

Acta Cryst. (1996). **A52**, 647–648

Optimizing bond lengths in a model using Patterson methods prior to the solution of the structure

C. C. WILSON at ISIS Facility, CLRC Rutherford Appleton Laboratory, Chilton, Didcot, Oxon OX11 0QX, England. E-mail: ccw@isise.rl.ac.uk

(Received 4 August 1995; accepted 25 March 1996)

Abstract

The potential use of Patterson-methods calculations in the limited refinement of parameters in a molecular fragment is discussed. It is found to be possible to optimize the bond lengths in a partial fragment model in an incomplete or unknown structure. This allows, for example, tautomer discrimination and provision of improved partial model information for further calculations.

1. Introduction

The nature of the Patterson function allows the orientation and position of a fragment in a crystal structure to be obtained separately. With the internal description of a molecular fragment thus independent of its location in the unit cell, this opens up the possibility of performing restricted refinements of aspects of the molecular structure without having achieved full structural solution. The use of Patterson methods has recently been exploited in several areas related to fragment building (Wilson, 1988, 1990) but the present work is concerned with optimizing the existing fragment rather than adding to it.

In the present case, the rotation function of Tollin & Cochran (1964),

$$\sigma(\theta_1, \theta_2, \theta_3) = \sum_{\mathbf{h}} |F_{\mathbf{h}}^*|^2 \left\{ \left[\sum_i \cos(2\pi \mathbf{h} \cdot \mathbf{r}_i) \right]^2 + \left[\sum_i \sin(2\pi \mathbf{h} \cdot \mathbf{r}_i) \right]^2 \right\}, \quad (1)$$

is maximized with respect to not only the orientation of the molecular fragment within the unit cell (defined by the Eulerian angles θ_1 , θ_2 and θ_3 ; the usual procedure) but also some subset of the coordinates of the atoms in the model group [the \mathbf{r}_i in (1)], by refining one or more bond lengths in the model. $|F_{\mathbf{h}}^*|^2$ are the sharpened squared structure factors.

The object of such limited refinements is to allow improved values to be obtained for certain critical parameters in an incomplete or unknown structure, for example to allow the distinction to be made between single and double bonds.

2. Results and discussion

Three examples were chosen to illustrate the potential of this procedure in distinguishing bond lengths in a model structure. In each of these cases, there are several exocyclic ring substituents of various types and the bond-refinement procedure has been used to distinguish between these. The calculations were performed using around 300 $|F_{\mathbf{h}}^*|^2$ in each case. The effects of data resolution were examined using different structure-factor sets and while there was evidence of slightly improved resolution in the bond-length optimization by restricting the calculation to the higher-resolution data, the

basic results are also reproduced with smaller less well resolved data sets.

2.1. 3-Deazauracil

The structure of 3-deazauracil (Low & Wilson, 1983) consists of a six-membered ring with two non-hydrogenous substituents, in this case two O atoms. One of the main interests in the structure of this material is the tautomeric form adopted, reflected in the bond order of the two exocyclic C—O bonds, C2—O2 and C4—O4. The Patterson bond-length refinement was used to differentiate between these two bond lengths in the model.

The complete molecule with idealized geometry (bond lengths 1.4 Å, bond angles 120°) was used to find the best orientation of the molecule, with the exocyclic C—O bonds subsequently allowed to vary in the range 1.0–1.5 Å for the purpose of optimizing the model (Fig. 1). There are clearly two distinct exocyclic bond lengths identified, with C2—O2 at 1.28 Å and C4—O4 at 1.38 Å (Table 1). This would indicate that there is more double-bond character in C2—O2 than in C4—O4, which is in fact the case (Low & Wilson, 1983; Wilson, 1994).

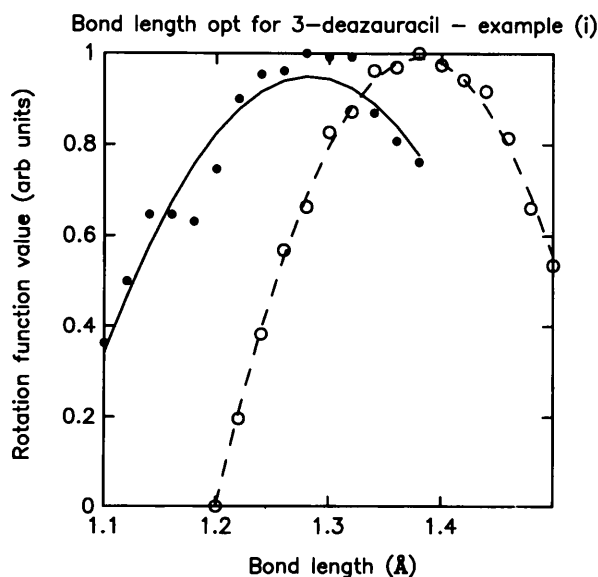


Fig. 1. Results of the Patterson bond-length refinement for 3-deazauracil [example (1)]. The values obtained for the rotation function as the C2—O2 (●) and C4—O4 (○) bond lengths in the model are varied are shown. The lines are simple quadratic fits to the data for C2—O2 (solid line) and C4—O4 (dashed line), respectively. Plots of a similar form were obtained for examples (2) and (3), with the results from all three summarized in Table 1.

