Table 18.5.8.1. $R = \langle|\Delta F|\rangle/\langle|F|\rangle$ *as a function of* $s\langle\Delta r\rangle$ *in the Luzzati model for three-dimensional noncentrosymmetric structures* $(s = 2\sin\theta/\lambda)$

$s\langle\Delta r\rangle$	R	$s\langle\Delta r\rangle$	R
0.00	0.000	0.10	0.237
0.01	0.025	0.12	0.281
0.02	0.050	0.14	0.319
0.03	0.074	0.16	0.353
0.04	0.098	0.18	0.385
0.05	0.122	0.20	0.414
0.06	0.145	0.25	0.474
0.07	0.168	0.30	0.518
0.08	0.191	0.35	0.548
0.09	0.214	∞	0.586

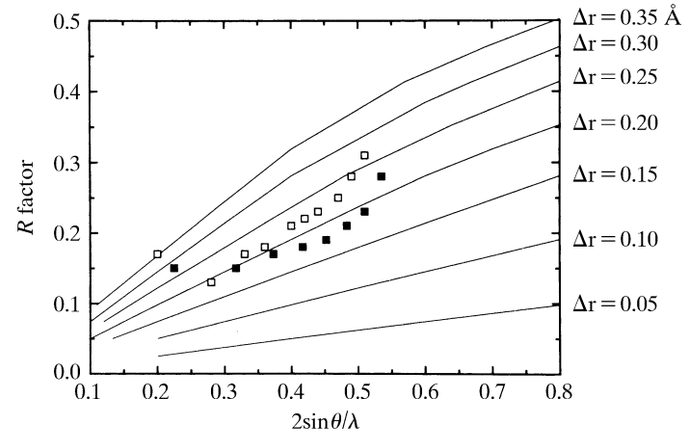

Fig. 18.5.8.1. Luzzati plots showing the refined R factor as a function of resolution for 1TGI (solid squares) and 1TGF (open squares) (Daopin *et al.*, 1994).

$F_{\rm calc}$ model (scattering factors, thermal parameters *etc.*) was perfect, apart from coordinate errors.

(2) The Gaussian probability distribution for these coordinate errors was assumed to be the *same for all atoms*, independent of Z or B.

(3) The atoms were not required to be identical, and the position errors were not required to be small.

Luzzati gave families of curves for R *versus* $2\sin\theta/\lambda$ for varying average positional errors $\langle\Delta r\rangle$ for both centrosymmetric and noncentrosymmetric structures. The curves do not depend on the number N of atoms in the cell. They all rise from $R = 0$ at $2\sin\theta/\lambda = 0$ to the Wilson (1950) values 0.828 and 0.586 for random structures at high $2\sin\theta/\lambda$. Table 18.5.8.1 gives $R = \langle|\Delta F|\rangle/\langle|F|\rangle$ as a function of $s\langle\Delta r\rangle$ for three-dimensional noncentrosymmetric structures.

In a footnote (p. 807), Luzzati suggested that at the end of a normal refinement (with R nonzero due to experimental and model errors, *etc.*), the curves would indicate an upper limit for $\langle\Delta r\rangle$. He noted that typical small-molecule $\sigma(r)$'s of 0.01–0.02 Å, if used as $\langle\Delta r\rangle$ in the plots, would give much smaller R's than are found at the end of a refinement.

As examples, the Luzzati plots for the two structures of TGF-β2 are shown in Fig. 18.5.8.1. Daopin *et al.* (1994) inferred average $\langle\Delta r\rangle$'s around 0.21 Å for 1TGI and 0.23 Å for 1TGF.

Of the three Luzzati assumptions summarized above, the most attractive is the third, which does not require the atoms to be identical nor the position errors to be small. For proteins, there are very obvious difficulties with assumption (2). Errors do depend very strongly on Z and B. In the high-angle data shells, atoms with large B's contribute neither to ΔF nor to $|F|$, and so have no effect on R in these shells. In their important paper on protein accuracy, Chambers & Stroud (1979) said 'the [Luzzati] estimate derived from reflections in this range applies mainly to [the] best determined atoms.'

Thus a Luzzati plot seems to allow a cautious upper-limit statement about the precision of the best parts of a structure, but it gives little indication for the poor parts.

One reason for the past popularity of Luzzati plots has been that the R values for the middle and outer shells of a structure often roughly follow a Luzzati curve. Evidently, the effective average $\langle\Delta r\rangle$ for the structure must be decreasing as $2\sin\theta/\lambda$ increases, since atoms of high B are ceasing to contribute, whereas the proportionate experimental errors must be increasing. This also suggests that the upper limit for $\langle\Delta r\rangle$ for the low-B atoms could be estimated from the lowest Luzzati theoretical curve touched by the experimental R plot. Thus in Fig. 18.5.8.1 the upper limits for the low-B atoms could be taken as 0.18 and 0.21 Å, rather than the 0.21 and 0.23 Å chosen by Daopin *et al.*

From the introduction of $R_{\rm free}$ by Brünger (1992) and the discussion of $R_{\rm free}$ by Tickle *et al.* (1998*b*), it can be seen that Luzzati plots should be based on a residual more akin to $R_{\rm free}$ than R in order to avoid bias from the fitting of data.

The mean positional error $\langle\Delta r\rangle$ of atoms can also be estimated from the σ_A plots of Read (1986, 1990). This method arose from Read's analysis of improved Fourier coefficients for maps using phases from partial structures with errors. It is preferable in several respects to the Luzzati method, but like the Luzzati method it assumes that the coordinate distribution is the same for all atoms. Luzzati and/or Read estimates of $\langle\Delta r\rangle$ are available for some of the structures in Tables 18.5.7.2 and 18.5.7.3. Often, the two estimates are not greatly different.

18.5.8.2. *Statistical reinterpretation of Luzzati plots*

Luzzati plots are fundamentally different from other statistical estimates of error. The Luzzati theory applies to an idealized incomplete refinement and estimates the average shifts needed to reach $R = 0$. In the least-squares method, the equations for shifts are quite different from the equations for estimating variances in a converged refinement. However, Luzzati-style plots of R *versus* $2\sin\theta/\lambda$ can be reinterpreted to give statistically based estimates of $\sigma(x)$.

During Cruickshank's (1960) derivation of the approximate equation (18.5.6.2) for $\sigma(x)$ in diagonal least squares, he reached an intermediate equation

$$\sigma^2(x) = N_i \Big/ \left[4\sum_{\rm obs}(s^2/R^2)\right]. \qquad (18.5.8.1)$$

He then assumed R to be independent of s $(= 2\sin\theta/\lambda)$ and took R outside the summation to reach (18.5.6.2) above.

Luzzati (1952) calculated the acentric residual R as a function of $\langle\Delta r\rangle$, the average radial error of the atomic positions. His analysis shows that R is a linear function of s and $\langle\Delta r\rangle$ for a substantial range of $s\langle\Delta r\rangle$, with

$$R(s,\ \langle\Delta r\rangle) = (2\pi)^{1/2} s\langle\Delta r\rangle. \qquad (18.5.8.2)$$

The theoretical Luzzati plots of R are nearly linear for small-to-medium $s = 2\sin\theta/\lambda$ (see Fig. 18.5.8.1). If we substitute this R in the least-squares estimate (18.5.8.1) and use the three-dimensional-Gaussian relation $\sigma(r) = 1.085\langle\Delta r\rangle$, some manipulation (Cruickshank, 1999) along the lines of Section 18.5.6 eventually yields a statistically based formula,

Copyright © 2006 International Union of Crystallography

references