Pramana – J. Phys., Vol. 36, No. 6, June 1991, pp. 581–588. © Printed in India.

Theoretical evaluation of the overall values of Booth type R-indices based on intensity variables

S PARTHASARATHY, N ELANGO[+] and K SEKAR

Department of Crystallography and Biophysics, University of Madras, Madras 600025, India
[+] Present address: Department of Chemistry, San Diego Super Computer Center, California, USA

MS received 11 January 1990; revised 27 February 1991

Abstract. Theoretical expressions for the overall values of three Booth type R-indices based on intensity variables are derived. The results are applicable to crystals of any space group containing any number and type of atoms at general positions in the asymmetric unit. The theoretical results were tested in the case of models of a few crystal structures.

Keywords. Booth type R-indices; intensity variables.

PACS No. 61·10

1. Introduction

In crystals containing dissimilar atoms the probability distribution of X-ray intensities is strongly influenced by the symmetry elements of the space group besides the centre of symmetry (Hauptman and Karle 1952; Karle and Hauptman 1953). For such crystals it is therefore necessary to derive the theoretical expressions for R-indices by taking into account the effect of space group symmetry. Though a number of R-indices have been proposed in the literature (Srinivasan and Parthasarathy 1976) only the Booth type indices based on intensity are easy to handle from a theoretical point of view, particularly when complications due to space group symmetry enter into picture. Furthermore, a comparative study of the behaviour of different types of R-indices has shown that, of all the R-indices, the Booth type index based on intensity is the best during the structure completion stage (Parthasarathy and Parthasarathi 1975). Hence Parthasarathy (1975) derived the theoretical expressions for the *local values* for Booth type indices for crystals containing any number and type of atoms at general positions in the unit cell. His results are applicable to only crystals of the triclinic, monoclinic and orthorhombic systems. It would be useful to derive the theoretical results which could be used to calculate the *overall values* of these R-indices for crystals of any space group. This can now be done owing to the availability of the higher order moments of intensity for crystals of all space groups (Shmueli and Kaldor 1981; Shmueli and Wilson 1981). In this paper we shall hence derive the theoretical expressions for the overall values of the three Booth type indices based on intensity variables [i.e., ${}_{B}\bar{R}_{1}(z)$, ${}_{B}\bar{R}_{1}(I)$ and ${}_{B}\bar{R}(I)$ – see (14)–(16) for a definition of these]. Since it is too difficult to consider the effect of coordinate errors in the theoretical treatment, particulary in crystals with dissimilar atoms, we shall derive results only for two

limiting situations, viz., a completely correct model (called related case or R-case) and a completely wrong model (called the unrelated case or UR-case). We shall also assume all the atoms in the asymmetric unit to be at general positions.

2. Notation and preliminary results

Consider a crystal belonging to any space group and containing n atoms in the asymmetric unit of which p are known and the remaining $q(=n-p)$ are yet to be located. We shall assume all the n atoms to occupy general positions. Let s be the number of equivalent general positions in the unit cell. Let $N(=ns)$ be the number of atoms in the unit cell of the crystal. Let $P(=ps)$ and $Q(=qs)$ respectively be the numbers of known and unknown atoms in the unit cell. The structure factor equation for a reflection $\mathrm{H}(=hkl)$ can be written as

$$F_N(\mathrm{H}) = \sum_{i=1}^{n} f_i J_i(\mathrm{H}), \tag{1}$$

where f_i is the atomic scattering factor of atom i and J_i is the trigonometric structure factor arising due to the contributions from the s equivalent atoms of which i is one (Shmueli and Kaldor 1981). We can also rewrite (1) in terms of the contributions from the P- and Q-atoms as

$$F_N(\mathrm{H}) = F_P(\mathrm{H}) + F_Q(\mathrm{H}). \tag{2}$$

Let $F_P^c(\mathrm{H})$ be the calculated structure factor of the model for reflection H. The intensity variables I_P^c, I_N, I_P and I_Q are defined as

$$I_P^c = |F_P^c(\mathrm{H})|^2 \text{ and } I_i = |F_i(\mathrm{H})|^2, \quad i = N, \quad P \text{ or } Q. \tag{3}$$

