Analysing Friedel averages and differences
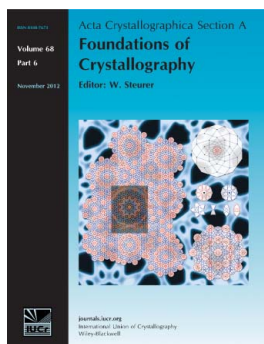
Simon Parsons, Phillip Pattison and Howard D. Flack

Acta Crystallographica Section A: Foundations of Crystallography covers theoretical and fundamental aspects of the structure of matter. The journal is the prime forum for research in diffraction physics and the theory of crystallographic structure determination by diffraction methods using X-rays, neutrons and electrons. The structures include periodic and aperiodic crystals, and non-periodic disordered materials, and the corresponding Bragg, satellite and diffuse scattering, thermal motion and symmetry aspects. Spatial resolutions range from the subatomic domain in charge-density studies to nanodimensional imperfections such as dislocations and twin walls. The chemistry encompasses metals, alloys, and inorganic, organic and biological materials. Structure prediction and properties such as the theory of phase transformations are also covered.

Crystallography Journals Online is available from **journals.iucr.org**

Analysing Friedel averages and differences

Simon Parsons,[a]* Phillip Pattison[b,c] and Howard D. Flack[d]

[a]School of Chemistry and Centre for Science at Extreme Conditions, University of Edinburgh, Kings Buildings, West Mains Road, Edinburgh EH9 3JJ, Scotland, [b]Laboratoire de Cristallographie, Ecole Polytechnique Fédérale de Lausanne, BSP, Lausanne, CH-1015, Switzerland, [c]Swiss–Norwegian Beamline, ESRF, BP 220, Grenoble, F-38043 Cedex 9, France, and [d]Chimie minérale, analytique et appliquée, University of Geneva, Geneva, Switzerland. Correspondence e-mail: s.parsons@ed.ac.uk

Various practical applications of the average (A) and difference (D) of Friedel opposites are described. Techniques based on the resonant-scattering contribution to Friedel differences are applied to see whether a crystal is centrosymmetric or not, and to determine the point group of the crystal. For the validation of a structural study, plots of A_{obs} against A_{model}, and D_{obs} against D_{model} are used extensively. Moreover, it is useful to display both plots on the same graph. Intensity measurements on a crystal of $NaClO_3$ were made at three different speeds, with two different radiations and two different diffractometers, and treated with two different software packages and four different absorption corrections. The evaluation of these numerous data sets reveals underlying deficiencies. For comparison, plots of A_{obs} against A_{model}, and D_{obs} against D_{model} are presented for two centrosymmetric crystals.

1. Introduction

In small-molecule crystallography it has been customary in crystal structure analysis to make no use of the contribution of resonant scattering other than in the specific area of absolute-structure and absolute-configuration determination. One may trace the causes of this situation to the weakness of the resonant-scattering contribution, to the high cost in time and labour of collecting intensity data sets containing measurements of all Friedel opposites and to the lack of any perceived or real need for the additional information that might be obtained from the effects of resonant scattering. A natural consequence of this situation is that reputable and authoritative texts on crystallography, introductory or advanced, contain no or very fragmentary information on this topic.

On the experimental side the turning point came with the widespread distribution of area detectors for small-molecule crystallography, giving the potential to measure, at no extra cost, the full-sphere data sets leading to the intensity differences between Friedel opposites, hkl and $\overline{hkl}$. Further impetuses for development were both the realization that standard uncertainties on the Flack parameter (Flack, 1983) were unreliable and the ever-increasing need for improvements in structure validation.

The validation of a structure analysis of a non-centrosymmetric crystal structure by way of plots of observed against model values of the average (A) and difference (D) intensities of Friedel opposites has been introduced by Flack *et al.* (2011). The average and difference of Friedel opposites are defined by

$$A(hkl) = \tfrac{1}{2}[|F(hkl)|^2 + |F(\overline{hkl})|^2], \tag{1a}$$

$$D(hkl) = |F(hkl)|^2 - |F(\overline{hkl})|^2. \tag{1b}$$

In general $D(hkl)$ is small compared to $A(hkl)$. A compound with an appreciable resonant-scattering contribution has $D(hkl)$ approximately $0.01A(hkl)$, whereas a compound with a small resonant-scattering contribution has $D(hkl)$ approximately $0.0001A(hkl)$. Flack *et al.* (2011) made a study of 29 crystal structures published in 2007. It was found that these crystal structure determinations could be separated into three categories based on the appearance of the D_{obs} against D_{model} plot of the acentric reflections. In the first category, these plots had the data points arranged around a straight line of slope 1 passing through the origin, and the conventional R values calculated on D were the lowest, ranging from 40 to 70%. It was deduced for structure determinations in this first category that the resonant-scattering contribution to the observed Friedel differences was significant, and that random uncertainties and systematic errors were minor in the intensity data. In the third category, the plots of D_{obs} against D_{model} had the data points arranged about the D_{obs} axis (where $D_{model} = 0$), the range of values of $|D_{obs}|$ was larger, even much larger, than the range of values of $|D_{model}|$, and the R values on D were high with values between 90 and 100%. Flack *et al.* (2011) interpreted the results of the third category as indicating that there was no problem with the D_{model} values but that the D_{obs} were entirely dominated by random uncertainties and systematic errors which combined to obscure, almost entirely,

the resonant-scattering contribution to the difference intensity of Friedel opposites. In the second category, the results were intermediate.

Various procedures are now described, arranged in the sequence in which they would be of use in a structure analysis. §2 deals with a novel method for the determination of the status of centrosymmetry of a crystal in which the contribution of resonant scattering is significant. §3 gives a detailed example of the extension of the R_{merge} technique to distinguish between the possible point groups in a chosen Laue class. In §4 we deal with the validation of the intensity data after structure refinement. This is carried out on our own carefully designed measurements, data correction and refinements on a single crystal of a model compound, $NaClO_3$. It is shown that these new data confirm and amplify the conclusions of Flack et al. (2011) and bring the analysis full circle. The problem in absolute-structure determination is shown to lie in the data measurements and corrections. The paper concludes in §5 with the presentation of the analyses of the intensity data from two centrosymmetric crystals.

2. Status of centrosymmetry and resonant scattering

We make use of the average and difference intensities of Friedel opposites given in equations (1a) and (1b). It is necessary to recall a few basic facts concerning these quantities. A data set of intensities needs to contain both reflections hkl and $\overline{hkl}$ in order to obtain the observed values of $A_{obs}(hkl)$ and $D_{obs}(hkl)$. $A(hkl)$ is a centrosymmetric function as $A(hkl) = A(\overline{hkl})$. On the other hand, $D(hkl)$ is antisymmetric as $D(hkl) = -D(\overline{hkl})$. In the model of a centrosymmetric crystal structure, $|F_{model}(hkl)|^2 = |F_{model}(\overline{hkl})|^2$ and consequently $D_{model}(hkl) = 0$ in this case. The values of $D_{obs}(hkl)$ of a centric reflection are entirely due to random uncertainties and systematic errors in the intensity measurements. On the other hand, the set of $D_{obs}(hkl)$ of acentric reflections contains contributions both from the random uncertainties and the systematic errors of the data measurements, and from the differences between $|F(hkl)|^2$ and $|F(\overline{hkl})|^2$ which arise through the effect of resonant scattering. Consequently, an appropriate study of the set of $D_{obs}(hkl)$ of potentially acentric reflections of a crystal structure of unknown space group may supply useful indications on whether the structure is centrosymmetric or non-centrosymmetric.

The Bijvoet ratio, defined by $\chi = \langle D^2 \rangle^{1/2}/\langle A \rangle$, is the ratio of the root-mean-square value of D to the mean value of A. At the outset of a structure analysis, two independent estimates of the Bijvoet ratio are available and their comparison leads to useful information as to whether the crystal structure is centrosymmetric or not.

The first estimate arises from considerations of intensity statistics through the analysis of an ensemble of random structures leading to the definition of the Bijvoet ratio as a value called Friedif_stat, whose functional form was derived by Flack & Shmueli (2007) and Shmueli & Flack (2009). One needs only to know the chemical composition of the compound and the wavelength of the X-radiation used in

order to calculate Friedif_stat. Flack & Bernardinelli (2008) provide a spreadsheet application for its evaluation.

The second estimate of the Bijvoet ratio, Friedif_obs, is obtained from the observed diffraction intensities. One tricky point in the evaluation of Friedif_obs is the different variation of A and D with $\sin(\theta)/\lambda$. One must thus work with suitably normalized values. The normalization of the A_{obs} values proceeds in the normal way by determining the values of the overall scale factor and isotropic displacement parameter from a Wilson plot of $\log(A_{obs}/\langle A \rangle)$ against $[\sin(\theta)/\lambda]^2$ and applying these to obtain normalized values of A_{obs} and $\langle A_{obs} \rangle$. D_{obs} values are normalized in the same way using the same values of the overall scale factor and isotropic atomic displacement parameter. However, for the D_{obs} a further adjustment is made to allow for the $\sin(\theta)/\lambda$ variation of Friedif_stat. To aid in this procedure, a further spreadsheet application, available as supplementary material to Flack et al. (2011), calculates Friedif_stat at various values of $\sin(\theta)/\lambda$ in the range 0.0 to 0.7 Å^{-1}. Expressing this variation as a polynomial, Friedif_stat $= c_0 + c_1\sin(\theta)/\lambda + c_2[\sin(\theta)/\lambda]^2$, the preliminary normalized D_{obs} are modified to become $D_{obs}/\{1 + (c_1/c_0)\sin(\theta)/\lambda + (c_2/c_0)[\sin(\theta)/\lambda]^2\}$. From these further-normalized D_{obs} values, $\langle D_{obs}^2 \rangle$ can be obtained to give Friedif_obs. A second tricky point in the calculation is to make sure that only acentric reflections of any of the non-centrosymmetric point groups in the chosen Laue class are selected for the calculation of Friedif_obs. In this way one is sure that if the point group of the crystal is centrosymmetric, all of the chosen reflections are centric, and if the point group of the crystal is non-centrosymmetric, all of the chosen reflections are acentric. The necessary selection is achieved by taking only those reflections that are general in the Laue group and these are indicated in Table 1. It is easy to prove that these will necessarily be acentric in any index-2 non-centrosymmetric subgroup of the Laue group. At the present time, the calculation of Friedif_obs is not available in distributed software. On comparison of Friedif_stat with Friedif_obs one is able to state with some confidence that:

(i) If Friedif_obs is much lower than Friedif_stat, then the crystal structure is either centrosymmetric, or non-centrosymmetric with the crystal twinned by inversion in a proportion close to 50:50, and random uncertainties and systematic errors in the intensity data set are minor.

