

## References

- BECKER, P., COPPENS, P. & ROSS, F. K. (1973). *J. Amer. Chem. Soc.* **95**, 7604.
- BRILL, H., DIETRICH, H. & DIERKS, H. (1971). *Acta Cryst.* **B27**, 2003–2018.
- COPPENS, P. (1967). *Science*, **158**, 1577.
- COPPENS, P. (1971). *Acta Cryst.* **B27**, 1931–1938.
- COPPENS, P. (1972). *Israel J. Chem.* pp. 85–91.
- COPPENS, P. & COULSON, C. A. (1967). *Acta Cryst.* **23**, 718–720.
- COPPENS, P. & HAMILTON, W. C. (1968). *Acta Cryst.* **B24**, 925–929.
- COPPENS, P., SABINE, T. M., DELAPLANE, R. G. & IBERS, J. A. (1969). *Acta Cryst.* **B25**, 2451–2458.
- COPPENS, P. & VOS, A. (1971). *Acta Cryst.* **B27**, 146–158.
- COULSON, C. A. & MOFFITT, W. E. (1949). *Phil. Mag.* **40**, 1–10.
- CRUICKSHANK, D. W. J. (1949). *Acta Cryst.* **2**, 65–82.
- CRUICKSHANK, D. W. J. & ROLLETT, J. S. (1953). *Acta Cryst.* **6**, 705–707.
- DUCKWORTH, J. A. K., WILLIS, B. T. M. & PAWLEY, G. S. (1970). *Acta Cryst.* **A26**, 263–271.
- HANSON, J. C., SIEKER, L. C. & JENSEN, L. H. (1973). *Acta Cryst.* **B29**, 797–808.
- HIGGS, P. W. (1953). *Acta Cryst.* **6**, 232–241.
- HIRSHFELD, F. L. (1971). *Acta Cryst.* **B27**, 769–781.
- JONES, D. S., PAUTLER, D. & COPPENS, P. (1972). *Acta Cryst.* **A28**, 635–645.
- MATTHEWS, D. A. & STUCKY, G. D. (1971). *J. Amer. Chem. Soc.* **93**, 5954–5959.
- REES, B. & COPPENS, P. (1973). *Acta Cryst.* **B29**, 2516–2528.
- STEWART, R. F. (1968). *Acta Cryst.* **A24**, 497–505.
- TAYLOR, J. C. & SABINE, T. M. (1972). *Acta Cryst.* **B28**, 3340–3351.
- WILLIS, B. T. M. (1969). *Acta Cryst.* **A25**, 277–300.

*Acta Cryst.* (1974). **B30**, 261

## The Reliability of Crystallographic Structural Information\*

BY S. C. ABRAHAMS

*Bell Laboratories, Murray Hill, New Jersey 07974, U.S.A.*

(Received by the Hamilton Symposium Committee 15 June 1973)

A review is presented, drawn largely from the work of Walter Hamilton, of the optimal design of a crystallographic experiment, of structure-factor measurement and the estimation of associated standard deviations, of pitfalls connected with the large scale computations required for structural refinement, of methods for choosing the best structural model derivable from the experiment, of means for assessing the quality of the information finally extracted, and of tests for recognizing aberrant data after refinement is complete.

### 1. Introduction

The field that the present paper seeks to review as part of the Hamilton Symposium is one that long held a central place in Walter Hamilton's crystallographic interests, and to which he made many major contributions. The field is concerned with the derivation of objective methods for assessing the quality of experimental integrated intensity measurements, and with the improvement both of the measurements and their associated correction factors: with the error in the theoretical models used in crystal structure refinement and with the associated computing methods; with choosing among alternative models derived from a crystallographic experiment; and with assessing the significance of the final parameters determined in the refinement.

It is hoped this review will help underscore the coherence and importance of this aspect of Walter Hamilton's work, and at the same time highlight the broad conceptual framework in which he habitually thought.

### 2. Optimum experimental design

#### 2(a). Integrated intensity measurement

The best strategy for achieving the maximum precision in a crystallographic experiment of given duration has been frequently discussed since Parrish (1956) and Mack & Spielberg (1958) showed, for the case of X-ray powder diffractometry, that the optimum division of time between background determination and peak scanning is proportional to the ratio of the square roots of their respective counting rates. This result is based on the assumption of a Gaussian distribution of diffracted X-ray quanta with variance given by the mean. Extension to the single-crystal diffractometry case was made in a group of three papers (Hamilton, 1967; Shoemaker, 1968; Shoemaker & Hamilton,

\*Editorial note: – This paper was presented at a memorial symposium in honour of Walter C. Hamilton, a former Co-editor of *Acta Crystallographica*, held on 15 June 1973 at Brookhaven National Laboratory, and sponsored jointly by the American Crystallographic Association and the Brookhaven National Laboratory.

1972), in which relationships were also developed giving the relative time that should be spent in measuring each reflection in order to maximize the precision of a given parameter in subsequent least-squares refinement. Hamilton and Shoemaker point out that the actual degree of parameter weight-optimization possible has not yet been determined, nor has the influence of systematic error on their equations. Counting statistics have also been the basis of attempts by Mackenzie & Williams (1973), Killean (1973) and Grant (1973) to determine the optimum counting time for a reflection giving maximum precision in the intensity.

The variance due to counting statistics alone is likely to dominate the total variance in the corrected integrated intensity only in the case of small absolute counts, *e.g.*, less than 1000. The relative importance of the counting-statistic variance in a given reflection is readily determined from the *variance ratio*  $Y$ , where  $Y = V'(\bar{I})/V''(\bar{I})$ , and the estimated variance  $V'(\bar{I})$  is given by:

$$V'(\bar{I}) = \frac{1}{j-1} \sum_{n=1}^j (\bar{I} - I_n)^2 \quad (1)$$

and

$$\bar{I} = \frac{1}{j} \sum_{n=1}^j I_n. \quad (2)$$

The  $j$  members of a form measured under conditions of pseudoreplication thus give an external measure of variance and  $V''(\bar{I})$ , the correctly calculated variance of the mean intensity due to the counting statistics (*cf.* Cetlin & Abrahams, 1963), gives an internal measure.