The normalized intensity variables z_P^c, z_N, z_P and z_Q are defined to be

$$z_P^c = I_P^c/\langle I_P^c\rangle \text{ and } z_i = I_i/\langle I_i\rangle, \quad i = N, \quad P \text{ or } Q. \tag{4}$$

It readily follows from (4) that

$$\langle z_P^c\rangle = 1 \text{ and } \langle z_i\rangle = 1, \quad i = N, \quad P \text{ or } Q. \tag{5}$$

Since all the atoms in the asymmetric unit occupy general positions, we have

$$\langle (z_P^c)^m\rangle = \langle z_P^m\rangle, \tag{6}$$

where m is a positive integer. Let us define σ_1^2 and σ_2^2 by*

$$\sigma_1^2 = \langle I_P^c\rangle/\langle I_N\rangle = \langle I_P\rangle/\langle I_N\rangle, \quad \sigma_2^2 = \langle I_Q\rangle/\langle I_N\rangle \tag{7}$$

so that

$$\sigma_1^2 + \sigma_2^2 = 1. \tag{8}$$

We shall denote the *local average* value of $|J|^k$ by m_k. That is,

$$m_k = \langle |J|^k\rangle \tag{9}$$

* Using the notation of (11) we can write (7) as $\sigma_1^2 = S_p(2)/S_n(2)$.

where the average is over a sufficiently large set of reflections in a narrow region of $\sin\theta/\lambda$. We shall define ρ to be the ratio

$$\rho = m_4/m_2^2. \tag{10}$$

It may be noted that ρ is a constant quantity for a given space group. It is convenient to define the quantities $S_p(k)$, $S_n(k)$, ψ_n and ψ_p by

$$S_i(k) = \sum_{j=1}^{i} f_{ij}^k, \quad i = n \text{ or } p, \tag{11}$$

$$\psi_i = S_i(4)/[S_i(2)]^2, \quad i = n \text{ or } p. \tag{12}$$

Following Shmueli and Wilson (1981) we can write the second moment of $z_i (i = n$ or $p)$ as

$$\begin{aligned} \langle z_N^2 \rangle &= L + (\rho - L)\psi_n, \\ \langle z_P^2 \rangle &= L + (\rho - L)\psi_p, \end{aligned} \tag{13}$$

where L is 3 for the centrosymmetric case (C-case, hereafter) and 2 for the non-centrosymmetric case (NC-case hereafter).

3. Theoretical derivation

The overall values of the three Booth type R-indices ${}_B\bar{R}_1(z)$, ${}_B\bar{R}_1(I)$ and ${}_B\bar{R}(I)$ are defined as

$${}_B\bar{R}_1(z) = \sum_{hkl}(z_N - z_P^c)^2 \Big/ \sum_{hkl} z_N^2, \tag{14}$$

$${}_B\bar{R}_1(I) = \sum_{hkl}(I_N - I_P^c/\sigma_1^2)^2 \Big/ \sum_{hkl} I_N^2, \tag{15}$$

$${}_B\bar{R}(I) = \sum_{hkl}(I_N - I_P^c)^2 \Big/ \sum_{hkl} I_N^2. \tag{16}$$

For the purpose of theoretical evaluation it is convenient to use (4) and rewrite (15) and (16) in terms of the normalized variables z_N and z_p^c as

$${}_B\bar{R}_1(I) = \sum_{hkl}\langle I_N \rangle^2 (z_N - z_P^c)^2 \Big/ \sum_{hkl}\langle I_N \rangle^2 z_N^2, \tag{17}$$

$${}_B\bar{R}(I) = \sum_{hkl}\langle I_N \rangle^2 (z_N - \sigma_1^2 z_P^c)^2 \Big/ \sum_{hkl}\langle I_N \rangle^2 z_N^2. \tag{18}$$