Table 1. *Optimized bond lengths from the Patterson refinement procedure and comparison with values from the full refinements*

E.s.d.'s in parentheses are estimated from curve fits.

| Compound, cell and data details | Bond | Optimized length (Å) | Actual length (Å) | Significance |
|--|--------|----------------------|-------------------|-------------------------|
| 3-Deazauracil, C ₅ H ₅ NO ₂ <i>P</i> 2 ₁ 2 ₁ , <i>a</i> = 8.638, <i>b</i> = 5.279, <i>c</i> = 11.220 Å Data to <i>d</i> _{min} = 1.0 Å used | C2—O2 | 1.28 (2) | 1.26/1.27 | Double bond |
| | C4—O4 | 1.38 (1) | 1.32/1.34 | Single bond |
| | | | | |
| 3-Methylxanthine, C ₆ H ₆ N ₄ O ₂ <i>P</i> 2 ₁ / <i>c</i> , <i>a</i> = 3.742, <i>b</i> = 14.710, <i>c</i> = 12.159 Å, β = 96.8° Data to <i>d</i> _{min} = 1.1 Å used | C2—O2 | 1.26 (3) | 1.22 | Double bond |
| | N3—C3 | 1.41 (2) | 1.47 | Single bond |
| | C6—O6 | 1.33 (2) | 1.23 | Double bond? |
| 2',3',5'-Tri- <i>O</i> -acetylguanosine, C ₁₆ H ₁₉ N ₅ O ₈ <i>P</i> 2 ₁ , <i>a</i> = 7.414, <i>b</i> = 11.491, <i>c</i> = 11.790 Å, β = 99.47° Data to <i>d</i> _{min} = 1.2 Å used | C2—N2 | 1.38 (2) | 1.37 | Single bond |
| | C6—O6 | 1.26 (1) | 1.24 | Double bond |
| | N9—C1' | 1.50 (1) | 1.47 | Single bond (aliphatic) |

2.2. 3-Methylxanthine

3-Methylxanthine (Low, Tollin, Brand & Wilson, 1986) consists of a nine-membered fused purine ring with three non-hydrogenous substituents, two O atoms (O2 and O6) and a methyl group (carbon C6). The Patterson bond-refinement technique was used in this case to confirm the stereochemistry of the molecule by determining the three exocyclic bond lengths in the model. The rotation-function calculations were carried out on the complete molecule, with the results of the bond-length optimization shown in Table 1. It is clear that these identify the distinct exocyclic bond lengths in the material, with C2—O2 at 1.26 Å, N3—C3 at 1.41 Å and C6—O6 at 1.33 Å. The latter two values are respectively slightly lower and higher than expected, perhaps reflecting a *trans* symmetry-averaging effect in the Patterson-function calculations. Nonetheless, the chemical nature of the structure is elucidated unambiguously from these results.

2.3. 2',3',5'-Tri-*O*-acetylguanosine

2',3',5'-Tri-*O*-acetylguanosine (Wilson, Low & Tollin, 1985) is a larger structure, a substituted nucleoside containing

a planar base ring [a nine-membered fused purine ring with two non-hydrogenous substituents – an O atom (O6) and an NH₂ group (nitrogen N2)] and a substituted ribose sugar, linked to the purine ring by the glycosidic N9—C1' bond. The 12-atom planar molecular fragment was used as a starting model. Patterson bond refinement was used in this case to identify these three exocyclic bond lengths in the model, in particular to identify the location of the substantial non-planar sugar substituent attached to the base ring. Once again, the distinct exocyclic bond lengths in the material are indicated (Table 1) and combining simple chemical knowledge with information gleaned from the Patterson-function calculations allows the location of the sugar group attached to the base ring to be found, facilitating further development of the model structure. It should also be noted that in this case the partial model used in the calculations is a minor part of the overall molecule.

3. Conclusions

The results from these first tests of the Patterson refinement procedure are promising, with refinement of bond lengths in a structural model shown to be possible using Patterson-method calculations. This refinement can be carried out on incomplete or unknown structures, lending the method potential utility in cases where full structural solution is difficult, for example when the available data are limited. The bond discrimination available can help define molecular geometry in an incomplete or unknown structure or to aid further structure-solution attempts using an improved or extended model.

References

- Low, J. N., Tollin, P., Brand, E. & Wilson, C. C. (1986). *Acta Cryst.* **C42**, 1447–1448.
 Low, J. N. & Wilson, C. C. (1983). *Acta Cryst.* **C39**, 1688–1690.
 Tollin, P. & Cochran, W. (1964). *Acta Cryst.* **17**, 1322–1324.
 Wilson, C. C. (1988). *Acta Cryst.* **A44**, 478–481.
 Wilson, C. C. (1990). *J. Appl. Cryst.* **23**, 132–134.
 Wilson, C. C. (1994). *J. Chem. Crystallogr.* **24**, 371–373.
 Wilson, C. C., Low, J. N. & Tollin, P. (1985). *Acta Cryst.* **C41**, 1123–1125.