(ii) If Friedif_obs is close in value to Friedif_stat, then the crystal is non-centrosymmetric, not twinned by inversion, and random uncertainties and systematic errors in the intensity data set are minor. However, data from a centrosymmetric crystal with large random uncertainties and systematic errors may also produce this result.

(iii) If Friedif_obs is much larger than Friedif_stat, then either the data set is dominated by random uncertainties and systematic errors or the chemical formula is erroneous.

Example 1. Measurements were made on the Swiss–Norwegian Beamline (BM01A) at the European Synchrotron Radiation Facility in Grenoble, France, on the compound 1-methyl-4-oxotetrahydro-2H-imidazol-2-iminium tetrachlorocopper(II) (Udupa & Krebs, 1979) [Cambridge Struc-
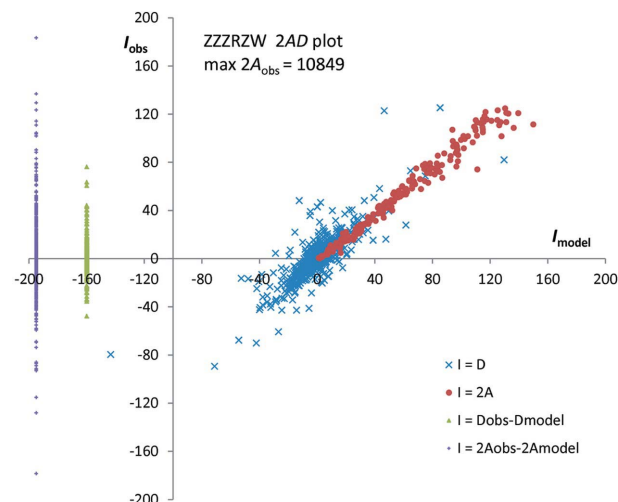
Table 1
Classification of reflections in centrosymmetric point groups.

All reflections are centric and all have a value of zero for the root-mean-square D. H.A. signifies hexagonal axes and R.A. rhombohedral axes.

Point group	h	ε, order of stabilizer	General (g) or special (s)
$\bar{1}$	hkl	1	g
$2/m$	hkl	1	g
	$h0l, 0k0$	2	s
mmm	hkl	1	g
	$0kl, h0l, hk0$	2	s
	$h00, 0k0, 00l$	4	s
$4/m$	hkl	1	g
	$hk0$	2	s
	$00l$	4	s
$4/mmm$	hkl	1	g
	$hk0, h0l, hhl$	2	s
	$hh0, h00$	4	s
	$00l$	8	s
$\bar{3}$ (H.A)	hkl	1	g
	$00l$	3	s
$\bar{3}$ (R.A)	hkl	1	g
	hhh	3	s
$\bar{3}m1$ (H.A.)	hkl	1	g
	$hh0, h0l$	2	s
	$00l$	6	s
$\bar{3}1m$ (H.A.)	hkl	1	g
	$h00, hhl$	2	s
	$00l$	6	s
$\bar{3}m$ (R.A.)	hkl	1	g
	$\bar{h}h0, hhl$	2	s
	hhh	6	s
$6/m$	hkl	1	g
	$hk0$	2	s
	$00l$	6	s
$6/mmm$	hkl	1	g
	$hk0, h0l, hhl$	2	s
	$h00, hh0$	4	s
	$00l$	12	s
$m\bar{3}$	hkl	1	g
	$0kl, hh0$	2	s
	hhh	3	s
	$00l$	4	s
$m\bar{3}m$	hkl	1	g
	$hhl, 0kl$	2	s
	$hh0$	4	s
	hhh	6	s
	$h00$	8	s

Figure 1
D_{obs} against D_{model} of all Friedel pairs with $2A_{obs}$ against $2A_{model}$ for weak Friedel pairs for ZZZRZW. On the left of the plot, $D_{obs} - D_{model}$ and $2A_{obs} - 2A_{model}$ of all Friedel pairs are displayed at constant abscissa.

tural Database (CSD; Allen, 2002) refcode: CRINCC] at 100 K with a radiation of wavelength 0.7000 Å. The crystal is known to be centrosymmetric (space group $P2_1/c$) and has a significant resonant-scattering contribution, Friedif$_{stat}$ = 498. The intensity data were merged and averaged in point group 1 giving 5372 measurements. The general reflections were separated from the special reflections ($h0l$ and $0k0$) for Laue group $2/m$ and collected into sets of reflections equivalent in $2/m$. This led to 724 sets (2896 reflections) for which all four $2/m$-symmetry-equivalent measurements were available. Normalized values of A_{obs} and D_{obs} were calculated for these 724 sets, from which $\langle A_{obs}\rangle$ and $\langle D_{obs}^2\rangle^{1/2}$ could be obtained to give Friedif$_{obs}$ = 164. The comparison of 498 for Friedif$_{stat}$ with 164 for Friedif$_{obs}$ indicates that the crystal structure is centrosymmetric. Other selected statistics on CRINCC are given in §3.

Example 2. Measurements and analysis of the compound potassium hydrogen (2R,3R) tartrate (CSD refcode: ZZZRZW) are described in §3. One finds there the values of $\langle A_{normalized}\rangle$ and $\langle D_{normalized}^2\rangle^{1/2}$ from the measurements of those sets of general reflections in the Laue group mmm which contain all eight mmm-symmetry-equivalent reflections. The value of Friedif$_{obs}$ is 217 compared to a Friedif$_{stat}$ value of 174. The agreement is good and allows the deduction that the crystal is neither centrosymmetric nor twinned by inversion in a proportion near to 50:50, nor that the data set is unsatisfactorily dominated by random uncertainty and systematic error. Clearly the crystal is non-centrosymmetric as justified by the results of the structure analysis and the excellent agreement between D_{obs} and D_{model} after refinement, as shown in Fig. 1. $2AD$ plots are described in detail in §4.2.2.

Example 3. Details of the relevant intensity measurements, structure refinement and data analysis for 1,3,4,6-tetra-O-acetyl-2-(trifluoromethylsulfonyl)-D-mannopyranose (Zhu & Jiang, 2007) (CSD refcode: UNEVAK01) are given in Flack & Bernardinelli (2008). The Laue group is $\bar{1}$ so all reflections are general. One finds Friedif$_{stat}$ = 70 and Friedif$_{obs}$ = 499. The huge discrepancy between the two shows that the observed values of D are dominated by random uncertainty and systematic error. This result is confirmed by the D_{obs} against D_{model} plot obtained after least-squares refinement of the crystal structure. The plots resemble those of SEZPUJ presented in Fig. 2 of Flack et al. (2011).

Table 2
Laue classes with corresponding merging groups for the calculation of A_{obs} and D_{obs}.

Laue class	Point groups	Merging group	Index (Laue/merging)
$\bar{1}$	$1, \bar{1}$	1	2
$2/m$	$2, m, 2/m$	1	4
mmm	$222, 2mm, m2m, mm2, mmm$	1	8
$4/m$	$4, \bar{4}, 4/m$	2	4
$4/mmm$	$422, \bar{4}2m, \bar{4}m2, 4mm, 4/mmm$	2	8
$\bar{3}$	$3, \bar{3}$	3	2
$\bar{3}m$	$32, 3m, \bar{3}m$	3	4
$6/m$	$6, \bar{6}, 6/m$	3	4
$6/mmm$	$622, \bar{6}2m, \bar{6}m2, 6mm, 6/mmm$	3	8
$m\bar{3}$	$23, m\bar{3}$	23	2
$m\bar{3}m$	$432, \bar{4}3m, m\bar{3}m$	23	4

Table 3
For the general reflections hkl of mmm, the number of measurements on ZZZRZW in each set and the corresponding number of sets of reflections symmetry equivalent under point group mmm.

No. of measurements in set	No. of sets
8	589
7	27
6	55
5	62
4	41
3	29
2	30
1	59
0	undetermined

3. Resolution of non-centrosymmetric ambiguities

It was shown in §2 that under certain circumstances it is possible to determine whether or not the space group of the crystal investigated is centrosymmetric. Suppose that the space group was found to be non-centrosymmetric. In each Laue class, there is one centrosymmetric point group and one or more non-centrosymmetric point groups. For example, in Laue class $2/m$, such point groups are 2 and m, and in Laue class mmm we need to distinguish between the point groups 222, $2mm$, $m2m$ and $mm2$, and of course between the space groups based on them. We shall show that it is possible in practice to distinguish between these non-centrosymmetric point groups using intensity differences between Friedel opposites caused by resonant scattering. The first technique, which works well in our test example, is a generalization of the use of conventional merging R factors applied to all point groups within a chosen Laue class. The second technique relies on intensity enhancement within specific zones or lines of reflections, and turns out not to be entirely satisfactory. The techniques are demonstrated by a practical example.

3.1. Practical example on ZZZRZW

Intensity measurements on a crystal of ZZZRZW were made on the Swiss–Norwegian Beamline (BM01A) at the European Synchrotron Radiation Facility in Grenoble, France. A wavelength of 0.7469 Å was used at 100 K. The crystal structure is well established and occurs in space group $P2_12_12_1$. Least-squares refinement on the data, merged and averaged in point group 222, displays conventional R factors $R_A = 3.1\%$, $R_D = 51.1\%$ and $R_{Aweak} = 10.4\%$ (*i.e.* for those reflections with $A_{obs} < |D_{obs}|_{max}$) (see Flack *et al.*, 2011), and the D_{obs} against D_{model} plot in Fig. 1 shows the data to be distributed about the straight line $D_{obs} = D_{model}$. The data are thus of very good quality with a clear signal from the resonant scattering in the Friedel opposites, Friedif$_{stat}$ = 174. There were 20 679 intensity measurements in the raw data with one reflection having been measured as many as nine times. The

experimental data are available as supplementary material (ktar.hkl).[1] The Laue group was assumed to be mmm.

3.1.1. R_{merge} on ZZZRZW. The steps in the data treatment are as follows:

(i) All reflections with identical reflection indices were collected together and averaged. This is a sorting and merging of reflections under point group 1 and makes no assumption at all as to the real point group of the intensity measurements. [For a crystal in Laue class mmm, the reflections were sorted and averaged in point group 1. In fact, the suitable merging group for reflections in any given Laue class is the highest subgroup common to all point groups in the Laue class. This merging group may be found by study of the figure of maximal subgroups of the three-dimensional point groups (Hahn & Klapper, 2002) and these are indicated in Table 2.]