From (14), (17) and (18) it is clear that we can rewrite these equations in a common format as

$$\bar{R} = \sum_{hkl} w(z_N - \alpha z_P^c)^2 \Big/ \sum_{hkl} w z_N^2. \tag{19}$$

provided we define w as

$$w = \langle I_N \rangle^2 \text{ for } \bar{\mathrm{R}} = {}_{\mathrm{B}}\bar{\mathrm{R}}_1(I) \text{ and } {}_{\mathrm{B}}\bar{\mathrm{R}}(I)$$
$$= 1 \text{ for } \bar{\mathrm{R}} = {}_{\mathrm{B}}\bar{\mathrm{R}}_1(z) \tag{20}$$

and α as

$$\alpha = 1 \text{ for } \bar{\mathrm{R}} = {}_{\mathrm{B}}\bar{\mathrm{R}}_1(z) \text{ and } {}_{\mathrm{B}}\bar{\mathrm{R}}_1(I)$$
$$= \sigma_1^2 \text{ for } \bar{\mathrm{R}} = {}_{\mathrm{B}}\bar{\mathrm{R}}(I). \tag{21}$$

Partitioning the region $0 \leqslant \mathrm{H} \leqslant \mathrm{H}_{\max}$ of the reciprocal space into v thin shells and replacing the summation over *hkl* into summation over the index i carried out over the various reflections in a *given shell* j and then over the index j carried over the v shells we can rewrite (19) as

$$\bar{\mathrm{R}} = \sum_j \sum_i w_{ij}(z_N - \alpha z_P^c)_{ij}^2 \Big/ \sum_j \sum_i w_{ij}(z_N^2)_{ij}. \tag{22}$$

If the shells are sufficiently thin, we can take the values of w_{ij} for the different reflections in any given shell j to be the same (w_{ij}, say). We can therefore rewrite (22) as

$$\bar{\mathrm{R}} = \sum_j w_j \sum_i (z_N - \alpha z_P^c)_{ij}^2 \Big/ \sum_j w_j \sum_i (z_N^2)_{ij}.$$

$$\bar{\mathrm{R}} = \sum_j w_j n_j \langle (z_N - \alpha z_P^c)^2 \rangle_j \Big/ \sum_j w_j n_j \langle (z_N^2) \rangle_j, \tag{23}$$

where n_j is the number of reflections in shell j and $\langle (z_N - \alpha z_P^c)^2 \rangle_j$ and $\langle z_N^2 \rangle_j$ are the mean values of $(z_N - \alpha z_P^c)^2$ and z_N^2 in the jth shell. It is convenient to partition the region $0 \leqslant \mathrm{H} \leqslant \mathrm{H}_{\max}$ of the reciprocal space into equi-volume shells instead of shells of arbitrary volumes. If there are v equi-volume shells then $n_1, n_2, \ldots n_v$ in (23) are equal so that we can cancel n_j from the numerator and denominator of (23). We thus obtain from (23) that

$$\bar{\mathrm{R}} = \sum_{j=1}^{v} w_j \langle (z_N - \alpha z_P^c)^2 \rangle_j \Big/ \sum_{j=1}^{v} w_j \langle z_N^2 \rangle_j. \tag{24}$$

In crystals with different species of atoms, the moments of z_N and $(z_N - \alpha z_P^c)$ depend on $\sin\theta/\lambda$ and hence the numerator and denominator in (24) are to be evaluated separately. The theoretical expression for $\langle z_N^2 \rangle_j$ is available in (13). We shall now derive an expression for $\langle (z_N - \alpha z_P^c)^2 \rangle_j$.