The presence of random error only in the integrated intensity measurement will give a variance ratio of about unity: systematic error generally results in  $V'(\bar{I}) \gg V''(\bar{I})$ . The normally small value of  $j$  leads to difficulties in analyzing the variance ratios in terms of the expected  $F$ -distribution. Schultz & Huber (1971) have shown the  $Y$  variance ratios to have an approximately  $\chi^2$  distribution with  $j-1$  degrees of freedom, with  $\chi^2$  values corresponding to each group of symmetry-equivalent reflections ordered as a function of an experimental variable such as intensity, or scattering angle.

*Quality control* of integrated intensity measurements in the course of the actual experiment (*i.e.*, in real on-line time) requires that counting statistics, as a criterion for optimization of counting times, be accepted only if an objective determination is made, for each reflection, that systematic error is not present. If systematic error is present, as indicated by the magnitudes of the variance ratios, analysis such as Schultz & Huber's should be made during the measurement sequence to identify individual terms leading to excessive  $V'(\bar{I})$  values. Such analysis should be followed by corrective action and improved remeasurement.

The variation of variance ratio, averaged over groups of ten symmetry-independent reflections ordered on the intensity magnitudes, with averaged intensity is shown in Fig. 1 for the cubic crystal  $\text{KTa}_x\text{Nb}_{1-x}\text{O}_3$

( $x \approx 0.77$ ). The percentage error in intensity for this data, due to counting statistics, ranges from about 0.1 to about 1.

### 2(b). Radiation damage detection

Successful control of intensity measurement leading to structure amplitudes of highest possible accuracy is predicated on complete immunity of the specimen crystal to radiation damage. A recent International Union of Crystallography (1973) study has shown that, in nine crystals of widely different chemical composition, changes in intensity as a function of radiation exposure were detectable in each case. In no case did all reflections studied change uniformly. A diagnostic procedure in which at least three strong, three medium and three weak reflections distributed throughout reciprocal space are measured periodically throughout the course of data collection was recommended. Such a procedure should be part of a good quality control program, since the results of the measurement are generally assumed to apply to the crystal in a state undisturbed by radiation damage.

## 3. The measured structure factor

### 3(a). Reduction of intensity to structure factor

The accurately measured integrated intensities are generally reduced to structure factors before calculations such as refinement of a structural model may be undertaken. Reduction requires correction for several processes, including absorption, extinction, thermal and other diffuse scattering and multiple scattering; Walter Hamilton was concerned about all of these, and especially about extinction, on which he wrote four papers (Hamilton, 1957, 1958, 1963; Coppens &

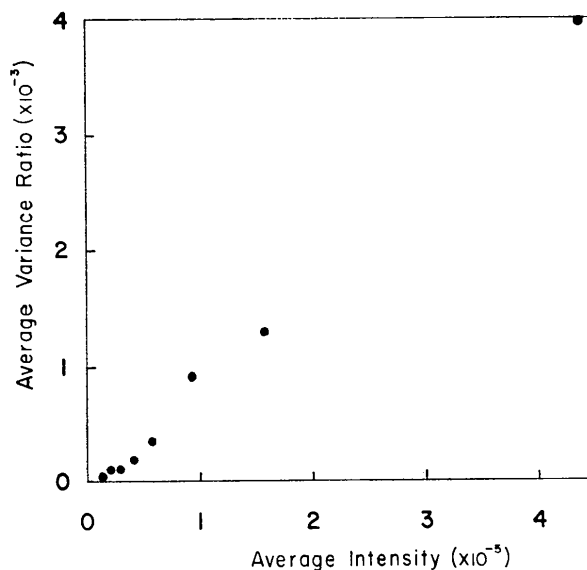


Fig. 1. Variation of averaged variance ratio with averaged intensity, for the cubic crystal  $\text{KTa}_x\text{Nb}_{1-x}\text{O}_3$  ( $x \approx 0.77$ )

Hamilton, 1970). He found that, based on Zachariasen's (1967) secondary extinction treatment, crystals not infrequently exhibit anisotropy in the extinction effect. Together with Coppens he presented a formalism by which anisotropic extinction coefficients could be determined either for Type I (extinction dominated by mosaic spread) or Type II (extinction dominated by particle size) crystals. He also modified his version of the *ORFLS* least-squares refinement program to refine the corresponding six components of a symmetrical second-order tensor.

### 3(b). Intensity measurement projects

Two interlaboratory projects have been designed to provide a measure of the accuracy of common methods of integrated intensity measurement. In the American Crystallographic Association (1967) project, the intensities of all *hkl* reflections, with  $(\sin \theta)/\lambda \leq 1.0 \text{ \AA}^{-1}$ , were measured on a ground sphere of  $\text{CaF}_2$  circulated among seven different laboratories. In the International Union of Crystallography (1970a) project, the intensities of all reflections, with  $(\sin \theta)/\lambda \leq 0.5 \text{ \AA}^{-1}$ , were measured on different crystals of (+)-tartaric acid by seventeen laboratories. Walter Hamilton enthusiastically undertook the statistical analysis of the data generated by both projects, primarily by analysis-of-variance and use of interexperimental *R* values.