(ii) The reflections were separated into seven classes corresponding to the general and special reflections of the Laue group mmm as indicated in Table 1. The special reflections $h00$, $0k0$ and $00l$ have a multiplicity of 2, the special reflections $0kl$, $h0l$ and $hk0$ have a multiplicity of 4, and the general reflections hkl have a multiplicity of 8. Reflections that are general reflections in point group mmm will necessarily be general reflections in all of its non-centrosymmetric subgroups (*i.e.* $2mm$, $m2m$, $mm2$, 222) within the same Laue class (mmm). The same technique was used for selecting reflections in §2.

(iii) Intensity measurements were gathered together into sets of reflections equivalent under the Laue group mmm. The count of reflections in each set and the corresponding number of sets are shown in Table 3 for the general reflections hkl in mmm. It is somewhat surprising that, despite the large redundancy of the raw data, there is nevertheless a considerable proportion of data for which the full complement of eight reflections had not been measured.

(iv) Restricting the analysis to the 589 sets containing eight general reflections (hkl, $\overline{hkl}$, $h\overline{kl}$, $\overline{h}kl$, $\overline{hk}l$, $h\overline{k}l$, $\overline{hk}\overline{l}$, $hk\overline{l}$), merging R factors were calculated on the $|F_{obs}|^2$ in the usual way. The values, given throughout in percentage units, are reported in the first line of Table 4. One sees that the true point group of the crystal 222 has a significantly lower value of $R_{|F|^2}$ than the other point groups. Next, the $|F_{obs}|^2$ of inversion-related reflections (*i.e.* hkl and $\overline{hkl}$; $h\overline{kl}$ and $\overline{h}kl$; $\overline{hkl}$ and $h\overline{kl}$;

Table 4
R merging values (%) for the 589 sets of general reflections of *mmm* which have all eight measurements on ZZZRZW in the set.

R_{merge}	*mmm*	*2mm*	*m2m*	*mm2*	222
$R_{|F|^2}$	2.42	2.31	2.29	2.31	1.80
R_A	1.30	1.30	1.30	1.30	1.30
R_D	100.0	254.4	235.7	258.1	82.9

Table 5
R merging values (%) for the 27 sets of general reflections of *mmm* which have seven measurements on ZZZRZW in the set, and the 55 and 27 sets that have six and seven measurements, respectively, in the set.

R_{merge}	*mmm*	*2mm*	*m2m*	*mm2*	222	No. of measurements
$R_{|F|^2}$	3.67	3.78	3.16	3.44	2.96	7 measured in set
$R_{|F|^2}$	3.90	3.73	3.39	3.36	3.19	6 and 7 measured in set

Table 6
R merging values (%) on $|F|^2$, A and D from ZZZRZW for special reflections $0kl$, $h0l$ and $hk0$ of *mmm* containing a full complement of four reflections.

	R_{merge}	*mmm*	*2mm*	*m2m*	*mm2*	222
$0kl$	$R_{|F|^2}$	1.85	1.85	1.25	1.55	1.85
$h0l$	$R_{|F|^2}$	1.20	1.01	1.20	0.98	1.20
$hk0$	$R_{|F|^2}$	1.52	1.29	1.18	1.52	1.52
$0kl$	R_A	0.68	0.68	0.68	0.68	0.68
$h0l$	R_A	0.72	0.72	0.72	0.72	0.72
$hk0$	R_A	0.75	0.75	0.75	0.75	0.75
$0kl$	R_D	100.0	100.0	72.8	137.3	100.0
$h0l$	R_D	100.0	107.4	100.0	93.1	100.0
$hk0$	R_D	100.0	115.8	86.4	100.0	100.0

Table 7
$\langle A_{normalized}\rangle$ and $\langle D^2_{normalized}\rangle^{1/2}$ for all sets of reflections of ZZZRZW with a full complement of measurements.

	No. of sets of reflections	$\langle A_{obs\ normalized}\rangle$	$\langle D^2_{obs\ normalized}\rangle^{1/2}$
hkl	589	1.73	0.038
$0kl$	75	1.59	0.027
$h0l$	52	1.43	0.019
$hk0$	129	1.70	0.021
$h00$	5	1.40	0.008
$0k0$	7	3.87	0.039
$00l$	1	0.15	0.009

$\overline{hkl}$ and $hk\bar{l}$) were converted into their average (A) and difference (D) values. The A and D values were then averaged separately under the five point groups of Table 4 using the relationships $A(hkl) = A(\overline{hkl})$ and $D(hkl) = -D(\overline{hkl})$. The merging R_A and R_D values are reported in lines 2 and 3 of Table 4. It should be noted that the merging R_D value in a centrosymmetric point group is 100%, not by coincidence, but by definition. Once again R_D merge of the true point group has the lowest value, which is now noticeably different from the other choices of point group. The R merging on D appears as a more sensitive value than the merging $R_{|F|^2}$.

(v) Similar calculations to those described in (iv) were performed on sets of general reflections of *mmm* which had less than the full complement of eight measurements. In this case it is not possible to undertake the calculation of the R_{merge} on A and D as these necessarily require all eight *mmm*-symmetry-equivalent reflections to be present to obtain the values of A and D. However it is possible to calculate the R_{merge} on $|F_{obs}|^2$. The results are given in Table 5. From a comparison of Tables 4 and 5 it is clear that the R merging values are more reliable with sets containing a full complement of eight reflections. Moreover, the calculation using A and D is preferable to that on $|F_{obs}|^2$.

(vi) Similar calculations to (iv) were performed on the *mmm* special reflections $0kl$, $h0l$ and $hk0$ where a full complement of four reflections had been measured in each set. There were 75 $0kl$, 52 $h0l$ and 129 $hk0$ reflections. The results are presented in Table 6. It is very clear from Table 6 that merging R factors either on $|F|^2$ or D from these special reflections do not provide reliable indications of the point group of the crystal.

3.1.2. Intensity enhancement analysis on ZZZRZW. This method is based on analytical expressions for $\langle D^2\rangle$, derived for an ensemble of random structures, and the classification of reflections for all non-centrosymmetric point groups given in Table 1 of Shmueli & Flack (2009). Special zones and lines of reflections are there shown to have an intensity enhancement which is characteristic of the point group of the crystal. In preparation for the applications to real data, the classification

presented in Table 1 of Shmueli & Flack (2009) has been supplemented to include the centrosymmetric point groups that are presented in Table 1.

To apply this technique one calculates $\langle A\rangle$ and $\langle D^2\rangle^{1/2}$ over various sets of data. Now, as the values of A and D vary with $\sin(\theta)/\lambda$, it is the normalized values of A and D that have to be prepared and averaged. The normalization technique that we used is exactly the same as that described in §2. Also, as the calculation requires values of A and D, one only uses those sets of reflections that have a full complement of measurements, *i.e.* eight for general reflections hkl, four for $0kl$, $h0l$ and $hk0$, two for $h00$, $0k0$ and $00l$. The results are given in Table 7.

The contents of Table 7 need to be consulted in conjunction with Table 1 of Shmueli & Flack (2009) and Table 1. The general reflections hkl give a baseline value of $\langle A\rangle$. The values of $\langle A\rangle$ for the reflection classes $0kl$, $h0l$ and $hk0$ clearly indicate that there is no intensity enhancement in these zones and preclude the point group of the crystal being *2mm*, *m2m* and *mm2* for which one zone has an enhancement factor of 2. For the classes $h00$, $0k0$ and $00l$ there are very few reflections indeed, making the results unreliable at the very best. It may just be a happy coincidence that $\langle A\rangle$ for $0k0$ is about twice the value for general reflections hkl as required for point group 222. Concerning the values of root-mean-square D, these do not give any clear indication. According to Table 1 of Shmueli & Flack (2009), r.m.s. D should have a value of zero for point group 222 but the $0kl$, $h0l$ and $hk0$ zones have approximately the value of the general hkl reflections.

Table 8
R merging values (%) for CRINCC for the 724 sets of general reflections of
$2/m$ which have all four measurements in the set.

R_{merge}	$2/m$	m	2
$R_{|F|^2}$	2.23	2.05	1.85
R_A	1.29	1.29	1.29
R_D	100.0	98.3	101.7

Table 9
$\langle A_{normalized} \rangle$ and $\langle D^2_{normalized} \rangle^{1/2}$ for all sets of reflections of CRINCC with a full complement of measurements.

	No. of sets of reflections	$\langle A_{obs\ normalized} \rangle$	$\langle D^2_{obs\ normalized} \rangle^{1/2}$
hkl	724	3.16	0.05
$h0l$	80	8.84	0.12

3.2. ZZZRZW and CRINCC

From all of the above information we note that it is the merging R factors on D, derived using only general reflections of mmm for which a full complement of eight reflections have been measured, that provide the clearest indication of the point group of the crystal. For applications to other Laue classes, Table 1 of Shmueli & Flack (2009) and Table 1 provide a complete list of general and special reflections.

For completeness, we report the principal statistics for the centrosymmetric structure CRINCC, Laue group $2/m$, already described as example 1 in §2. Table 8, similar to Table 4, gives the R merging values for the 724 sets of general reflections of $2/m$ which have all four measurements in the set. R_D values show no preference between the three point groups, whereas from $R_{|F|^2}$ one might be tempted to believe that the point group is 2. Table 9, similar to Table 7, gives normalized $\langle A_{obs} \rangle$ and $\langle D_{obs}^2 \rangle^{1/2}$ for sets of reflections with a full complement of measurements. Space-group-absent reflections have been omitted. No pairs of special non-space-group-absent reflections $0k0$ were available. The normalized $\langle A_{obs} \rangle$ is compatible with a symmetry-enhancement value of 2 for the $h0l$ reflections in agreement with Table 1.

4. Data validation demonstrated on NaClO₃

We here present the use of A and D in data validation following structure solution and least-squares refinement. A simple model compound, $NaClO_3$, has been used in this enterprise.

4.1. Intensity measurements and corrections on NaClO₃

Diffraction intensity measurements were made on one single crystal (dimensions $0.01 \times 0.16 \times 0.18$ mm) of $NaClO_3$ crystallized from aqueous solution. $NaClO_3$ crystallizes in space group $P2_13$ (No. 198) ($a \simeq 6.54$ Å), $Z = 4$ with both the Na and the Cl atoms in special positions $4a$ (x, x, x) on the threefold axis and an O atom in a general position $12b$. Measurements were made at 150 K. The value of $Friedif_{stat}$ (see, for example, Flack et al., 2011) is 114 for Mo $K\alpha$ and 492 for Cu $K\alpha$ radiation. The goniometers used allowed full orientational freedom of movement of the crystal and the intensities were recorded on area detectors. A full sphere of data was measured giving between 3260 and 3412 reflections with Mo $K\alpha$ radiation, and between 3581 and 4622 reflections with Cu $K\alpha$ radiation.

