Acta Cryst. (1995). D**51**, 853–855

A modified Crowther and Blow T_1 translation function for partial search models.

By CATHERINE L. LAWSON, *Biology Department, Brookhaven National Laboratory, Building* 463, *Upton, NY* 11973-5000, *USA*

(*Received* 14 *November* 1994; *accepted* 14 *February* 1995)

Abstract

A simple modification to the fast T_1 translation function of Crowther & Blow [Crowther & Blow (1967). *Acta Cryst.* **23**, 544–548] can reduce systematic error in cases where the input set of oriented search models represents a fraction of the scattering matter of the unknown crystal unit cell. The observed Patterson function is modified by a partial self-vector subtraction (SVS) and then the residual Patterson origin is removed. In a test case where a protein/DNA complex crystal was searched using a protein-only model, the highest signal-to-noise ratios for translation-search maps were obtained either with origin removal alone or partial SVS combined with origin removal. Origin removal is likely to be a generally useful alternative to SVS for the T_1 function.

Introduction

Molecular-replacement methods employ searches based on the Patterson function to find the orientations and positions of a model in an unknown crystal unit cell (Rossmann, 1972; Lattman, 1985). The success of these methods ultimately hinges on obtaining a consistent set of relative translation vectors between well oriented search models. The fast T_1 function of Crowther & Blow (1967) has traditionally been encoded to handle problems where the set of well oriented search models represents all of the scattering matter of the crystal. For problems where only a fraction of the crystal asymmetric unit can be represented by a known model, *e.g.* the protein of a protein/DNA complex, the traditional function requires a simple modification to minimize systematic error and maximize signal-to-noise (*S/N*).

The T_1 function evaluates, using a reciprocal-space function, the correlation between the calculated Patterson of a pair of search models separated by vector **t** and the Patterson of the unknown crystal. The observed Patterson is modified by subtracting intramolecular (self) vectors of the set of M oriented search models. In reciprocal space, the modification is accomplished by subtracting the sum of calculated intensities, $I_{\rm self} = \sum_{n=1}^{M} I_n$, from each observed intensity, $I_{\rm obs}$. Although Crowther & Blow recommended that observed and calculated intensity data be placed on an absolute scale, the tradition is to set the average calculated intensity $\langle I_{\rm self}\rangle$ equal to the average observed intensity $\langle I_{\rm obs}\rangle$. This approximation is valid only if the input set of M oriented search models fully represents the scattering matter of the crystal unit cell. If the model set represents only a fraction ($p_{\rm model}$) of the scattering matter, the calculated and observed intensity data should be scaled proportionately to avoid over-subtraction of self-vectors (Schierbeek, Renetseder, Dijkstra & Hol, 1985).

With the traditional T_1 function, the self-vector subtraction procedure (SVS) fully removes the observed Patterson origin, *i.e.*, the average modified intensity $\langle I'_{\rm obs}\rangle$ is set to zero. Because the height of the calculated Patterson origin peak increases for incorrect translation positions with molecular overlap (Harada, Lifchitz & Berthou, 1981), the residual observed Patterson origin that results from partial SVS is likely to contribute to map noise rather than to the signal. Residual origin removal may be performed with the following function,

$$I'_{\rm obs}(hkl) = I_{\rm obs}(hkl) - [pI_{\rm self}(hkl)] - [(1-p)\langle I_{\rm obs}\rangle]. \quad (1)$$

This form of the modified observed intensity is essentially identical to those found in translation functions for small to medium-size molecules (Karle, 1972; Langs, 1975; Beurskens, Gould, Bruins Slot & Bosman, 1987).

The T_1 subroutine of the program package *MERLOT, Version* 2.4 (Fitzgerald, 1988) was modified to calculate (1). The modified function was key to the success of molecular-replacement structure determinations for three different protein/DNA complex crystals using protein-only search models (Otwinowski *et al.*, 1988; Lawson & Carey, 1993; Klimasauskas, Kumar, Roberts & Cheng, 1994). In each case, searches with the traditional T_1 function ($p = 1.0$) failed to produce consistent relative translation-vector sets for protein orientations obtained from strong rotation-function peaks. Modified T_1 searches with origin removal ($p = 0.0$) and/or partial self-vector subtraction combined with origin removal ($p \le p_{\rm model}$) improved the signal-to-noise ratio enough to produce clean interpretable maps with consistent translation-vector sets.