The conclusion was drawn, for the ACA project, that most experiments determined  $F^2$  values within 5% of the probable true values, and that few experiments measured  $F^2$  to better than 2%. The largest systematic errors were found to be associated with the scattering angle: a dependence on intensity was also found, in a separate analysis, by Mackenzie & Maslen (1968). The IUCr project allowed the inference to be drawn that two scaled experimental sets of structure factors, measured under circumstances similar to those of the project, would differ most probably by 6%, agree no better than 3% and usually differ by no more than 10% (one set however differed by nearly 50%). The absolute accuracy of any set was probably not better than 5 to 6%. Several experiments had significant correlations between the experimental deviations from the average and intensity: other error sources also appeared to be present in addition to those simply associated with angle and intensity, such as variation in specimen characteristics.

### 3(c). Standard deviation of the structure factor

Determination of the fully corrected set of structure factors,  $F_{\text{meas}}$ , as in § 3(a), is incomplete without simultaneous estimation of the standard deviation  $\sigma F_{\text{meas}}$  in the sense that any quantitative determination is incomplete without a reliable measure of the associated uncertainty. A difficulty in applying analysis-of-variance in 3(b) was the lack of reliable  $\sigma F_{\text{meas}}$  for the various sets of experimental  $F_{\text{meas}}$ . An objective estimate of  $\sigma F_{\text{meas}}$  has been given (Abrahams, Bernstein & Keve, 1971; Abrahams, 1964), in which the vari-

ance in  $(F_{\text{meas}})^2$  is a sum of identifiable component variances:

$$\begin{aligned} \sigma^2(F_{\text{meas}})^2 = & \sigma^2(F_{\text{meas}_{\text{counting statistics}}})^2 \\ & 4 \sum_j k_j^2 \times 10^{-4} (F_{\text{meas}})^4 + \text{larger of} \\ & \{4 \sum_j k_j'^2 \times 10^{-4} (F_{\text{meas}})^4 \text{ or } V'(\bar{F})^2\} \quad (3) \end{aligned}$$

where  $V'(\bar{F})^2$  is defined as in equation (1), and the  $k$  and  $k'$  terms are estimated percentage errors. The larger of the third and fourth terms in equation (3) is used, since both estimate the variance due to anisotropy in measurement conditions, hence causing deviation from known symmetry (*cf.* Schultz, 1971).

The expression for the variance  $\sigma^2(F_{\text{meas}})^2$  used by Walter Hamilton, especially in the extensive series of investigations on precise neutron diffraction structural studies of protein and nucleic acid components that he was engaged with to the very end, was:

$$\begin{aligned} \sigma^2(F_{\text{meas}})^2 = & \sigma^2(F_{\text{meas}_{\text{counting statistics}}})^2 + K^2 \\ & \times 10^{-4} (F_{\text{meas}})^4. \quad (4) \end{aligned}$$

In this series (*e.g.*, Frey, Koeztle, Lehmann & Hamilton, 1973; Koeztle, Hamilton & Parthasarathy, 1972)  $K^2$  was derived after least-squares structural refinement was essentially complete: values between 1 and 49 resulted in  $\{|F_{\text{meas}}| - |F_{\text{calc}}|/\sigma F_{\text{meas}}\}$  becoming, on average, independent of  $|F_{\text{meas}}|$ . These  $K^2$  values may be interpreted in terms of equation (3) as indicating the largest single error in addition to counting statistics, for this series, to lie between 0.5 and 3.5% of  $(F_{\text{meas}})^2$ , including that due to  $V'(\bar{F})^2$ . The standard deviation of an observation of unit weight, as given by:

$$S = \sum \{(|F_{\text{meas}}| - |F_{\text{calc}}|/\sigma F_{\text{meas}})^2 \div (m - n)\}^{1/2} \quad (5)$$

where the summation is over all  $m$  observations, and  $n$  variables are used in least-squares minimization, was approximately 1.5 for both examples cited. The average value of  $\sigma F_{\text{meas}}$  obtained from equation (4) is hence about one-third too small.

Systematic error in  $F_{\text{meas}}$  and  $\sigma F_{\text{meas}}$  can be detected definitively by comparison with a second, independently measured, data set by use of normal probability plots (Abrahams & Keve, 1971). This method is discussed further in § 5(b).

## 4. Crystal structure refinement

Walter Hamilton's interests in the use of the fast digital computer for controlling crystallographic experiments in data acquisition and for refining crystal structure models stimulated numerous advances in technique (*cf.* Hamilton, 1968). He improved many standard programs in addition to others he wrote himself. The possibility that some versions of widely used standard programs might contain logical error concerned him, and in his term on the IUCr Commission on Crystallographic Computing he became deeply involved in the

early discussions on planning and testing that led to formulation of a set of standard tests for crystallographic programs (Ahmed, Cruickshank, Larson & Stewart, 1972). Looking to the future, he organized a symposium on April 18, 1972 on *Computational Needs and Resources in Crystallography* (1973) under the auspices of the National Research Council. Of particular interest was a discussion on the possibility of establishing regional computer centers or networks of computers for large scale crystallographic computations.

Crystal structure refinement of materials under conditions for which the effects of anomalous dispersion are significant may be handled either by correcting the  $F_{\text{meas}}$  for dispersion or by including the effects in the  $F_{\text{calc}}$ . Ibers & Hamilton (1964) pointed out the advantages of the latter approach, and accordingly modified a standard least-squares refinement program in a manner which is now widely and routinely used.