Expanding the square term in $\langle (z_n - \alpha z_P^c)^2 \rangle_j$ and then using the known result that the expectation of a sum of random variables is the sum of their expectations, we obtain

$$\langle (z_N - \alpha z_P^c)^2 \rangle_j = \langle z_N^2 \rangle_j + \alpha_j^2 \langle (z_P^c)^2 \rangle_j - 2\alpha_j \langle z_N z_P^c \rangle_j, \tag{25}$$

where α_j is the value of α corresponding to shell j. The results obtained so far are valid for both R and UR cases. For further simplications we have to consider these two cases separately.

3.1 *Results for the related case*

From (2) we have

$$I_N = |F_P(\mathrm{H}) + F_Q(\mathrm{H})|^2 = I_P + I_Q + 2(I_P I_Q)^{1/2}\varepsilon, \tag{26}$$

where

$$\varepsilon = \cos\phi \quad \text{for NC case}$$
$$= s_P s_Q \quad \text{for C case} \tag{27}$$

Here ϕ is the angle between the structure factors F_P and F_Q in the NC case and s_P and s_Q are the signs of F_P and F_Q for the C case respectively. Since the P- and Q-groups are independent, it readily follows that

$$\langle\varepsilon\rangle = 0 \text{ for both the C and NC cases.} \tag{28}$$

Making use of the normalized variables z_P and z_Q we can rewrite (26) as

$$z_N = \sigma_1^2 z_P + \sigma_2^2 z_Q + 2\sigma_1\sigma_2(z_P z_Q)^{1/2}\varepsilon \tag{29}$$

so that

$$\langle z_N z_P^c\rangle_j = (\sigma_1^2)_j\langle z_P z_P^c\rangle_j + (\sigma_2^2)_j\langle z_Q z_P^c\rangle_j + 2(\sigma_1\sigma_2)\langle(z_P z_Q)^{1/2} z_P^c\varepsilon\rangle_j, \tag{30}$$

where $(\sigma_1^2)_j$ and $(\sigma_2^2)_j$ are the values of these quantities corresponding to shell j. For the related case $z_P = z_P^c$ so that (30) yields

$$\langle z_N z_P^c\rangle_j = (\sigma_1^2)_j\langle z_P^2\rangle_j + (\sigma_2^2)_j. \tag{31}$$

For obtaining (31) from (30) we have used (5), (28) and the fact that z_P, z_Q and ε are independent random variables. Making use of (31) in (25) and the results that $\langle z_P^2\rangle = \langle(z_P^c)^2\rangle$ (see (6)) we obtain

$$\langle(z_N - \alpha z_P^c)^2\rangle_j = \langle z_N^2\rangle_j + (\alpha_j^2 - 2\alpha_j^2(\sigma_1^2)_j)\langle z_P^2\rangle_j - 2\alpha_j(\sigma_2^2)_j. \tag{32}$$

Making use of (13) in (32) and simplifying we can show that

$$\langle(z_N - \alpha z_P^c)^2\rangle_j = [C_1 + (\rho - L)(\psi_n + C_2\psi_p)]_j \tag{33}$$

C_1 and C_2 occurring in (33) are defined in table 1. Making use of (33) and (13) in (24) we obtain

$$\bar{R} = \frac{\sum_j w_j[C_1 + (\rho - L)(\psi_n + C_2\psi_p]_j}{\sum_j w_j[L + (\rho - L)\psi_n]_j}. \tag{34}$$

From (11) and (20) we obtain w_j to be

$$w_j = [sS_n(2)]^2 \text{ for } \bar{R} = {}_B\bar{R}_1(I) \text{ and } {}_B\bar{R}(\bar{I})$$
$$= 1 \text{ for } {}_B\bar{R}_1(z). \tag{35}$$

Table 1. Definition of the quantities C_1, C_2 and w occurring in eq. (38).