Using Mo $K\alpha$ radiation, three distinct data sets were measured on one instrument with different exposure times per frame: i.e. 60 s (labelled Mo$K\alpha$-slow), 10 s (labelled Mo$K\alpha$-medium) and 1 s (labelled Mo$K\alpha$-fast). Two different data-reduction software packages were used, labelled SftW1 and SftW2. These produce a list of integrated intensities from the raw frames and apply corrections for standard systematic effects such as Lorentz–polarization.

Using Cu $K\alpha$ radiation, two distinct data sets were measured on an instrument different from the one used for the Mo $K\alpha$ data sets. The set labelled Cu$K\alpha$-slow was measured over a period of 24 h with times per frame of 4 and 20 s for low- and high-angle detector positions. The set labelled Cu$K\alpha$-medium used 2 and 5 s per frame. These two data sets were processed only with the SftW1 data-reduction software package.

The five data sets of intensity measurements (Mo$K\alpha$-fast, Mo$K\alpha$-medium, Mo$K\alpha$-slow, Cu$K\alpha$-medium and Cu$K\alpha$-slow) were corrected for absorption following four different procedures. These are now described.

AbsNo: No absorption correction was applied.

AbsSe: A semi-empirical absorption correction was applied. A semi-empirical absorption correction uses redundancy in the intensity data to undertake its absorption correction. Symmetry-equivalent reflections are used to produce a correction which makes the corrected intensities as nearly equal as possible in a least-squares sense. The correction function is the sum of a series of spherical harmonic functions each with an adjustable coefficient, as described for example in Blessing (1995). The maximum order of spherical harmonics is generally a parameter chosen by the user. As the criterion of a semi-empirical correction is the near-equality of symmetry-equivalent reflections, the correction produced is not a pure absorption correction but rather a general-purpose correction. The origin of the disparity in intensity between symmetry-equivalent reflections is not identified. The point group used to identify equivalent reflections when applying this procedure to $NaClO_3$ was 23 for all data sets.

AbsGr: For SftW1, an absorption correction was applied using the method of Clark & Reid (1995), whereas for SftW2 the correction was obtained by numerical integration based on a Gaussian grid over the volume of the crystal. For these procedures it is necessary to index the crystal faces and measure their distance from the centre of the crystal.

AbsGrSe: An absorption correction as carried out for AbsGr was followed by a low-order semi-empirical absorption correction.

Table 10
Residuals for structure refinements on NaClO$_3$ for the five data sets (MoKα-fast, MoKα-medium, MoKα-slow, CuKα-medium and CuKα-slow) each treated with four different absorption-correction procedures (AbsNo, AbsSe, AbsGr and AbsGrSe) and two data-reduction software options (SftW1 and SftW2).

$R_A = \sum_{hkl}^{\text{paired acentric}} |A_{\text{obs}}(hkl) - A_{\text{model}}(hkl)| / \sum_{hkl}^{\text{paired acentric}} |A_{\text{obs}}(hkl)|$, $R_D = \sum_{hkl}^{\text{paired acentric}} |D_{\text{obs}}(hkl) - D_{\text{model}}(hkl)| / \sum_{hkl}^{\text{paired acentric}} |D_{\text{obs}}(hkl)|$. $R_{A\text{weak}}$ is an R value on A limited to reflections with $A_{\text{obs}} < |D_{\text{obs}}|_{\text{max}}$. R_{int} is an internal R factor with respect to point symmetry 23. $R_{|F|^2}$ is a conventional R factor on all data.

Data	Absorption correction	Software	R_A (%)	R_D (%)	$R_{A\text{weak}}$ (%)	$R_{D/A}$ (%)	R_{int} (%)	$R_{|F|^2}$ all (%)	Flack x
MoKα-fast	AbsNo	SftW1	2.3	90.9	10.5	90.0	7.79	2.03	0.09 (13)
MoKα-fast	AbsSe	SftW1	2.3	89.0	16.7	92.7	6.30	2.04	0.11 (12)
MoKα-fast	AbsGr	SftW1	2.3	88.5	9.1	89.5	7.63	2.00	0.11 (13)
MoKα-fast	AbsGrSe	SftW1	2.3	89.1	16.7	92.9	6.26	2.04	0.11 (12)
MoKα-med	AbsNo	SftW1	2.0	65.7	6.0	62.2	3.32	1.32	0.02 (8)
MoKα-med	AbsSe	SftW1	2.0	68.1	5.7	67.6	2.98	1.30	0.02 (8)
MoKα-med	AbsGr	SftW1	2.0	62.3	6.9	61.5	3.30	1.30	0.01 (8)
MoKα-med	AbsGrSe	SftW1	1.9	70.3	5.6	69.8	2.98	1.30	0.02 (8)
MoKα-slow	AbsNo	SftW1	2.3	59.6	7.8	52.9	2.91	1.31	0.05 (8)
MoKα-slow	AbsSe	SftW1	2.3	80.3	6.1	75.1	4.64	1.55	0.04 (9)
MoKα-slow	AbsGr	SftW1	2.3	63.4	6.7	54.1	2.97	1.31	0.05 (8)
MoKα-slow	AbsGrSe	SftW1	2.3	79.2	6.3	74.9	4.64	1.55	0.04 (9)
MoKα-fast	AbsNo	SftW2	2.9	96.4	11.4	96.0	16.26	2.52	0.13 (14)
MoKα-fast	AbsSe	SftW2	2.2	86.9	21.6	94.3	4.41	1.85	0.09 (12)
MoKα-fast	AbsGr	SftW2	2.9	96.5	6.2	96.2	15.95	2.51	0.15 (14)
MoKα-fast	AbsGrSe	SftW2	2.2	86.9	21.6	94.3	4.41	1.85	0.09 (12)
MoKα-med	AbsNo	SftW2	2.7	94.3	5.1	88.0	13.86	2.03	0.02 (11)
MoKα-med	AbsSe	SftW2	1.8	67.9	5.5	61.4	2.59	1.21	0.02 (8)
MoKα-med	AbsGr	SftW2	2.7	94.2	4.8	88.9	13.53	1.98	0.01 (11)
MoKα-med	AbsGrSe	SftW2	1.8	67.8	5.7	61.1	2.60	1.21	0.02 (8)
MoKα-slow	AbsNo	SftW2	3.6	95.5	4.6	92.1	14.76	2.28	0.01 (11)
MoKα-slow	AbsSe	SftW2	2.1	52.0	6.6	46.9	2.16	1.20	0.03 (8)
MoKα-slow	AbsGr	SftW2	3.1	94.9	4.1	91.5	14.41	2.12	0.01 (11)
MoKα-slow	AbsGrSe	SftW2	2.1	51.3	6.6	45.8	2.17	1.21	0.03 (8)
CuKα-med	AbsNo	SftW1	3.4	65.5	5.0	49.6	8.99	2.23	0.03 (3)
CuKα-med	AbsSe	SftW1	3.0	39.1	4.2	35.8	5.11	2.07	0.02 (3)
CuKα-med	AbsGr	SftW1	2.6	59.2	3.8	45.6	7.31	1.96	0.02 (3)
CuKα-med	AbsGrSe	SftW1	2.8	41.0	4.5	31.7	4.95	1.90	0.02 (3)
CuKα-slow	AbsNo	SftW1	3.7	82.0	3.9	49.0	10.76	2.60	0.02 (3)
CuKα-slow	AbsSe	SftW1	4.1	36.9	3.4	25.4	5.23	2.28	0.01 (3)
CuKα-slow	AbsGr	SftW1	3.2	69.3	4.1	39.5	7.89	2.30	0.01 (3)
CuKα-slow	AbsGrSe	SftW1	3.7	32.2	3.0	23.0	5.18	2.05	0.00 (2)

After each of these absorption corrections, the intensity data were merged and averaged in point group 23. Refinement of the 32 absorption-corrected data sets was carried out by linear(ized) least squares working on $|F|^2$. 18 parameters were refined: scale factor, Flack parameter (Flack, 1983), isotropic extinction parameter, x(Na), x(Cl), x(O), y(O), z(O), U^{11}(Na), U^{12}(Na), U^{11}(Cl), U^{12}(Cl), U^{11}(O), U^{22}(O), U^{33}(O), U^{12}(O), U^{23}(O), U^{13}(O). There were 247 or 248 (95 acentric pairs, 1 unpaired acentric and 56 or 57 centric) reflections in the Mo Kα, and 197 (74 acentric pairs, 1 or 0 unpaired acentric and 48 or 49 centric) reflections in the Cu Kα least-squares refinements.

Table 10 contains the R values and Flack parameter of the various refinements. File AFAD.NaClO3.allplots.pdf in the supplementary material contains the plots of A_{obs} against A_{model}, D_{obs} against D_{model} with $2A_{\text{obs}}$ against $2A_{\text{model}}$ in the same range, and $(D/A)_{\text{obs}}$ against $(D/A)_{\text{model}}$ for $A_{\text{obs}} > 0.01A_{\text{max}}$ of all of the refinements. The A_{obs} against A_{model} plots are on logarithmic axes whilst the others are on linear axes. In all cases the x and y axes are arranged to span the same domain of values. In this way, a plot with a satisfactory agreement between the observed and model values follows a straight line of slope 1 passing through the origin. Some of the

plots have been reproduced in the body of the paper by way of examples.

