Acta Cryst. (1984). A**40**, 364–366

A Comparison of Two Methods for the Calculation of Pseudorotation Parameters

BY MARIUSZ JASKÓLSKI

Department of Crystallography, Faculty of Chemistry, A. Mickiewicz University, ul. Grunwaldzka 6, 60-780 *Poznań, Poland*

(*Received* 16 *November* 1983; *accepted* 18 *January* 1984)

Abstract

A comparison of the two methods available for the calculation of exact values of pseudorotation parameters P, τ_m, *i.e.* the Fourier-series (FS) method [Rao, Westhof & Sundaralingam (1981). *Acta Cryst.* A**37**, 421–425] and the least-squares (LS) method [Jaskólski (1983). *J. Chem. Educ.* **60**, 980–981] reveals that they yield identical values for τ_m and P. The variances $\sigma^2(\tau_m)$ and $\sigma^2(P)$ calculated by these methods are, however, different since they have different interpretations.

The concept of pseudorotation introduced by Kilpatrick, Pitzer & Spitzer (1947) and applied to five-membered rings (Geise, Altona & Romers, 1967; Altona, Geise & Romers, 1968) has been widely used for the description of the conformation of five-membered rings and became particularly popular for the description of the ribose pucker in nucleosides and nucleotides when Altona & Sundaralingam (1972) extended it to furanose systems. For several years, however, its practical applications suffered from some inexactitude connected with the possibility of calculating several slightly different τ_m values for the same ring (Rao, Westhof & Sundaralingam, 1981; hereafter RWS). To overcome this shortcoming, RWS proposed an ingenious procedure (hereafter FS) in which the five endocyclic torsion angles ($\theta_1, \ldots, \theta_5$) are represented by a second-order Fourier series. In an independent attempt to calculate exact values of τ_m and P the present author used the (nonlinear) method of least squares (LS) to fit τ_m and P to the five observations θ_i (Jaskólski, 1983). The final formulae for calculating the improvements $\Delta\tau_m$ and ΔP in the LS procedure with unit weights (*i.e.* in the most common case when the observed endocyclic torsion angles θ_i have similar e.s.d.'s) are the following:

$$\Delta\tau_m = (\tfrac{2}{5}) \sum_{i=1}^{5} \theta_i \cos(P + 2\alpha_i) - \tau_m \qquad (1)$$

and

$$\Delta P = -(\tfrac{2}{5}) \frac{1}{\tau_m} \sum_{i=1}^{5} \theta_i \sin(P + 2\alpha_i), \qquad (2)$$

where

$$\alpha_i = \tfrac{2}{5}\pi(i-1). \qquad (3)$$

The two methods (FS and LS) can be shown to yield identical final values of τ_m and P. The parameters from the LS approach are considered final when the LS process has converged, *i.e.* when

$$\Delta\tau_m = 0 \qquad (4)$$

and

$$\Delta P = 0. \qquad (5)$$

From (1) and (4) we get

$$\tau_m = (\tfrac{2}{5}) \sum_{i=1}^{5} \theta_i \cos(P + 2\alpha_i) \qquad (6)$$

and from (2) and (5)

$$\sum_{i=1}^{5} \theta_i \sin(P + 2\alpha_i) = 0. \qquad (7)$$

We note that (6) and (7) can be used *directly* to calculate τ_m and P in the LS method without the necessity of running several least-squares cycles. Equation (7) can be readily transformed to give

$$\tan P = -\frac{\sum_{i=1}^{5} \theta_i \sin 2\alpha_i}{\sum_{i=1}^{5} \theta_i \cos 2\alpha_i}, \qquad (8)$$

which is identical with (11) of RWS expressing P in the FS method.

To show the identity of the τ_m values obtained from both methods, we first multiply (8) and (9) of RWS by $\cos P$ and $\sin P$, respectively, and add them together to obtain

$$\tau_m = (\tfrac{2}{5})\left(\cos P \sum_{i=1}^{5} \theta_i \cos 2\alpha_i - \sin P \sum_{i=1}^{5} \theta_i \sin 2\alpha_i \right), \qquad (9)$$

which readily transforms to (6).