## 5. Crystal structure model hypothesis testing

### 5(a). Hamilton $\mathcal{R}$ -test

It frequently happens in course of refining a crystal structure, especially by the method of least squares, that a choice must be made between two or more models differing in the number of variable parameters with which the structure is described. The assumption that one model fits the observations better than another, with due allowance for the differences in parameters varied, forms a statistical hypothesis capable of being tested by statistical methods. In one of the most widely cited papers in modern structural crystallography, Walter Hamilton (Hamilton, 1961 and especially 1965) developed a significance test based on comparison of weighted  $R$  values ( $\mathcal{R} = wR_1/wR_2$ ), together with tables for the significance level of the determination. Unlike use of the Student  $t$ -distribution (for small samples) for testing deviations on individual parameters derived from least-squares refinement of the various models, the  $\mathcal{R}$ -test does not presuppose linearity in the equations of condition.

In the absence of systematic error the  $\mathcal{R}$ -test allows the hypothesis that Model 1 is correct, of Models 1 and 2 corresponding to final refinement values of  $wR_1$  and  $wR_2$ , to be rejected at the  $100\alpha\%$  significance level if

$$\mathcal{R}_{\text{experimental}} > \mathcal{R}_{\text{calculated}, \alpha} \quad (6)$$

As the inequality of (6) approaches equality, the possibility of rejecting the wrong hypothesis increases, for the assumption that the crystallographic experiment contains random error only is unlikely to be correct. A cautious approach would reject Model 1 only if

$$\mathcal{R}_{\text{experimental}} - 1 \gg \mathcal{R}_{\text{calculated}, \alpha} - 1 \quad (7)$$

A typical case in which the wrong hypothesis might be rejected would be an experiment in which angle-dependent systematic error is present that differs in magnitude in different but corresponding segments of

reciprocal space: the resulting anisotropy could simulate and be fitted with apparent anisotropic thermal parameters, even if these are genuinely isotropic. If inequality (7) is not satisfied, additional information should be sought before either hypothesis is rejected.

Walter Hamilton's analysis showed that the  $\mathcal{R}$ -significance test is strictly applicable only if either the normal weighted  $R$  values or the generalized weighted  $R$  values for the two models are compared. Ford & Rollett (1970) point out that use of simple unweighted  $R$  values, as is sometimes found in the literature, can lead to highly deceptive results. The tabulation of significance points for the  $\mathcal{R}$ -test given by Hamilton extends to a maximum value of 120 degrees of freedom: larger values are obtainable by interpolation. A simple analytic expression has been given by Pawley (1970) which closely approximates the exact values, and which is readily adapted to inclusion in existing refinement programs.

### 5(b). $\delta R$ -Normal probability plot

Readily generated additional information for aiding hypothesis testing is available by computing the  $\delta R$ -normal probability plot for each model. The statistic  $\delta R_i = |F_{\text{meas}_i} - F_{\text{calc}_i}| / \sigma F_{\text{meas}_i}$ , obtained from the final least-squares refinement cycle with  $F_{\text{meas}_i}$  and  $\sigma F_{\text{meas}_i}$  on the same scale as  $F_{\text{calc}_i}$ , is ordered and plotted against the quantiles expected for a normal distribution (Abrahams & Keve, 1971; Hamilton & Abrahams, 1972). If the distribution of  $\delta R_i$  is Gaussian, the resulting plot is a straight line of unit slope and zero intercept. The greater the departure from Gaussian, the less acceptable is the model.

Use of the  $\delta R$  plot is illustrated by some recent results obtained in a study of the piezoelectric nonlinear optic chalcopyrite  $\text{CuGaS}_2$ . A total of 134 symmetry-independent  $F_{\text{meas}}$ , with  $\sigma F_{\text{meas}}$  obtained by equation (3), were measured on a sphere with  $r = 0.0945$  mm and  $\mu r = 1.724$ . Refinement with isotropic temperature factors gave  $R = 0.031$ ,  $wR = 0.0460$  and  $S = 0.97$ . Further refinement with anisotropic temperature factors led to  $R = 0.030$ ,  $wR = 0.0416$  and  $S = 0.89$ . The Hamilton  $\mathcal{R}$ -ratio is hence 1.107 and the calculated ratio  $\mathcal{R}_{5,124,0.005}$  is 1.068: equality (6) is obeyed, but not (7). The  $\delta R$ -plot for the isotropic case is shown in Fig. 2, that for the anisotropic case in Fig. 3. Initial examination shows no major systematic deviations from linearity for either plot: systematic error, of the type expected for inadequate absorption correction for example, is not of major consequence. Excluding two or three outliers which, if present, are automatically at the extrema of probability plots, the improved linearity in Fig. 3 as compared with Fig. 2 greatly increases the confidence with which the hypothesis that the thermal vibrations in  $\text{CuGaS}_2$  are isotropic may be rejected. Closer examination of Figs. 2 and 3 reveals intercepts close to zero, but slopes of 0.82 and 0.77 respectively.

The unweighted slope of the  $\delta R$ -plot is equivalent to the magnitude of  $S$ , the standard deviation of an

observation of unit weight, given by equation (5): for the anisotropic case,  $S=0.79$ . The inference may hence be made that the average value of  $\sigma F_{\text{meas}}$  for  $\text{CuGaS}_2$  is overestimated by about 28%, but that distribution of the  $\delta R_i$  appears to be normal random.

### 6. Quality of structural information from final model

Critical evaluation of the quality of structural information obtained crystallographically is very close to the subject on which Walter Hamilton had planned to speak this month at the Conference on Critical Evaluation (1973) in Dartmouth College, and on which David Shoemaker will instead speak. It will be assumed in this section that the structural model refined was free

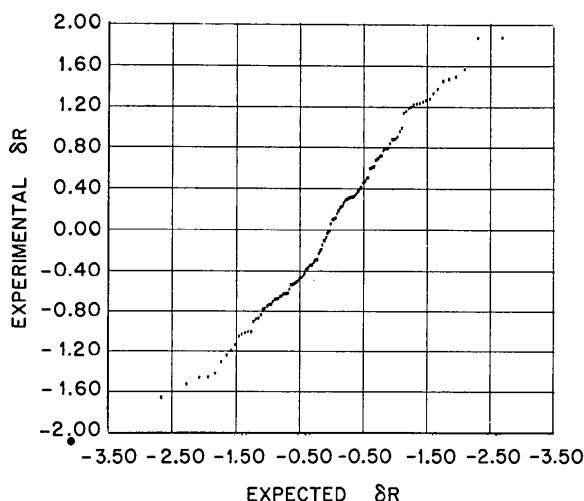


Fig. 2.  $\delta R$ -normal probability plot for a fully refined isotropic model ( $R=0.031$ ), cf. Fig. 3.