Results

The following example illustrates the improvement obtained by use of (1) for modified observed intensities in the T_1 translation search. The 'unknown' is a complex between *trp* repressor and a 16 base-pair DNA duplex (Lawson & Carey, 1993) where the stoichiometry is two protein dimers to one DNA duplex (Carey, Lewis, Lavoie & Yang, 1991; Haran, Joachimiak & Sigler, 1992). The crystal belongs to space group $C2$ and has unit-cell dimensions $a = 112.34$, $b = 90.16$, $c = 58.65$ Å, and $\beta = 113.92°$; there is one complex 'equivalent' in the asymmetric unit. The search model was the first *trp* repressor dimer from PDB entry 1TRO (Otwinowski *et al.*, 1988). Rotation-function searches yielded three strong peaks, leading to three oriented search models for input to the translation function (M1, M2 and M3). Molecules M1 and M3 are in special positions, with molecular dyads coincident with crystallographic twofolds, while M2 is in a general position. Of the six possible pairwise translation searches with three oriented molecules, (M1–M1sym) and (M3–M3sym) produced trivial unit vector solutions; (M2–M2sym), (M2–M1), (M3–M2), and (M3–M1) produced fractional relative translation vector solutions (see Fig. 1 legend for further details). Because conclusions drawn from the latter four searches are more generally applicable to other molecular-replacement problems, the trivial searches are not considered further here.

Varying p and/or removing terms from (1) had significant effects on the interpretability of the set of four pairwise translation vectors. The first panel of Fig. 1 shows that when the observed Patterson was not modified, the correct solution signal (black bar) was hidden beneath noise peaks (white bars) for three out of four searches. Full SVS (*i.e.*, a traditional

T_1 search, panel 2) hid two of the four correct translation vectors. Partial SVS ($p = 0.5$) with the residual origin left in place (panel 3) brought out the correct translation vectors as the highest peaks for all search pairs, and comparable results were obtained with SVS plus residual origin removal (panel 4). Origin removal (panel 5) was as effective as partial SVS in improving S/N.

The apparent equivalence of the three successful search strategies was somewhat puzzling since, from theoretical considerations, SVS with an accurate value for p, plus residual origin removal, should give the highest S/N. Recalculation of each of the four pairwise searches with fine increments in p showed that there was indeed a range of values where partial SVS plus residual origin removal does give slightly higher S/N than origin removal alone, and much higher S/N than SVS alone. Fig. 2 shows the results for search pair (M2–M2sym), where S/N is maximized at $p = 0.2$; the other search pairs give qualitatively similar results but with maximum S/N at values of p ranging between 0.0 and 0.6 (data not shown). Assuming that each atom scatters proportional to the square of its atomic number, $p_{\rm model}$ is estimated to be 0.7. One must therefore conclude that the optimal p for SVS may be significantly less than $p_{\rm model}$.