The equivalence of the two methods is illustrated in Table 1, which presents the pseudorotation

0108-7673/84/040364-03$01.50

Table 1. *Comparison of τ_m, P and their standard deviations calculated by the FS and LS methods*

			FS method				LS method		
Entry	Reference*	$\sigma(\theta_i)$† (°)	A_0‡	A_1‡	$P(\sigma P)$§ (°)	$\tau_m(\sigma\tau_m)$§ (°)	$P(\sigma P)$ (°)	$\tau_m(\sigma\tau_m)$ (°)	R¶(%)
i	(*b*)	0·4	0·16	0·67	159·0 (4)	39·7 (2)	159·0 (6)	39·7 (4)	1·62
ii	(*d*)	0·9–1·0	0·21	1·02	163·3 (9)	38·3 (6)	163·3 (9)	38·3 (6)	2·90
iii	(*f*)(1)	5	0·33	1·48	147 (4)	43 (3)	147 (1)	43·1 (9)	3·52
iv	(*h*)(1)	0·6	−0·01	0·08	27 (5)	3·9 (4)	26·7 (7)	3·90 (5)	2·06
v	(*h*)(3)	0·7	0·01	0·06	75 (1)	25·4 (5)	74·95 (8)	25·43 (4)	0·24
vi	(*i*)	0·3	0·10	0·92	132·8 (2)	42·1 (2)	132·8 (7)	42·1 (5)	2·13
vii**		0·3	2·10	4·84	138·1 (2)	42·7 (2)	138 (4)	43 (3)	12·3
viii††		0·03	0·01	0·09	132·8 (2)	4·21 (2)	132·8 (7)	4·21 (5)	2·13
ix‡‡		0·3	0·00	0·41	131 (2)	4·1 (2)	131 (3)	4·1 (2)	10·2

* Reference in Table 1 of the paper by RWS: (*b*) Emerson & Sundaralingham (1980); (*d*) Swaminathan & Sundaralingam (1980); (*f*) Hogle, Sundaralingam & Lin (1980), (1) first line; (*h*) Sprang, Rohrer & Sundaralingam (1978), (1) first line, (3) third line; (*i*) Smith, Chwang & Sundaralingam (1980).

† Average $\sigma(\theta_i)$ redetermined by us from the original positional parameters and their e.s.d.'s using the method of Shmueli (1974).

‡ See equation (1) of RWS.

§ Recalculated by us from the redetermined θ_i's and $\sigma(\theta_i)$'s (see footnote †). The standard deviations have been recalculated using equations (13) and (14).

¶ Conventional R factor, $R = 100\,\Sigma|\theta_{obs} - \theta_{calc}|/\Sigma|\theta_{obs}|$.

** θ_i and $\sigma(\theta_i)$ same as in (vi) except for θ_i which has been artificially biase by 10°: θ_1 (vii) = θ_1 (vi) + 10°.

†† θ_i and $\sigma(\theta_i)$ taken from (vi) and divided by 10: θ_i (viii) = θ_i (vi)/10, $\sigma[\theta_i$ (viii)] = $\sigma[\theta_i$ (vi)]/10; the ribofuranose ring (viii) is thus 'ten times flatter', and its torsion angles are ten times more precise than in the case of (vi).

‡‡ The values of θ_i (ix) have been generated by dividing the corresponding θ_i (vi)'s by a factor of ten and introducing random errors in the decimal place corresponding to $\sigma[\theta_i$ (vi)]: θ_i (ix) = θ_i (vi)/10 + $r \times \sigma[\theta_i$ (vi)], where r is a random number from (0,1); $\sigma[\theta_i$ (ix)] = $\sigma[\theta_i$ (vi)].

parameters calculated by the FS and LS methods for several sample structures used by RWS and for three extra artificial ribofuranose rings (entries vii, viii, ix).