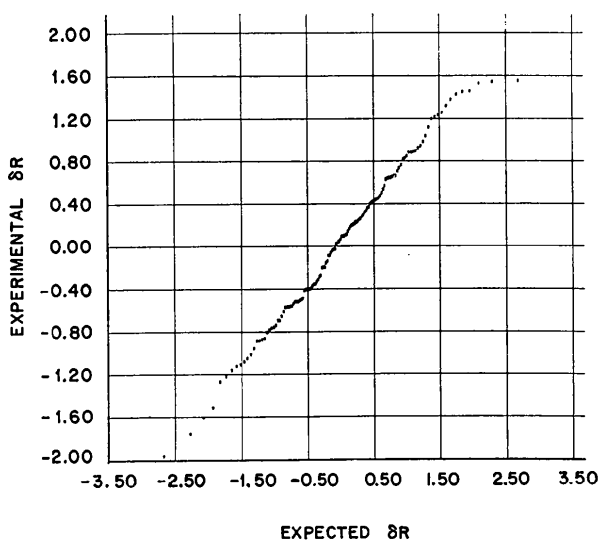


Fig. 3.  $\delta R$ -normal probability plot for a fully refined anisotropic model ( $R=0.030$ ). The major central section of the plot is noticeably more linear than in Fig. 2.

of gross blunders, and that neither the crystal measured nor the measurement system suffered from perceptible signs of instability.

#### 6(a). Comparison of results from interlaboratory intensity measurement project

A direct measure of the quality of the final results of a given structural study is obtainable by comparing the positional and vibrational coordinates from the fully refined model with other independently measured and refined sets of coordinates. Seventeen independently measured sets of  $F_{\text{meas}}$  for (+)-tartaric acid were generated in the International Union of Crystallography (1970a) intensity project: two sets consisted only of one zone of data. Walter Hamilton (International Union of Crystallography, 1970b) used the remaining 15 data sets in separate refinements of the known model, by means of the method of least squares. It was immediately apparent that the magnitude of the usual agreement factor  $R$  is highly insensitive to systematic error in the  $F_{\text{meas}}$ : the range, over all data sets, is 0.034 to 0.112 and the value of  $R$  for the set that systematically differed most from average was 0.057. Hence, a low value of  $R$  is not necessarily an indicator of high quality.

Analysis of the variation in the 29 positional and 60 thermal vibrational parameters, excluding the hydrogen atoms, obtained for each data set is given in the published report, and additional analysis is expected to be presented by W. L. Nicholson and J. W. Tukey at the Dartmouth Conference (1973). It is sufficient here only to point out the conclusion that the estimated standard deviations obtained from the inverse matrix in the least-squares refinement for these data sets were not infrequently too small by a factor of about 2, and all were on average too small by about the factor  $\sqrt{2}$ . The underestimation in the thermal vibration standard deviations is even greater, by an *additional* factor of about 2.

#### 6(b). $\chi^2$ Comparison of two parameter sets

In case two independent sets of  $F_{\text{meas}}$  are available for a given material, with two resulting sets of position and thermal vibration parameters, the differences between corresponding parameters may be defined in terms of the statistic:

$$\delta p_i = \left| |p(1)_i| - |p(2)_i| \right| / \left\{ \sigma^2 p(1)_i + \sigma^2 p(2)_i \right\}^{1/2}. \quad (8)$$

Walter Hamilton (1969) proposed examining the hypothesis that the  $\delta p_i$  are drawn from a normal population with unit variance and zero mean by testing:

$$\delta^2 p = \sum_{i=1}^N \delta^2 p_i \quad (9)$$

against  $\chi^2$  with  $N$  degrees of freedom. If the calculated values of  $\delta^2 p$  exceeds the tabulated value of  $\chi^2_{N, \alpha}$ , the hypothesis may be rejected with a  $100\alpha\%$  probability of falsely rejecting a true hypothesis. It is not necessary in (9) to sum over all  $N$  parameters: a subset  $m$  may be

used, and tested as  $\chi^2_{m,\alpha}$ . The Student *t*-test is not an adequate substitute for examining the significance of an individual value of  $\delta p_i$ . In his 1969 paper, Hamilton used the  $\chi^2$  test to compare the parameters obtained in studying six crystals of different materials by independent X-ray and neutron diffraction investigations. He found  $\delta^2 p$  to be comparable to or less than  $\chi^2_{0.05}$  for most nonhydrogen atom position coordinates in these crystals: an increase by a factor of 2 in the estimated standard deviation [see § 6(a)] for these parameters would make all but one  $\sigma^2 p < \chi^2_{0.05}$ , i.e., there are no significant differences at the 5% level in the two sets of investigations, except for the *z* coordinates of completely deuterated  $\alpha$ -oxalic acid dihydrate. Large and highly significant differences, however, remain in many of the thermal and hydrogen atom position parameters.

