of the valinomycin complexes superimposed on the known structure in a wire frame.

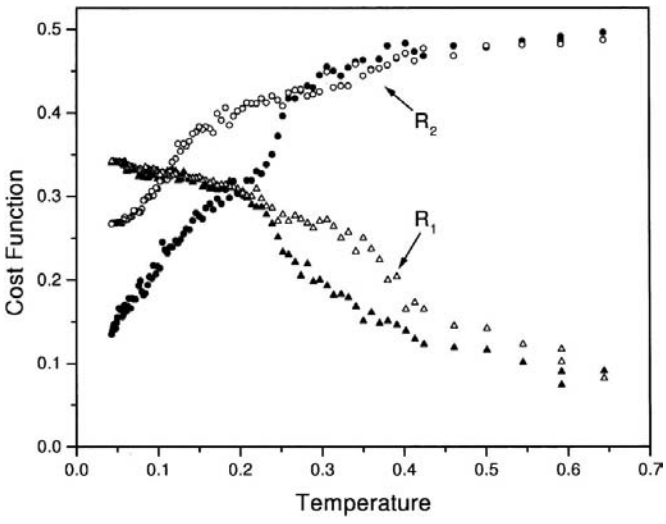


**Figure 5**
Annealing curves of a typical unsuccessful run *versus* those of Fig. 3, the successful run. The black symbols (▲, ●) represent the successful run; the open symbols (△, ○) represent the unsuccessful run.

## 4. Discussion

Owing to the apparent similarity of the present work with some previous publications, two comments are in order for contrast and for clarification. (i) Although our work on Sayre's equation (Chen & Su, 2000) also involves real-space filtering in the form of peak picking, it is performed only at the (low-temperature) end of an annealing cycle. Thus the role of real-space filtering in the formation of a correct structure is rather limited compared to the extensive and important role it plays above. (ii) We have employed simulated annealing in most of our minimization calculations. Annealing has proven to be essential in order to avoid being trapped in local minima. In a similar vein, Sheldrick (1990) has invoked annealing to prevent the tangent formula from converging to an overconsistent phase set.

The $X - Y$ method is a variant of the $X + Y$ method, which is a variational formulation of the classical tangent formula. Real-space filtering has also been part of the classical direct methods. Here we have combined these two tools into a single automated procedure. The trial calculations presented above indicate that this hybrid method is quite efficient considering the simplicity of the algorithm. It is conceivable that the real-space filtering can be combined with other phase-filtering techniques in a similar way and leads to a significant improvement of the original methods.

This work was partially supported by the Texas Center for Superconductivity, the Texas Advanced Research Program under grant no. 003652-707-1997, the Texas Advanced Technology Program under grant no. 003652-0222-1999 and the Robert A. Welch Foundation. We gratefully acknowledge D. A. Langs for useful correspondence and for providing the actual diffraction data.

## References

Chen, Y. & Su, W.-P. (2000). *Acta Cryst.* A**56**, 127–131.
Chen, Y.-S., Su, W.-P., Mallela, S. P. & Geanangel, R. A. (1997). *Acta Cryst.* A**53**, 396–399.
Debaerdemaeker, T. & Woolfson, M. M. (1983). *Acta Cryst.* A**39**, 193–196.
Debaerdemaeker, T. & Woolfson, M. M. (1989). *Acta Cryst.* A**45**, 349–353.
DeTitta, G. T., Weeks, C. M., Thuman, P., Miller, R. & Hauptman, H. A. (1994). *Acta Cryst.* A**50**, 203–210.
Giacovazzo, C. (1998). *Direct Phasing in Crystallography*, pp. 198–201. Oxford University Press.
Kirkpatrick, S., Gelatt, C. D. Jr & Vecchi, M. P. (1983). *Science*, **220**, 671–680.
Langs, D. A., Blessing, R. H. & Duax, W. L. (1992). *Int. J. Peptide Protein Res.* **39**, 291–299.
Liu, X. & Su, W.-P. (2000). *Acta Cryst.* A**56**, 525–528.
Miller, R., DeTitta, G. T., Jones, R., Langs, D. A., Weeks, C. M. & Hauptman, H. A. (1993). *Science*, **259**, 1430–1433.
Pletenev, V. Z., Galitskii, N. M., Smith, G. D., Weeks, C. M. & Duax, W. L. (1980). *Biopolymers*, **19**, 1517–1534.
Pletenev, V. Z., Ivanov, V. T., Langs, D. A., Strong, P. & Duax, W. L. (1992). *Biopolymers*, **32**, 819–827.
Sheldrick, G. M. (1990). *Acta Cryst.* A**46**, 467–473.
Su, W.-P. (1995). *Acta Cryst.* A**51**, 845–849.