Another advantage of the two methods is the possibility of estimating the variances in the derived pseudorotation parameters. However, the variances have quite different interpretations in these two approaches. In the LS method the variances are obtained by the inversion of the least-squares matrix and for the simplified practical case when $\sigma(\theta_i)$ = constant (unit weights in the least-squares equations), they are expressed (in rads) by

$$\sigma^2_{LS}(\tau_m) = (\tfrac{2}{5})k^2 \qquad (10)$$

$$\sigma^2_{LS}(P) = (\tfrac{2}{5})\frac{1}{\tau_m^2}k^2 \qquad (11)$$

(Jaskólski, 1983). The factor k corresponds to the well known goodness of fit and is calculated by

$$k^2 = \sum_{i=1}^{5}(\theta_{obs} - \theta_{calc})_i^2/3. \qquad (12)$$

Thus, the variances from the LS method provide a measure of how well the observed quantities (θ_i) fit the pseudorotation model of a puckered five-membered ring.

In the FS formalism the variances $\sigma^2_{FS}(\tau_m)$ and $\sigma^2_{FS}(P)$ are obtained from the variances in the observed torsion angles $\sigma^2(\theta_i)$ by the application of the error-propagation law. They provide therefore a measure of how reliable the derived values (τ_m, P) are. It should be pointed out here that the formulae for $\sigma^2_{FS}(\tau_m)$ and $\sigma^2_{FS}(P)$ given by RWS [(15) and (16) of RWS] are incorrect, since they treat A and B [see (8) and (9) of RWS] as independent variables. Below are given the correct expressions for $\sigma^2_{FS}(\tau_m)$ and $\sigma^2_{FS}(P)$ in the FS formalism with no correlation between individual θ_i's

$$\sigma^2_{FS}(\tau_m) = \left(\frac{1}{\tau_m}\right)^2 (A^2\sigma^2 A + B^2\sigma^2 B - C) \qquad (13)$$

$$\sigma^2_{FS}(P) = \left(\frac{1}{\tau_m}\right)^4 (A^2\sigma^2 B + B^2\sigma^2 A + C), \qquad (14)$$

where

$$C = 2(\tfrac{2}{5})^2 AB \sum_{i=1}^{5} \sigma^2(\theta_i)\cos 2\alpha_i \sin 2\alpha_i. \qquad (15)$$

For practical applications with $\sigma(\theta_i)$ = constant = $\sigma(\theta)$, the correction term C is 0 and the formulae (15) and (16) proposed by RWS become valid. In such a case, however, they can be further simplified to give

$$\sigma^2_{FS}(\tau_m) = (\tfrac{2}{5})\sigma^2(\theta) \qquad (16)$$

$$\sigma^2_{FS}(P) = (\tfrac{2}{5})\left(\frac{1}{\tau_m}\right)^2 \sigma^2(\theta) = \left(\frac{1}{\tau_m}\right)^2 \sigma^2_{FS}(\tau_m). \qquad (17)$$

Sample calculations revealed that for practical purposes the correction term C can be neglected unless the individual $\sigma(\theta_i)$'s show significant scatter.

We note here the similarity between (16), (17) and (10), (11). Their forms are identical but the true discrepancies between the observed and calculated quantities in (10) and (11) are in (16) and (17) replaced by the estimated variances $\sigma^2(\theta)$ of the observations. It seems reasonable to calculate σ^2_{FS}'s and σ^2_{LS}'s for each determination of the pseudorotation parameters and to use the bigger ones as the variances in τ_m and P.

Table 1 presents a compilation of the standard deviations calculated using the two methods. We decided to redetermine the values reported by RWS since we found some inconsistencies in their original paper [for instance, the ratio $\sigma(\tau_m)/\sigma(P)$ for entry v was $1{\cdot}8\ \tau_m$ (rad) while it should be close to $1{\cdot}0\ \tau_m$ (rad)]. Since individual $\sigma(\theta_i)$'s are not available in the original references, we had to calculate them also. For the calculations we used the original positional parameters and their e.s.d.'s and the method of Shmueli (1974). From a comparison of the standard deviations calculated with the two methods, three cases can be distinguished (Table 1): (i) σ_{FS}'s and σ_{LS}'s are roughly the same (entry ii); (ii) low-precision observations fit the model very well (entry v); and (iii) precise observations give very poor fit (entry vii).