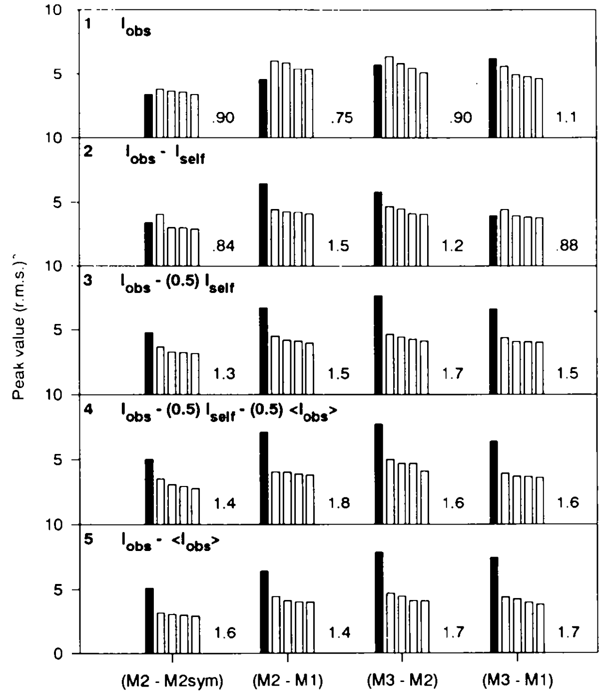

Fig. 1. Comparison of modifications to the observed Patterson function in the T_1 translation function calculated with *MERLOT* (Fitzgerald, 1988). Panel 1: no modification; panel 2: traditional T_1 function; panel 3: partial self-vector subtraction; panel 4: partial self-vector subtraction plus residual origin removal; panel 5: origin removal. Left to right: four independent relative translation searches needed for determining the absolute position vectors of the three molecules M1, M2, and M3. (M2–M2sym) is a two-dimensional search between molecule M2 and its crystallographic twofold symmetry mate; (Mm–Mn) are full three-dimensional searches. Black bars indicate the r.m.s. peak height of the true solution, white bars indicate the r.m.s. values of the first four incorrect (noise) peaks. The signal-to-noise ratio, defined as the r.m.s. value of the solution peak divided by the r.m.s. value of the highest noise peak, is displayed to the right of the noise peaks. For self-vector subtraction calculations, molecules in special positions (M1 and M3) were given half the weight of the one in a general position (M2). Each translation search was summed over 6691 multiplicity-corrected observed reflections between resolution limits 10 to 4 Å. Structure factors of the oriented search models were calculated by interpolation of an initial data set with the unoriented model in a $P1$ unit cell ($a = b = c = 90$ Å, $\alpha = \beta = \gamma = 90°$).

Discussion

The results presented here verify that SVS can be useful for improving the translation map S/N when performed with a sensible scaling ratio (p) between $\langle I_{\rm self}\rangle$ and $\langle I_{\rm obs}\rangle$. SVS is likely to have a detrimental effect, however, when the input value of p is larger than $p_{\rm model}$, *e.g.*, $p = 1$ for a set of search models that represents a fraction of the scattering matter of the crystal. The surprising result of this study is that optimal S/N may be produced with values of p much lower than an estimated $p_{\rm model}$ based on the fraction of scattering atoms. Several factors common to 'typical' molecular-replacement search problems could produce this effect, including inaccuracies of the search model set (*e.g.*, orientation errors and/or conformational differences), and neglect of the solvent contribution to the observed intensities.

A clear benefit of full SVS is the complete removal of the observed Patterson origin. The importance of downweighting the contribution of the origin in translation searches has long been recognized. The TO/O translation search of Harada *et al.* (1981) uses an origin overlap function to downweight incorrect translation positions with molecular overlap; Blundell & Tickle (1985) suggest an inverse weighting of translation functions with the integrated calculated Patterson origin peak. Origin subtraction is also implicit in the standard linear correlation coefficient between calculated and observed amplitudes, a function successfully used for translation searches by Fuginaga & Read (1987). The latter approach is also implemented in *X-PLOR* (Brünger, 1992). An advantage of the T_1 function over correlation coefficient calculation is significantly faster

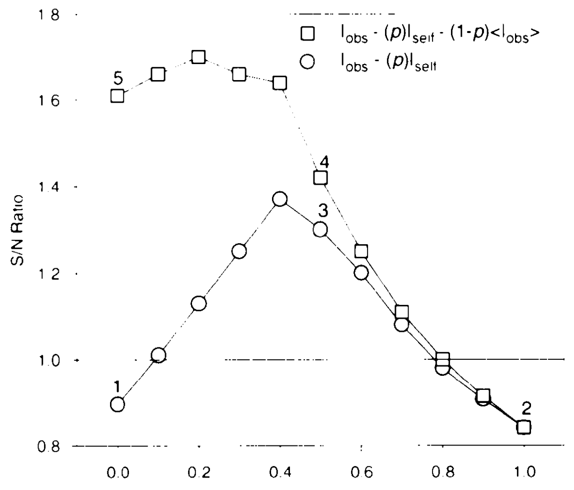

Fig. 2. Effect of systematic variation of p on S/N. Calculations were performed for search pair (M2–M2sym) as described in Fig. 1. The resulting S/N ratio is plotted for T_1 searches with p values between 0.0 and 1.0 with increments of 0.1. Points plotted with $p = 0$, 0.5 and 1.0 are labelled with reference numbers indicating the corresponding panel of Fig. 1.

calculation (on a Silicon Graphics 340VGX, the T_1 searches reported here took approximately one minute each; on the same machine, one correlation coefficient *X-PLOR* search on a different problem of similar scale took several hours).