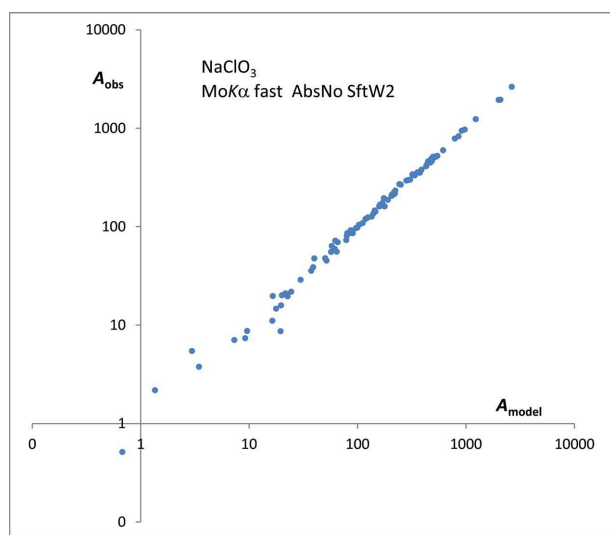### 4.2. Interpretation of NaClO$_3$ measurements

4.2.1. A_{obs} against A_{model} plots. These plots are on logarithmic axes, which entail both advantages and disadvantages. A distinct advantage is that the data may be presented compactly in one plot. However, the logarithmic axes have a tendency to exaggerate the spread of the data points at low A values and to compress the spread at high A values. One of the poorest plots is that of MoKα-fast-AbsNo-SftW2 shown in Fig. 2(a) and one of the best is that of CuKα-slow-AbsGr-SftW1 shown in Fig. 2(b). For MoKα-fast-AbsNo-SftW2 there is a slight spread of data points around the ideal line of slope 1 passing through the origin whereas CuKα-slow-AbsGr-SftW1 is almost perfect. In none of the plots is there any 'outlier' or 'rogue' reflection. Table 10 shows that the range of R_A values is between 1.8 and 4.1%. One may say that these plots reveal decently refined crystal structures. In accordance with expectation, the noisiest plots are those of data measured at high speed and the best ones have been measured at low speed. A notable feature is that data-reduction software SftW1 consis-

tently produces a positive intercept on the A_{obs} axis (indicating that the intensity of weak reflections is overestimated), largest for the fast data collection and smallest for the slow data collection. The plots for SftW2 do not show this effect and seem to pass very close to the origin of the plot. These effects are very clearly visible in the 2AD plots presented in §4.2.2. However, the R_A values produced by SftW1 are consistently lower than those of SftW2. The absorption corrections have worked well on the A data for which the software has most probably received the most tests and validation.

4.2.2. The 2AD plots. These plots contain all of the D_{obs}, D_{model} data points plotted on linear axes of identical length. The plots also contain $2A_{obs}$, $2A_{model}$ data points lying within the domain of values of the D_{obs}, D_{model} data. On the left of each plot, $D_{obs} - D_{model}$ and $2A_{obs} - 2A_{model}$ of all Friedel pairs are displayed at constant abscissa. These show the spread of residual 2A and D obtained after refinement, and hence the

overall uncertainties and errors in the data. For each individual Friedel pair of reflections, 2A and D are, respectively, the sum and difference of $|F(hkl)|^2$ and $|F(\overline{hkl})|^2$ and, consequently, their standard uncertainties $u(2A)$ and $u(D)$ are identical. In the 2AD plots, the (weak) D_{obs}, D_{model} data are being compared with the weak $2A_{obs}$, $2A_{model}$ data. If the $2A_{obs}$, $2A_{model}$ plot has a different appearance to the D_{obs}, D_{model} plot, one is led to suspect that systematic errors are affecting the data. Table 10 shows R_D, R_A and $R_{A weak}$ values, the latter being R_A limited to those reflections with $A_{obs} < |D_{obs}|_{max}$ which appear in the 2AD plots.

One of the poorest cases is that of MoKα-fast-AbsNo-SftW2 shown in Fig. 3(a). The range of values in D_{obs} ($|D_{obs}|_{max}$ = 117) is much larger than in D_{model} ($|D_{model}|_{max}$ = 16). The arrangement of the D_{obs}, D_{model} data points is very different from that of the $2A_{obs}$, $2A_{model}$ points. It would appear that the
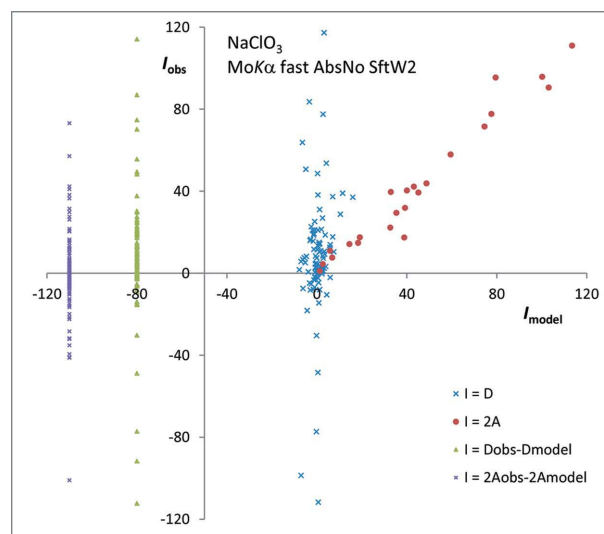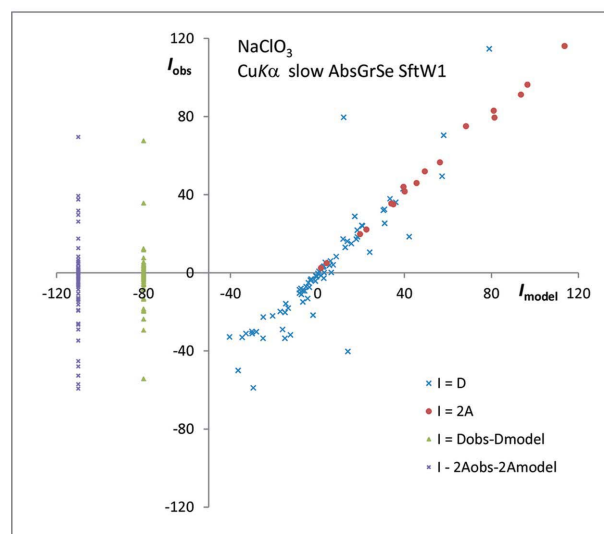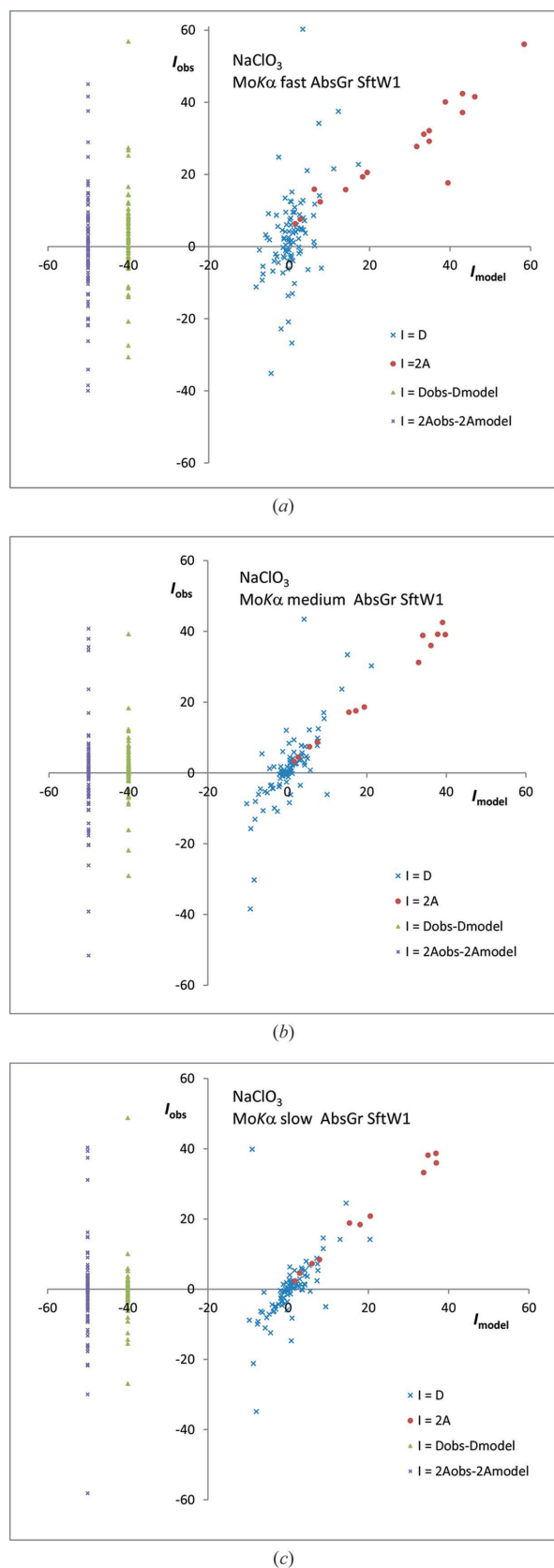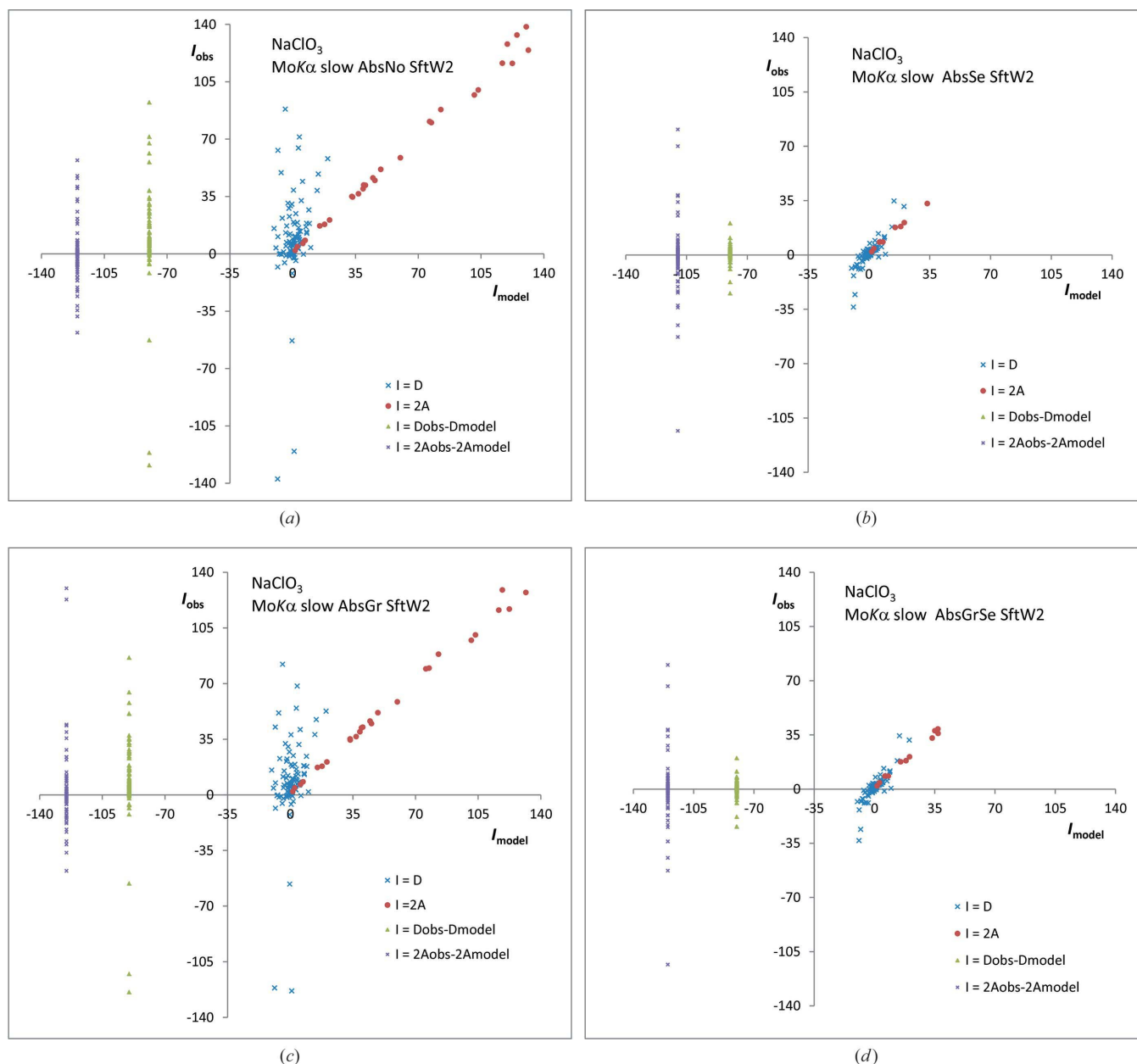

(a)

(b)

Figure 2
A_{obs} against A_{model} on logarithmic axes for MoKα-fast-AbsNo-SftW2 in (a) and CuKα-slow-AbsGr-SftW1 in (b).