As discussed above, the σ^2_{LS}'s are a measure of the fit between the real and calculated worlds. Another measure of this agreement in the LS formalism is the conventional R factor. Entries vii and ix of Table 1 show that R is sensitive to both systematic and random errors. As pointed out by RWS, the departure of the A_0 and A_1 coefficients of their Fourier series [see (1) of RWS] from 0 furnishes the FS method with a measure of the deviation from the ideal pseudorotation description. Table 1 reports the A_0 and A_1 values for each entry. Although it is possible to trace a very poor fit by the large values of $|A_0|$ and $|A_1|$ (entry vii), it is in general not obvious how to establish the discrepancy between the observations and the model from individual $|A_0|$ and $|A_1|$ values, and how to compare the deviations for different systems. It seems that the R factor provides a more convenient measure of the deviation from the ideal pseudorotation description.

References

ALTONA, C., GEISE, H. J. & ROMERS, C. (1968). *Tetrahedron*, **24**, 13–32.

ALTONA, C. & SUNDARALINGAM, M. (1972). *J. Am. Chem. Soc.* **94**, 8205–8212.

EMERSON, J. & SUNDARALINGAM, M. (1980). *Acta Cryst.* B**36**, 537–543.

GEISE, H. J., ALTONA, C. & ROMERS, C. (1967). *Tetrahedron Lett.* pp. 1383–1386.

HOGLE, J., SUNDARALINGAM, M. & LIN, G. H. (1980). *Acta Cryst.* B**36**, 564–570.

JASKÓLSKI, M. (1983). *J. Chem. Educ.* **60**, 980–981.

KILPATRICK, J. E., PITZER, K. S. & SPITZER, R. (1947). *J. Am. Chem. Soc.* **64**, 2483–2488.

RAO, S. T., WESTHOF, E. & SUNDARALINGAM, M. (1981). *Acta Cryst.* A**37**, 421–425.

SHMUELI, U. (1974). *Acta Cryst.* A**30**, 848–849.

SMITH, J., CHWANG, A. & SUNDARALINGAM, M. (1980). *Acta Cryst.* B**36**, 833–837.

SPRANG, S., ROHRER, D. C. & SUNDARALINGAM, M. (1978). *Acta Cryst.* B**34**, 2803–2810.

SWAMINATHAN, P. & SUNDARALINGAM, M. (1980). *Acta Cryst.* B**36**, 2590–2597.

Acta Cryst. (1984). A**40**, 366–373

Rules for Estimating the Values of Triplet Phase Invariants in Multiwavelength Anomalous Dispersion Experiments

BY J. KARLE

Laboratory for the Structure of Matter, Naval Research Laboratory, Washington, DC 20375, *USA*

(*Received* 12 *October* 1983; *accepted* 23 *January* 1984)

Abstract

Several simple rules, $R_{ano,4}$, $R_{ano,5}$, $R_{ano,6}$ and $R_{ano,7}$, have been derived on the basis of the mathematical and physical characteristics of anomalous dispersion experiments that permit the estimation of values for triplet phase invariants. They apply to two-wavelength experiments and concern a variety of values defined in terms of the real and imaginary corrections to atomic scattering factors. The rules apply to the case of a single type of predominant anomalous scatterer. The generalization to more than one type of predominant anomalous scatterer is also described. Test examples show that large numbers of invariants may be evaluated by these means with reliabilities that, in certain circumstances, are at a potentially useful level, but the ultimate applicability depends, of course, on the reliability of the experimental data. The only information required besides the measurements of the diffraction intensities is the chemical composition of the anomalously scattering atoms. If there is more than one type of predominant anomalous scatterer, information concerning the relative proportion of the different types is also required.

Introduction

In a previous article (Karle, 1984*b*), rules were presented for selecting triplet phase invariants whose values are close to certain anticipated values. The rules arise from considerations of a mathematical and physical

0108-7673/84/040366-08$01.50