	Related case			Unrelated case		
$\bar{R}$	C_1	C_2	w	C_1	C_2	w
${}_B\bar{R}(I)$	$L - L\sigma_1^4 - 2\sigma_1^2\sigma_2^2$	$-\sigma_1^4$	$[S_n(2)]^2$	$L + L\sigma_1^4 - 2\sigma_1^2$	σ_1^4	$[S_n(2)]^2$
${}_B\bar{R}_1(I)$	$2\sigma_2^2(L-1)$	$1 - 2\sigma_1^2$	$[S_n(2)]^2$	$2(L-1)$	1	$[S_n(2)]^2$
${}_B\bar{R}_1(z)$	$2\sigma_2^2(L-1)$	$1 - 2\sigma_1^2$	1	$2(L-1)$	1	1

3.2 *Results for the unrelated case*

For the unrelated case z_N and z_P^c are independent random variables so that

$$\langle z_N z_P^c \rangle_j = \langle z_N \rangle \langle z_P^c \rangle_j = 1. \tag{36}$$

Using (36) in (25) we obtain

$$\langle (z_N - \alpha z_P^c)^2 \rangle_j = \langle z_N^2 \rangle_j + \alpha_j^2 \langle (z_P^c)^2 \rangle_j - 2\alpha_j \tag{37}$$

Making use of (13) in (37) and simplifying we can show that (34) applies for the unrelated case as well, provided C_1 and C_2 are redefined for the present case as in table 1.

4. Evaluation of theoretical overall values of R-indices

The theoretical expressions for the overall values ${}_B\bar{R}_1(z)$, ${}_B\bar{R}_1(I)$ and ${}_B\bar{R}(I)$ for both R and UR cases take the common format

$$\bar{R} = \frac{\sum_j w_j [C_1 + (\rho - L)(\psi_n + C_2 \psi_p)]_j}{\sum_j w_j [L + (\rho - L)\psi_n]_j}. \tag{38}$$

The quantities σ_1^2, σ_2^2, ψ_n and ψ_p occurring in (38) can be computed from a knowledge of the unit cell content of the crystal and the model (see (11), (12) and the footnote on page 582). L is 3 for the C-case and 2 for the NC-case. The values of ρ are available in table 1 of Shmueli and Kaldor (1981) under the column entitled q/p^2. The theoretical overall values of the three Booth type R-indices for the related and unrelated models can hence be computed from (38).

5. Test of the theoretical results

The theoretical results obtained here were tested using the observed scaled intensities (I_0, say) of a few crystal structures belonging to different space group symmetry. The details of these structures are given in table 2. In each crystal the reflections for which $\sin\theta/\lambda \geqslant 1/a_{\min}$, where $a_{\min}$ is the least unit cell parameter, were used in the calculation due to theoretical reasons (Wilson 1949). In each case the overall temperature factor (B, say) obtained from MULTAN 80 (Main *et al* 1980) was used to calculate the intensities I_P^c of the trial structure. Cubic spline functions with smoothing (Carl de Boor 1978) were used to fit the data ($\sin\theta/\lambda$, $\langle I_0 \rangle$), ($\sin\theta/\lambda$, $\langle I_P^c \rangle$) and ($\sin\theta/\lambda$, σ_1^2) where σ_1^2 is taken to be $\langle I_P^c \rangle / \langle I_0 \rangle$. For each reflection, the values of $\langle I_0 \rangle$, $\langle I_p^c \rangle$ and σ_1^2 were obtained from the corresponding analytical expressions of the cubic spline functions using the value of $\sin\theta/\lambda$ appropriate to the reflection. The values of $z_N (= I_0/\langle I_0 \rangle)$, $z_P^c (= I_P^c/\langle I_P^c \rangle)$ and σ_1^2 needed for computing the experimental values of ${}_B\bar{R}_1(I)$ and ${}_B\bar{R}_1(z)$ were computed for each reflection. The experimental values of R-indices where then computed and these are given against the symbol "E" in table 2. The corresponding theoretical values for the related and unrelated cases are shown

Table 2. Experimental verification of the theoretical results.