### 6(c). $\delta p$ -Half-normal probability plot

The  $\delta p_i$ -statistic distribution in equation (8) may also be conveniently analyzed by normal probability plotting methods [see section 5(b)]. The  $\delta p_i$  are ordered (note all  $\delta p_i \geq 0$ ) in increasing magnitude and plotted against the expected quantiles for half-normal order statistics (i.e., ranked moduli of normal observations). Tables of ranked moduli, for sample sizes to 41, are given in Hamilton & Abrahams (1972). For larger samples, the percentage points of the normal distribution may be used, with  $P(X_i) = (2j - 2i + 1)/2j$ ,\* where *j* is the number of  $\delta p_i$  in the sample and corresponding values of  $X_i$  are tabulated, for example, in *Tables of Normal Probability Functions* (1953). Use of these percentage points results in small errors in the extrema of the ranked moduli: these are eliminated in tables appearing in Vol. IV of *International Tables for X-ray Crystallography* (1974).

Examples of  $\delta p$  half-normal probability plots are given in Fig. 4, taken from Verbist, Lehmann, Koetzle & Hamilton (1972): three independent investigations of the crystal structure of L-asparagine monohydrate are compared. The solid line represents an ideal normal distribution of  $\delta p_i$ . The linearity of Figs. 4(a), (b), and (c) indicate that the  $\delta p_i$  for these cases are approximately normally distributed, but with pooled standard deviations [see equation (8)] underestimated by a factor of about 1.6 for experiments 1 and 2 and overestimated by about 0.8 for experiments 1 and 3. The comparison of position parameters for experiments 1 and 3 in Fig. 4(d) indicates the presence of systematic error, and an underestimation of the pooled standard deviation by about the factor 1.4. These factors may be used in more realistically estimating the standard deviations.

## 7. Model-guided recognition of aberrant data

At the conclusion of crystal structure refinement, based on the best *F*<sub>meas</sub> and  $\sigma F$ <sub>meas</sub> available, large values

of  $\delta R$  [§ 5(b)] may be examined for evidence of error. Large  $\delta R$  terms necessarily accumulate at the extrema of the  $\delta R$ -normal probability plot. It is generally inappropriate to single out a  $\delta R$  term, no matter how large, if it forms part of a smooth distribution since the remaining terms are clearly subject to similar error. The  $\delta R$ -plot, however, provides an unmistakable diagnostic for the presence of unexpected systematic error in a class of measurement.

An illustrative example of the diagnostic use of  $\delta R$ -plots comes from recent measurements on the piezoelectric chalcopyrite CuInS<sub>2</sub>, isomorphous with the material referred to in § 5(b). Based on 310 pairs of *F*<sub>meas</sub>,  $\sigma F$ <sub>meas</sub> the 'final' least-squares refinement cycle gave *R* = 0.047, *wR* = 0.060 and *S* = 1.08: the corresponding  $\delta R$ -plot shown in Fig. 5 contains a large and systematic departure from linearity over the entire central portion. Evaluation of the experimental conditions showed that one complete layer of reciprocal space, {*h*3*l*}, had been measured with a lossy co-axial connection between preamplifier and amplifier. The scale factor for this layer was aberrantly different from the remainder of the data. Correction by a scale factor 9.0% larger than the single overall scale factor led to further refinement, with *R* = 0.038, *wR* = 0.052 and *S* = 0.99: the corresponding  $\sigma R$ -plot shown in Fig. 6 is close to linear, except for a few outliers all of which are of low intensity with  $\sigma F$ <sub>meas</sub> that are probably systematically underestimated. It may be noted that although the single position parameter changed by only 0.3 $\sigma$ , all the thermal parameters *decreased* by a maximum of 1.8 $\sigma$ .

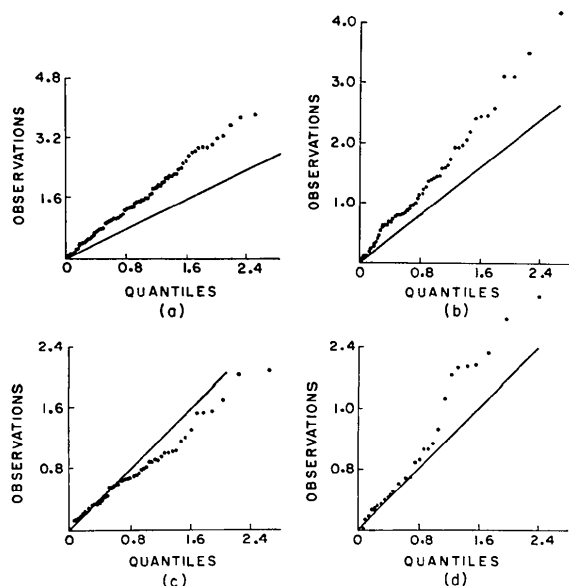


Fig. 4. Four  $\delta p$ -half normal probability plots comparing (a) thermal and (b) position parameters for experiments 1 and 2 and (c) thermal and (d) position parameters for experiments 1 and 3. The three experiments were made in different laboratories. [from Verbist *et al.* (1972)].

\* A more accurate value is given (private communication from J. W. Tukey) by  $P(X_i) = (3j - 3i + 2)/(3j + 1)$ .

## 8. Conclusions

1. If structure factors are measured for the purpose of crystal structure determination by the method of least squares, the corresponding  $\sigma F_{\text{meas}}$  should be estimated with a comparable level of care.

2. Diagnostic procedures for detection of possible radiation damage should be followed in all crystallographic experiments.

3. The hypothesis that one structure model should be rejected in favor of another, based on Hamilton's (1965)  $\mathcal{R}$ -test, assumes the presence only of random error. A more cautious approach permits rejection only

if  $\mathcal{R}_{\text{experimental}} - 1 \gg \mathcal{R}_{\text{calculated}, \alpha} - 1$ , in the absence of further independent data or tests.

4. In addition to the inclusion of  $R$ ,  $wR$  and  $S$  factors in published crystal structure papers, the  $\delta R$ -normal probability plot should always be given as a visual indicator of the quality of the entire data.