(a)

(b)

Figure 3
D_{obs} against D_{model} of all Friedel pairs with $2A_{obs}$ against $2A_{model}$ for weak Friedel pairs for MoKα-fast-AbsNo-SftW2 in (a) and CuKα-slow-AbsGrSe-SftW1 in (b). On the left of the plot, $D_{obs} - D_{model}$ and $2A_{obs} - 2A_{model}$ of all Friedel pairs are displayed at constant abscissa.

Simon Parsons et al. · Analysing Friedel averages and differences **743**

(a)

(b)

(c)

Figure 4
D_{obs} against D_{model} of all Friedel pairs with $2A_{obs}$ against $2A_{model}$ for weak Friedel pairs for Mo$K\alpha$-XXX-AbsGr-SftW1. (a) XXX = fast, (b) XXX = medium and (c) XXX = slow. On the left of the plot, $D_{obs} - D_{model}$ and $2A_{obs} - 2A_{model}$ of all Friedel pairs are displayed at constant abscissa.

random uncertainties and systematic errors in the D_{obs} values are masking the resonant-scattering signal. Since this data set has been measured with a very short exposure time per frame, it is natural that the data are noisy and R_D is high at 96.4% (R_A = 2.9%).

One of the best cases is Cu$K\alpha$-slow-AbsGrSe-SftW1 shown in Fig. 3(b). The range of values in D_{obs} (−58 to 115) is similar to that in D_{model} (−40 to 79). Moreover, most of the D_{obs}, D_{model} data points are distributed about the ideal line of slope 1 passing through the origin. With R_D at 32.2% (R_A = 3.7%), the result is most satisfactory. It is apparent that the D_{obs} values are dominated by the resonant-scattering contribution with random uncertainties and systematic errors making only a small contribution.

Let us examine the effect of the speed of data collection by comparing plots in which the same method of absorption correction has been applied. We have chosen the series Mo$K\alpha$-fast-AbsGr-SftW1, Mo$K\alpha$-medium-AbsGr-SftW1 and Mo$K\alpha$-slow-AbsGr-SftW1 shown in Fig. 4. For Mo$K\alpha$-fast-AbsGr-SftW1, the D_{obs}, D_{model} data points show a bunch of values around the origin with a vague tendency to follow a line of slope greater than 1 passing through the origin, R_D is 88.5%. In Fig. 4(a) one D_{obs}, D_{model} data point at (−120.5, −7.6) has been omitted. The range of $|D_{obs}|$ is approximately seven times that of $|D_{model}|$. For Mo$K\alpha$-medium-AbsGr-SftW1, the D_{obs}, D_{model} data points show a more extended bunch around the origin but with a clear tendency to follow a line of slope 1 with R_D at 62.3%. The range of $|D_{obs}|$ is twice that of $|D_{model}|$. For Mo$K\alpha$-slow-AbsGr-SftW1, the arrangement of the D_{obs}, D_{model} data points is very clearly a straight line of slope 1 passing through the origin. There are a few outliers. Comparison with the other sets of three plots (fast, medium and slow) with identical absorption correction shows the same behaviour as the AbsGr set. This experiment makes it very clear that random uncertainties can very easily produce D_{obs} against D_{model} plots in which the data points are arranged around the D_{obs} axis and that the plot becomes more satisfactory if the intensities are measured with more care and more slowly.

The effect of an absorption correction can be seen in the series of four plots of the data collection at slow speed: Mo$K\alpha$-slow-AbsNo-SftW2, Mo$K\alpha$-slow-AbsSe-SftW2, Mo$K\alpha$-slow-AbsGr-SftW2 and Mo$K\alpha$-slow-AbsGrSe-SftW2 seen in Fig. 5. In AbsNo, even excluding some outliers, the data points are arranged around the D_{obs} axis and the range of $|D_{obs}|$ is much larger than that of $|D_{model}|$. An absorption correction by numerical integration (AbsGr) improves things very little. In the plot of the semi-empirical correction AbsSe, the D_{obs} and D_{model} values are now about on the same scale. AbsSe and AbsGrSe show straight lines of slope 1. For most data sets, one sees that the Se and GrSe corrections produce very similar D_{obs}, D_{model} plots, suggesting that the Gr part of the correction is not helping very much. Two explanations spring to mind. Maybe the indexing of the crystal faces and the distance measurements are inadequate or maybe the systematic error present in the Gr plots is not due to absorption.
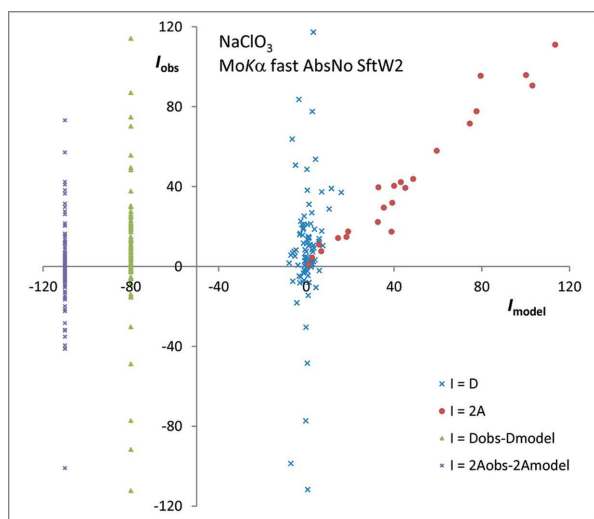
Figure 5
D_{obs} against D_{model} of all Friedel pairs with $2A_{obs}$ against $2A_{model}$ for weak Friedel pairs for MoKα-slow-YYY-SftW2. (a) YYY = AbsNo, (b) YYY = AbsSe, (c) YYY = AbsGr and (d) YYY = AbsGrSe. On the left of the plot, $D_{obs} - D_{model}$ and $2A_{obs} - 2A_{model}$ of all Friedel pairs are displayed at constant abscissa.

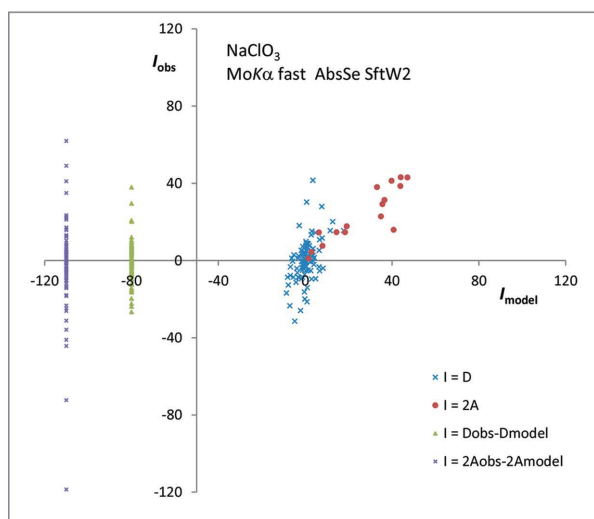There is also an effect of the speed of data collection on the quality of a semi-empirical absorption correction. Fig. 6 shows plots for MoKα-fast-AbsNo-SftW2, MoKα-fast-AbsSe-SftW2 and MoKα-medium-AbsSe-SftW2. For the fast data set there is a slight improvement on applying a semi-empirical correction, R_D changing from 96.4 to 86.9%, though the apparent positive gradient is actually the result of just a few data points. By contrast, with the exception of one or two outliers, the D_{obs}, D_{model} data for the medium data set follow the $2A_{obs}$, $2A_{model}$ data tolerably well, with $R_D = 67.9\%$. For the slow data set $R_D = 52.0\%$. The ability of the semi-empirical method to correct for systematic errors depends on the quality of the individual intensity measurements. If the latter are subject to high

random uncertainties because they have been collected rapidly, the resulting semi-empirical correction is inadequate. This shows that a data-collection strategy aimed at obtaining rapidly a data set with a very high redundancy is not appropriate for absolute-structure determination when using a semi-empirical absorption correction.
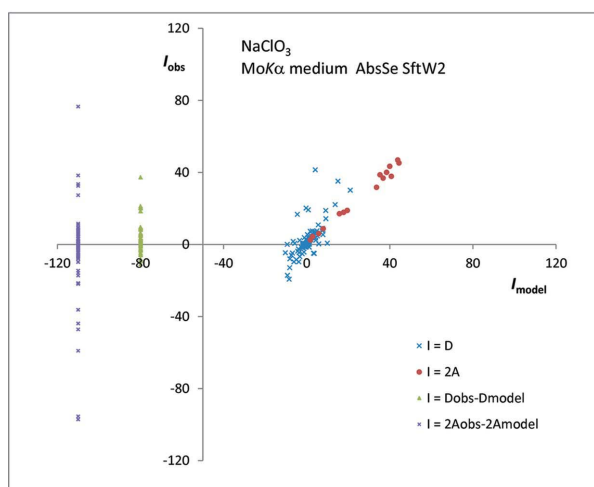
The effect of a change of data-reduction software can be seen by comparing MoKα-slow-AbsNo-SftW1 and MoKα-slow-AbsNo-SftW2 shown in Fig. 7. In both D_{obs} against D_{model} plots there are outliers, but those of SftW2 have considerably larger values of $|D_{obs}|$ than SftW1. Moreover, the D_{obs}, D_{model} data points of SftW1 follow reasonably closely a straight line of slope 1 passing through the origin, whereas
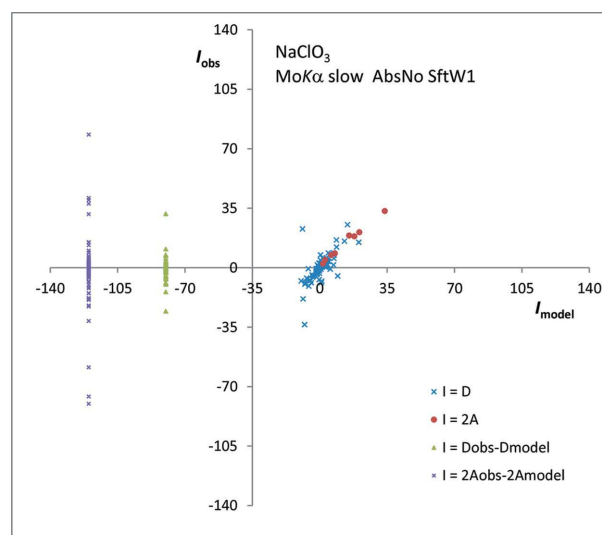
Simon Parsons *et al.* · Analysing Friedel averages and differences