The molecular-replacement program package *AmoRe* (Navaza, 1994) implements origin-removal in place of SVS for its T_1 function. The modification may be responsible for the dramatic success of the program in finding translation vectors for multiple partial search models. Origin removal is certainly a more flexible option than SVS, because no assumption about the content of the crystal unit cell is required for the calculation. SVS should still be useful when applied properly, since removal of self-vectors will reduce noise, but for difficult problems the optimal value of p will need to be established by trial-and-error. Given the speed at which the T_1 function is calculated, there should be little difficulty employing searches with several test values for p.

The basic theory behind the present study was derived in the period of the first *trp* repressor/operator structure determination (Otwinowski *et al.*, 1988), in collaboration with Paul Sigler and Zbyszek Otwinowski. I thank Robert M. Sweet and Venki Ramakrishnan for careful reading of the manuscript. This research was conducted at Brookhaven National Laboratory with the support of the US Department of Energy Office of Health and Environmental Research. Coordinates and observed reflection data of the tandem *trp* repressor/operator complex have been deposited in the Protein Data Bank.* The modified *MERLOT* T_1 subroutine may be requested from the author by e-mail, addressed to 'lawson@bnl.gov'.

References

Beurskens, P. T., Gould, R. O., Bruins Slot, H. J., & Bosman, W. P. (1987). *Z. Kristallogr.* **179**, 127–159.

Blundell, T. L. & Tickle, I. J. (1985). *Molecular Replacement: Proceedings of the Daresbury Study Weekend* 15–16 *February* 1985, edited by P. E. Machin, pp. 89–91. Daresbury, England: SERC.

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

Carey, J., Lewis, D. E. A., Lavoie, T. A & Yang, J. (1991). *J. Biol. Chem.* **266**, 24509–24513.

Crowther, R. A. & Blow, D. M. (1967). *Acta Cryst.* **23**, 544–548.

Fitzgerald, P. M. D. (1988). *J. Appl. Cryst.* **21**, 273–278.

Fuginaga, M. & Read, R. J. (1987). *J. Appl. Cryst.* **20**, 517–521.

Harada, Y., Lifchitz, A. & Berthou, J. (1981). *Acta Cryst.* A**37**, 398–406.

Haran, T. E., Joachimiak, A. & Sigler, P. B. (1992). *EMBO J.* **11**, 3021–3030.

Karle, J. (1972). *Acta Cryst.* B**28**, 820–824.

Klimasauskas, S., Kumar, S., Roberts, R. J. & Cheng, X. (1994). *Cell*, **76**, 357–369.

Langs, D. A. (1975). *Acta Cryst.* A**31**, 543–550.

Lattman, E. E. (1985). *Methods Enzymol.* **115**, 55–77.

Lawson, C. L. & Carey, J. (1993). *Nature (London)*, **366**, 178–182.

Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163

Otwinowski, Z., Schevitz, R. W., Zhang, R.-G., Lawson, C. L., Joachimiak, A., Marmorstein, R. Q., Luisi, B. F. & Sigler, P. B. (1988). *Nature (London)*, **335**, 321–329.

Rossmann, M. G. (1972). Editor. *The Molecular Replacement Method* New York: Gordon & Breach.

Schierbeek, A. J., Renetseder, R., Dijkstra, B. W. & Hol, W. G. J. (1985). *Molecular Replacement: Proceedings of the Daresbury Study Weekend* 15–16 *February*, 1985, edited by P. E. Machin, pp. 16–21. Daresbury, England: SERC.

---

* Atomic coordinates, observed reflection data and structure factors have been deposited with the Protein Data Bank, Brookhaven National Laboratory (Reference: 1TRR). Free copies may be obtained through The Managing Editor, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England (Reference: GR0415).