No.	Asymmetric unit	Space group	B	*p*-group		${}_B\bar{R}(I)$	${}_B\bar{R}_1(I)$	${}_B\bar{R}_1(z)$
1.	$C_{13}O_2Br$	$P\bar{1}$	4·25	(i) Br	E	43	44	45
					R	38	37	29
					UR	69	76	74
				(ii) C_2O_2Br	E	31	32	37
					R	24	23	19
					UR	80	88	83
2.	$C_{20}NO$	$P\bar{1}$	4.24	(i) ONC_8	E	55	52	65
					R	58	65	67
					UR	90	128	129
				(ii) ONC_{16}	E	23	25	26
					R	21	22	23
					UR	112	131	131
3.	$C_{11}N_2S$	$P2_1$	3·19	(i) S	E	49	50	44
					R	52	45	43
					UR	66	73	73
				(ii) SN_2C_2	E	38	43	32
					R	33	31	30
					UR	70	85	83
4.	$C_{10}NOS$	$P2_1/n$	3.41	(i) S	E	64	59	67
					R	62	64	59
					UR	84	109	109
				(ii) $SONC_6$	E	37	36	34
					R	19	20	19
					UR	110	126	125
5.	C_9NOSCL	$P2_12_12_1$	3·53	(i) CLS	E	26	31	36
					R	30	30	27
					UR	75	94	94
				(ii) $CLSNC_2$	E	19	22	29
					R	22	22	20
					UR	79	95	95

Structures 1 to 5 are those determined by (1) Sivakumar and Natarajan (1990); (2) Sekar *et al* (1990); (3) Eswaramoorthy *et al* (1990c); (4) Eswaramoorthy *et al* (1990a); (5) Eswaramoorthy *et al* (1990b) respectively. Values of R-indices are given in per cent.

in table 2 against the symbols "R" and "UR" respectively. In the case of each crystal, results were obtained for two different *p*-groups and these are shown under the designations of (i) and (ii) in table 2. A study of table 2 shows that the experimental values are reasonably close to the corresponding theoretical values expected for the related models in all the cases.

Acknowledgements

One of authors (SP) thanks the CSIR (India) for financial support. The financial support of the UGC (India) and CSIR (India) to Dr N Elango and Mr K Sekar respectively is gratefully acknowledged.

References

Carl de Boor 1978 *A practical guide to splines* (New York: Springer-Verlag)

Eswaramoorthy S, Ponnuswamy M N, Raju K S and Edmund W Czerwinski 1990a *Acta Crystallogr.* (in press)

Eswaramoorthy S, Ponnuswamy M N, Raju K S and Edmund W Czerwinski 1990b (Personal communication)

Eswaramoorthy S, Ponnuswamy M N, Raju K S, Umarani R and Edmund W Czerwinski 1990c (Personal communication)

Hauptman H and Karle J 1952 *Acta Crystallogr.* **11** 724

Karle J and Hauptman H 1953 *Acta Crystallogr.* **6** 131

Main P, Fiske S J, Hull S E, Lessinger L, Germain G, Declercq J-P and Woolfson M M 1980 MULTAN80 *A system of computer programs for the automatic solution of crystal structures from x-diffraction data* (Universities of York, England and Louvain, Belgium)

Parthasarathy S 1975 *Acta Crystallogr.* **A31** 375

Parthasarathy S and Parthasarathi V 1975 *Acta Crystallogr.* **A31** 178

Sekar K, Parthasarathy S and Radhakrishnan T R 1990 (Personal Communication)

Shmueli U and Kaldor U 1981 *Acta Crystallogr.* **A37** 76

Shmueli U and Wilson A J C 1981 *Acta Crystallogr.* **A37** 342

Sivakumar K and Natarajan S 1990 (Personal Communication)

Srinivasan R and Parthasarathy S 1976 *Some statistical applications in X-ray crystallography* (New York: Pergamon Press)

Wilson A J C 1949 *Acta Crystallogr.* **2** 318