(a)

(b)

(c)

Figure 6
D_{obs} against D_{model} of all Friedel pairs with $2A_{obs}$ against $2A_{model}$ for weak Friedel pairs for MoKα-fast-AbsNo-SftW2 in (a), MoKα-fast-AbsSe-SftW2 in (b) and MoKα-medium-AbsSe-SftW2 in (c). On the left of the plot, $D_{obs} - D_{model}$ and $2A_{obs} - 2A_{model}$ of all Friedel pairs are displayed at constant abscissa.
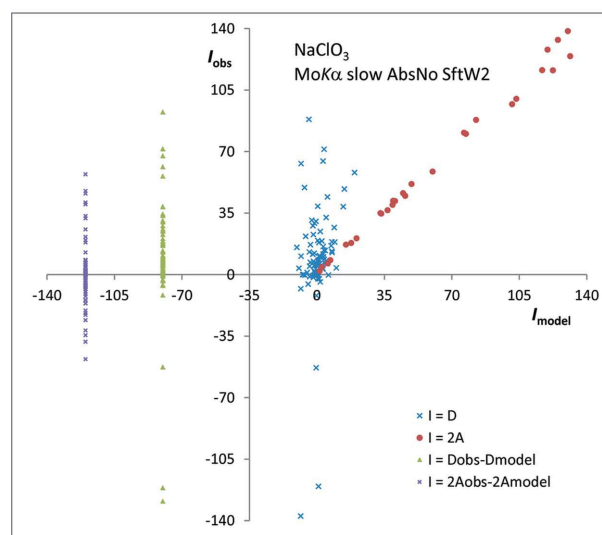
those of SftW2 are bunched around the D_{obs} axis. Simultaneously the $2A_{obs}$, $2A_{model}$ data points of SftW1 appear to be more noisy than those of SftW2.

4.2.3. $(D/A)_{obs}$ against $(D/A)_{model}$ plots. In an attempt to improve the quality of absolute-structure determination, Parsons & Flack (2004) and Parsons (2011) have developed and investigated the use of the ratio D/A as observables in a least-squares technique. The hypotheses underlying this approach are that: (i) there are significant systematic errors on all intensities under the restriction that the difference in these errors between reflections hkl and $\overline{hkl}$ is small, and (ii) random uncertainties in the intensity measurements are assumed to be small.

Indeed, in the past, using the now obsolete four-circle serial diffractometer, it was possible under certain conditions, see e.g. Le Page et al. (1990), to produce intensity measurements

(a)

(b)

Figure 7
D_{obs} against D_{model} of all Friedel pairs with $2A_{obs}$ against $2A_{model}$ for weak Friedel pairs for MoKα-slow-AbsNo-ZZZ. (a) ZZZ = SftW1 and (b) ZZZ = SftW2. On the left of the plot, $D_{obs} - D_{model}$ and $2A_{obs} - 2A_{model}$ of all Friedel pairs are displayed at constant abscissa.
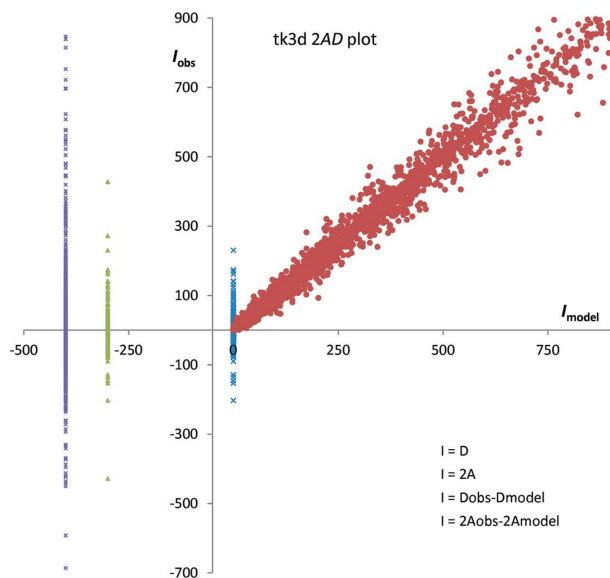
Figure 8
D_{obs} against D_{model} of all Friedel pairs with $2A_{obs}$ against $2A_{model}$ for weak Friedel pairs for tk3d. On the left of the plot, $D_{obs} - D_{model}$ and $2A_{obs} - 2A_{model}$ of all Friedel pairs are displayed at constant abscissa.

obeying the above conditions very closely. We may make a simple model encapsulating these hypotheses by writing

$$|F(hkl)|^2_{obs} = [s(hkl) + \Delta s(hkl)]|F(hkl)|^2_{model},$$
$$|F(\overline{hkl})|^2_{obs} = [s(hkl) - \Delta s(hkl)]|F(\overline{hkl})|^2_{model}. \quad (2)$$

In these, $s(hkl)$ and $\Delta s(hkl)$ are, respectively, the average and half-difference of the systematic errors to reflections hkl and $\overline{hkl}$. This leads to

$$(D/A)_{obs} \simeq (D/A)_{model} + \frac{2\Delta s(hkl)}{s(hkl)} - \left[\frac{\Delta s(hkl)}{s(hkl)}\right]^2 (D/A)_{model}$$
$$- \frac{\Delta s(hkl)}{2s(hkl)}(D/A)^2_{model}, \quad (3)$$

from which it may be seen that for small relative systematic errors, $\Delta s/s$, between reflections hkl and $\overline{hkl}$, $(D/A)_{obs} \simeq (D/A)_{model}$. The major part of the systematic errors, s, cancels out in this approach. In the supplementary material we also present plots of $(D/A)_{obs}$ against $(D/A)_{model}$ in order to be able to judge how well this approach applies to modern-day measurement techniques. So that the plots conform to the premises of this procedure, we eliminate all intensity measurements which have $A_{obs} < 0.01A_{max}$. These few eliminated reflections are thus weak. In Table 10 one can see that for all data sets other than MoKα-fast, the R values on (D/A) are smaller than the corresponding ones on D. Indeed, the slow data sets are superior to the medium ones, as are the Cu Kα ones to those using Mo Kα. Moreover, comparison of the plots of (D/A) to those on D shows that the former are cleaner and closer to the ideal plot of slope 1 passing through the origin. The ratio procedure is clearly providing some degree of correction of the systematic errors not undertaken by the

various absorption corrections we applied. The plots of the MoKα-fast refinements show that the ratio procedure does very little for data which are dominated by random uncertainties. In all, the D/A ratio is performing according to expectation, with the best performance being obtained for data sets with a small random uncertainty and appreciable resonant-scattering contribution.

5. 2AD plots from centrosymmetric crystal structures

In the course of the investigations described in §§2 to 4, it became clear that we had neglected to produce 2AD plots of any centrosymmetric crystal structure to serve for the purposes of comparison and reference. As a precaution, only reflections which are general in the point group and for which both the hkl and $\overline{hkl}$ reflections have been measured have been used to produce these plots. Technically, these are more difficult to produce, as the file of observed and model intensities from the final least-squares cycle contains reflections merged and averaged in the centrosymmetric point group of the model. Now, although the available $|F_{obs}|^2$ and $|F_{model}|^2$ values are identical to A_{obs} and A_{model}, respectively, and one knows that all D_{model} are zero, the D_{obs} values cannot be obtained from this file. They may, however, be obtained from a file of observed intensity data which have been merged and averaged either in point group 1 or in the merging group indicated in Table 2. In general, the latter intensity data are not on an absolute scale and a suitable scale factor is determined by comparison with the data in the file obtained at the completion of least-squares refinement. As the averaging procedures for producing the A_{obs} and the D_{obs} values are different, one needs to plot $(2N)^{1/2}A$ values to compare with D values. We have nevertheless continued to call this a 2AD plot. N is equal to the order of the point group of the crystal if the merging and averaging have been carried out in point group 1, and is the value of *index* given in Table 2 if the merging group of Table 2 has been used.