5. In case two independent sets of  $F_{\text{meas}}$  (and  $\sigma F_{\text{meas}}$ ) are available for the same material,  $\delta p$  half-normal and  $\delta m$ -normal probability plot analysis provide the basis for a realistic estimate of the joint error distribution.

## References

- ABRAHAMS, S. C. (1964). *Acta Cryst.* **17**, 1327-1328.
- ABRAHAMS, S. C., BERNSTEIN, J. L. & KEVE, E. T. (1971). *J. Appl. Cryst.* **4**, 284-290.
- ABRAHAMS, S. C. & KEVE, E. T. (1971). *Acta Cryst.* **A27**, 157-165.
- AMERICAN CRYSTALLOGRAPHIC ASSOCIATION (1967). *Acta Cryst.* **22**, 1-6.
- AHMED, F. R., CRUICKSHANK, D. W. J., LARSON, A. C. & STEWART, J. M. (1972). *Acta Cryst.* **A28**, 365-393.
- CETLIN, B. B. & ABRAHAMS, S. C. (1963). *Acta Cryst.* **16**, 943-946.
- Computational Needs and Resources in Crystallography* (1973). National Academy of Sciences, Washington, D.C.
- Conference on Critical Evaluation of Chemical and Physical Structural Information* (1973). Edited by D. R. LIDE and M. A. PAUL. National Academy of Sciences, Washington, D.C.
- COPPENS, P. & HAMILTON, W. C. (1970). *Acta Cryst.* **A26**, 71-83.
- FORD, G. C. & ROLLETT, J. S. (1970). *Acta Cryst.* **A26**, 162.
- FREY, M. N., KOETZLE, T. F., LEHMANN, M. S. & HAMILTON, W. C. (1973). *J. Chem. Phys.* **58**, 2547-2556.
- GRANT, D. F. (1973). *Acta Cryst.* **A29**, 217.
- HAMILTON, W. C. (1957). *Acta Cryst.* **10**, 629-634.
- HAMILTON, W. C. (1958). *Acta Cryst.* **11**, 585-594.
- HAMILTON, W. C. (1961). Amer. Cryst. Assoc. Summer Meeting, Boulder, Colorado. Abstract H-5.
- HAMILTON, W. C. (1963). *Acta Cryst.* **16**, 609-611.
- HAMILTON, W. C. (1965). *Acta Cryst.* **18**, 502-510.
- HAMILTON, W. C. (1967). Amer. Cryst. Assoc. Summer Meeting, Minneapolis, Minnesota. Abstract E-6.
- HAMILTON, W. C. (1968). *J. Comp. Phys.* **2**, 417-438.
- HAMILTON, W. C. (1969). *Acta Cryst.* **A25**, 194-206.
- HAMILTON, W. C. & ABRAHAMS, S. C. (1972). *Acta Cryst.* **A28**, 215-218.
- IBERS, J. A. & HAMILTON, W. C. (1964). *Acta Cryst.* **17**, 781-782.
- International Tables for X-ray Crystallography* (1974). Vol. IV. Birmingham: Kynoch Press. (In preparation.)
- INTERNATIONAL UNION OF CRYSTALLOGRAPHY (1970a). *Acta Cryst.* **A26**, 1-18.
- INTERNATIONAL UNION OF CRYSTALLOGRAPHY (1970b). *Acta Cryst.* **A26**, 18-24.
- INTERNATIONAL UNION OF CRYSTALLOGRAPHY (1973). *Acta Cryst.* **A29**, 111-116.
- KILLEAN, R. C. G. (1973). *Acta Cryst.* **A29**, 216-217.
- KOETZLE, T. F., HAMILTON, W. C. & PARTHASARATHY, R. (1972). *Acta Cryst.* **B28**, 2083-2090.
- MACK, M. & SPIELBERG, N. (1958). *Spectrochim. Acta*, **12**, 169-178.

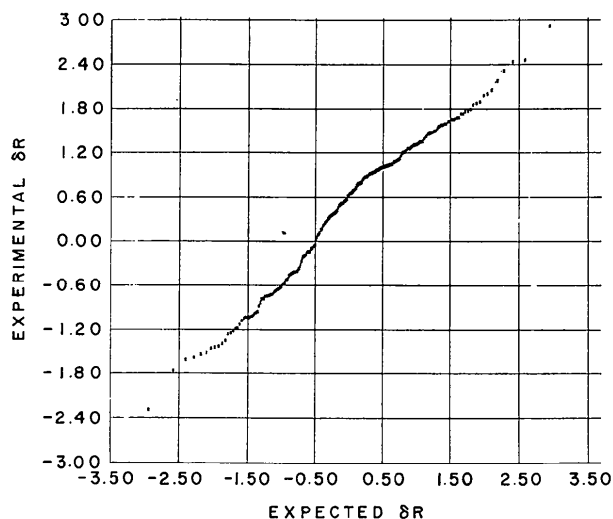


Fig. 5.  $\delta R$ -normal probability plot for a fully refined anisotropic model ( $R=0.047$ ). The 'hump' in the central section was indicative of a major experimental or model error (cf. Fig. 6).

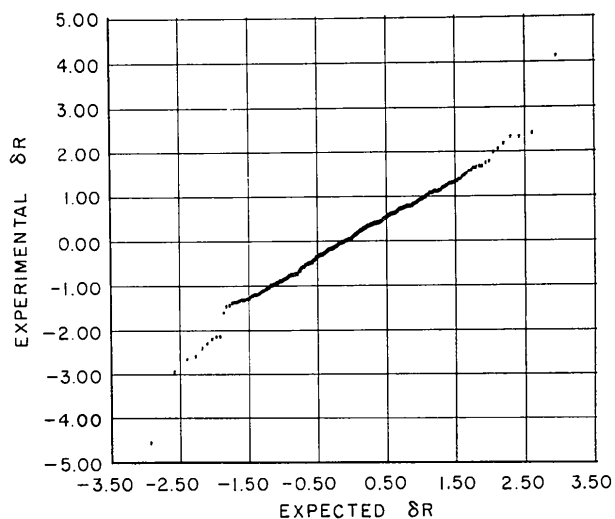


Fig. 6.  $\delta R$ -normal probability plot for model used in Fig. 5, on full refinement and after elimination of the experimental error.

- MACKENZIE, J. K. & MASLEN, V. W. (1968). *Acta Cryst.* A24, 628–639.
- MACKENZIE, J. K. & WILLIAMS, E. J. (1973). *Acta Cryst.* A29, 201–204.
- PARRISH, W. (1956). *Philips Tech. Rev.* 17, 206–221.
- PAWLEY, G. S. (1970). *Acta Cryst.* A26, 691–692.
- SCHULTZ, H. (1971). *Acta Cryst.* A27, 540–544.
- SCHULTZ, H. & HUBER, P. J. (1971). *Acta Cryst.* A27, 536–539.
- SHOEMAKER, D. P. (1968). *Acta Cryst.* A24, 136–142.
- SHOEMAKER, D. P. & HAMILTON, W. C. (1972). *Acta Cryst.* A28, 408–411.
- Tables of Normal Probability Functions* (1953). National Bureau of Standards, Applied Mathematics Series No. 23. U. S. Government Printing Office, Washington, D.C.
- VERBIST, J. J., LEHMANN, M. S., KOETZLE, T. F. & HAMILTON, W. C. (1972). *Acta Cryst.* B28, 3006–3013.
- ZACHARIASEN, W. H. (1967). *Acta Cryst.* 23, 558–564.

*Acta Cryst.* (1974). B30, 268

## Stereochemistry of Cooperativity Effects in the Prosthetic Group of Coboglobin\*

BY JAMES A. IBERS, JOSEPH W. LAUHER AND ROBERT G. LITTLE

*Department of Chemistry, Northwestern University, Evanston, Illinois 60201, U.S.A.*

(Received by the Hamilton Symposium Committee 15 June 1973)

Based on diffraction studies of model cobalt-porphyrin systems an estimate of 0.38 Å is obtained for the maximum movement of the proximal histidine residue relative to the mean plane of the porphyrin ring on oxygenation of coboglobin, the cobalt analogue of hemoglobin. This estimate is about one-half that currently believed to exist in hemoglobin. It is suggested that an additional stereochemical basis for cooperativity of oxygenation in coboglobin may be the transition of the initially non-planar porphyrin core to essential planarity on oxygenation of the cobalt-heme.

The oxygenation of hemoglobin is a cooperative phenomenon in which the last oxygen molecule enters with less difficulty than the others because of an interaction between protein sub-units. In a series of classic studies on model systems Hoard and coworkers have provided a stereochemical explanation for the expected changes in the prosthetic group of hemoglobin upon oxygenation. [See Hoard (1971) for a recent review of the subject.] Perutz, with direct experimental results on various hemoglobins as well as Hoard's studies of model systems, has postulated a 'trigger' mechanism for this cooperativity (Perutz, 1970*a, b*; Perutz & TenEyck, 1971). In essence, oxygenation results in a change of the state of the iron atom of the heme group from high-spin Fe(II) to low-spin Fe(III) or low-spin Fe(II), with a concomitant movement of the Fe atom toward† the mean plane of the porphyrin ring estimated to be 0.75 to 0.95 Å. Since this Fe atom is covalently bonded to the proximal histidine residue of the globin, this residue is shifted toward the porphyrin plane with a resultant change in the structure of the protein. In this way the oxygenation of one heme

unit leads to effects at the other heme units, despite the fact that the Fe atoms in the hemes are separated by more than 25 Å.

In the past few years it has been shown by a number of workers that it is possible to prepare and study complexes of the type BCoL.O<sub>2</sub>, where B is a base and where L is a Schiff base (Crumbly & Basolo, 1970; Hoffman, Diemente & Basolo, 1970) or a porphyrin (Misono, Koda & Uchida, 1969; Walker, 1970; Stynes & Ibers, 1972*a*). Following the initial studies in this area Hoffman & Petering (1970) reconstituted hemoglobin with a cobalt porphyrin. Hoffman, Spilburg & Petering (1971) showed that this material, which they called 'coboglobin', exhibits cooperative, allosteric oxygen binding. This has since been confirmed (Hsu, Spilburg, Bull & Hoffman, 1972). As Hoffman, Spilburg & Petering (1971) note, the oxygenation of coboglobin involves a change from low-spin Co(II) to low-spin Co(III). From stereochemical considerations they conclude that the cobalt ligation in coboglobin should cause a substantially smaller, perhaps negligible, motion of the proximal histidine residue relative to the mean plane of the porphyrin molecule than is postulated for hemoglobin. Perutz & TenEyck (1971) have discussed the implications of cooperative oxygen binding in coboglobin on the trigger mechanism for hemoglobin.

Here, on the basis of studies of model systems, we present a more quantitative picture of the stereochemistry of cooperative effects in the prosthetic group of

\* *Editorial note:* – This paper was presented at a memorial symposium in honour of Walter C. Hamilton, a former Co-editor of *Acta Crystallographica*, held on 15 June 1973 at Brookhaven National Laboratory, and sponsored jointly by the American Crystallographic Association and the Brookhaven National Laboratory.

† We choose the mean plane of the porphyrin ring (excluding substituents) as the reference point for a description of these relative motions.