Intensity measurements (14 997 Bragg reflections) on a crystal of 2-(4-(di-*p*-tolylmethylene)cyclohexa-2,5-dien-1-ylidene)malononitrile (local code: tk3d) were made on the Swiss–Norwegian Beamline (BM01A) at the European Synchrotron Radiation Facility in Grenoble, France. A wavelength of 0.69830 Å was used at 100 K. The crystal structure occurs in space group $P2_1/c$, the compound has composition $C_{24}H_{18}N_2$ giving Friedif$_{stat}$ = 2. Least-squares refinement on the data, merged and averaged in point group $2/m$, displays conventional R factors $R_{|F|all} = 3.9\%$, $R_{|F|} = 3.7\%$, $R_A = 5.6\%$ and $R_{Aweak} = 6.3\%$ (*i.e.* for those reflections with $A_{obs} < |D_{obs}|_{max}$) (see Flack *et al.*, 2011). The data are thus of very good quality. The 2AD plot is in Fig. 8. Some systematic error is affecting the D data. (Three outliers have not been reproduced in Fig. 8.) Moreover, as with the NaClO$_3$ data treated with data-reduction software SftW1, the mean line of the $2A_{obs}$, $2A_{model}$ data has a positive intercept at $A_{model} = 0$. The 2AD plot of CRINCC, described in §§2 and 3, shows similar features, as seen in Fig. 9.
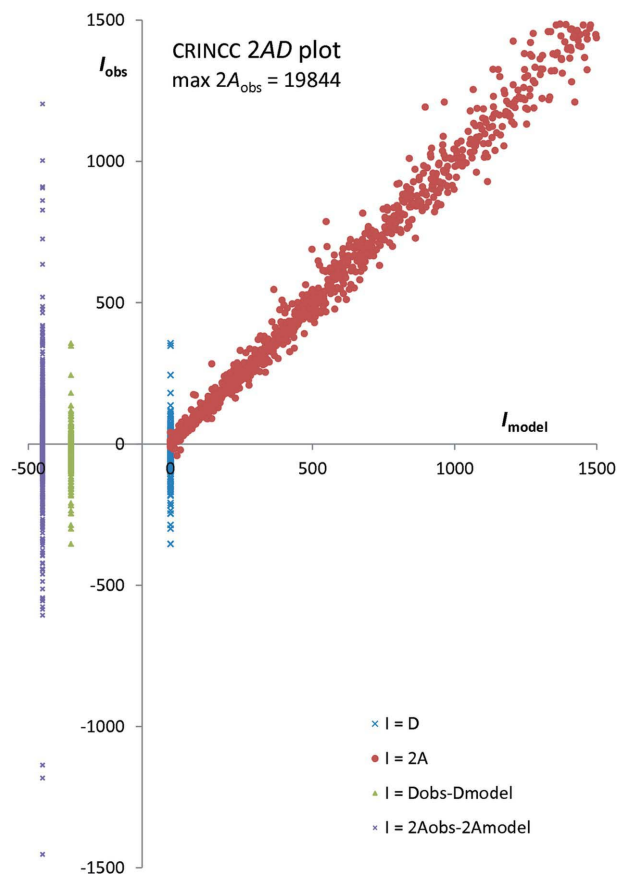
Figure 9
D_{obs} against D_{model} of all Friedel pairs with $2A_{obs}$ against $2A_{model}$ for weak Friedel pairs for CRINCC. On the left of the plot, $D_{obs} - D_{model}$ and $2A_{obs} - 2A_{model}$ of all Friedel pairs are displayed at constant abscissa.

6. Concluding remarks

The technique demonstrated in §2 shows that information on the status of centrosymmetry of a crystal may be obtained from the observed average and differences of the Friedel opposites without the need for a model of the crystal structure. The necessary condition for success is that the resonant-scattering contribution to the Friedel differences be clearly expressed in the observed diffraction data. Unfortunately this contribution is often masked by the random uncertainties and systematic errors in the diffraction data. Nevertheless, the technique of §2 has its place in the toolkit of the structure analyst.

The measurements on ZZZRZW described in §3 reveal a weakness in the data-collection strategy as required for the type of analysis described therein. One recalls that one reflection was measured as many as nine times. On the other hand, Table 3 shows that for a significant number of sets of reflections, symmetry equivalent under *mmm*, not all of the eight reflections in each set had been measured. This lack of measurements is to the detriment of the use of R_{merge} in determining the point group of the crystal. The measurements were performed with the crystal in one single orientation on its mount. This is the method that seems to be used universally.

Although it should be possible to collect a complete sphere of data on a diffractometer equipped with a four-circle goniometer, this can add substantially to data-collection times, and so is rarely achieved in practice. A more even coverage of the reflections in reciprocal space could doubtless be obtained by undertaking measurements with the crystal in more than one orientation on its mount. One would require to know what are optimal values for the number of different orientations of the crystal and their mutual angular offsets. We know of no publications dealing with such matters.

Following the introduction of the Wilson intensity statistics (Wilson, 1949), Rogers (1950) presented a detailed analysis of new methods for determining crystal classes and space groups. Rogers' method made use both of the established techniques of merging R and space-group absences, together with the symmetry enhancement of the intensity of zones and lines of special reflections. In practice, the latter proved to be of limited reliability and the method is not used in practice nowadays. Our analysis of symmetry enhancement of intensity in §3.1.2 confirms the unreliability of this method.

We have shown that the use of R_{merge} for determining the point group of a crystal is viable with a good data set for a crystal in point group 222. In the future we shall investigate the use of this technique to point groups with fewer symmetry-equivalent reflections in the Laue group. In particular, we think that it will be worthwhile to investigate crystals in the following point groups: 2, *m* and *mm*2.

Concerning the contents of §4, we have paid the greatest attention not to identify the producers of the instrumentation and software that have been used in the study on $NaClO_3$. The intention of the current paper is to reveal general problems of data collection and correction arising in absolute-structure determination and not to undertake a witch-hunt of instrumentation and software in their design, manufacture, implementation and documentation. We believe that the effects that we describe in this paper are general phenomena and not specific shortcomings of a particular instrument or software. Consequently, the instrumentation and software have not been identified.

It is our intention to investigate the behaviour of the normal probability plots of A and D as a means of providing further information for the validation of absolute-structure determinations.

The results presented in §5 on centrosymmetric crystals show that the $2A_{obs}$ against $2A_{model}$ plots show similar features to those of non-centrosymmetric crystals. Watkin (2011) has found similar results with other centrosymmetric crystal structures. Somewhat to our surprise, there appear to be no results on this matter in the extensive literature devoted to the experimental study of deformation electron densities, which has been carried out in the main on centrosymmetric crystals.

The techniques described in §4 provide an objective method for the validation of the observed contribution of resonant scattering to the measured diffraction intensities. This leads us to suggest the following protocol of alternative choices for the publication of a structure analysis on a non-centrosymmetric crystal structure. Choice (*a*): Average all Friedel opposites. A

$2AD$ plot, R_A, R_D and $R_{A\mathrm{weak}}$ values should not be reported. Make no report or comment on the Flack paramter (Flack, 1983), absolute structure or absolute configuration. Authors should not be required to justify this choice of procedure. Currently there are implementation problems with this procedure. Most single-crystal structure-factor least-squares software neither directly allows the calculation and use of A nor provides a way of simultaneously using data sets containing both A and $|F|^2$ values, the latter arising from reflections for which only one member of a Friedel pair was measured. Also, up to version 2.4.2, the CIF core dictionary (http://www.iucr.org/resources/cif/dictionaries/cif_core) has no data items for recording A. Choice (b): Do not average Friedel opposites. Provide a $2AD$ plot, R_A, R_D and $R_{A\mathrm{weak}}$ values which should attest to a satisfactory agreement between D_{obs} and D_{model}. Provide a value of the Flack parameter (Flack, 1983), and comment on absolute structure and absolute configuration as appropriate.

These choices provide the structure analyst with a free hand in dealing with the compound under study and justifying the results in accordance with the objectives of the study. A disturbing aspect of the analysis of Friedel averages and differences is that one may readily obtain an excellent fit on the averages whilst having a poor or nonexistent fit on the differences. One must not assume that a good fit of the averages implies a good fit of the differences.

For the determination of the absolute structure of a non-centrosymmetric crystal structure, our study shows clearly that one obtains the best results by measuring slowly and to a high redundancy. Systematic effects have to be corrected for with care and may easily hide the resonant-scattering effect in the Friedel opposites. The $2AD$ plots provide a very powerful method for the validation of the absolute-structure determination.

Some words of a more general nature are in order. Over the years, procedures have appeared in the literature which attempt to improve the quality of absolute-structure determination. These procedures rely on advanced statistical methodology resulting in the reduction of the uncertainty of the absolute-structure determination. Little or no account is made of intensity-measurement and data-correction methodology in the description of these improved techniques. The results on the $NaClO_3$ crystal point in exactly the opposite direction. It is rather the intensity measurement and correction which need to be studied and undertaken with improved precision to come to a more satisfactory result. For the purposes of absolute-structure determination, the model of a crystal twinned by inversion is physically acceptable and realizable, and has stood the test of time. Our own plots of the weighted sum of squares against the Flack parameter (Flack, 1983) have always shown that full-matrix refinement in an iterative linearized least-squares procedure leads to a Flack parameter at the minimum of the weighted sum of squares. There are nevertheless indications that the weak point in the iterative linearized least-squares procedure is the calculation of the standard uncertainty of the Flack parameter, which apparently is sometimes too large and sometimes too small. Smaller standard uncertainties are not necessarily better or more realistic.

We suggest some lines of action which might lead to improvement in absolute-structure determination. With regard to semi-empirical absorption correction, it would appear that the algorithms used need more study and improvement in order to obtain more realistic corrections to the D_{obs} values. Moreover, it is expected that the weakness in absorption correction by numerical or analytical integration arises from the indexing of the crystal faces and the measurement of the crystal dimensions. More powerful and appropriate optical instrumentation on the diffractometer may well provide a solution to this problem. Moreover, the techniques used, as implemented in software, for the transformation of frames of diffraction images into lists of integrated intensities need more study and evaluation.

References

Allen, F. H. (2002). *Acta Cryst.* B**58**, 380–388.
Blessing, R. H. (1995). *Acta Cryst.* A**51**, 33–38.
Clark, R. C. & Reid, J. S. (1995). *Acta Cryst.* A**51**, 887–897.
Flack, H. D. (1983). *Acta Cryst.* A**39**, 876–881.
Flack, H. D. & Bernardinelli, G. (2008). *Acta Cryst.* A**64**, 484–493.
Flack, H. D., Sadki, M., Thompson, A. L. & Watkin, D. J. (2011). *Acta Cryst.* A**67**, 21–34.
Flack, H. D. & Shmueli, U. (2007). *Acta Cryst.* A**63**, 257–265.
Hahn, Th. & Klapper, H. (2002). *Point groups and crystal classes*, Figure 10.1.3.2. *International Tables for Crystallography*, Vol. A, *Space-Group Symmetry*, edited by Th. Hahn, 5th ed. Dordrecht: Kluwer Academic Publishers.
Le Page, Y., Gabe, E. J. & Gainsford, G. J. (1990). *J. Appl. Cryst.* **23**, 406–411.
Parsons, S. (2011). *Acta Cryst.* A**67**, C191.
Parsons, S. & Flack, H. (2004). *Acta Cryst.* A**60**, s61.
Rogers, D. (1950). *Acta Cryst.* **3**, 455–464.
Shmueli, U. & Flack, H. D. (2009). *Acta Cryst.* A**65**, 322–325.
Udupa, M. R. & Krebs, B. (1979). *Inorg. Chim. Acta*, **33**, 241–244.
Watkin, D. J. (2011). Personal communication.
Wilson, A. J. C. (1949). *Acta Cryst.* **2**, 318–321.
Zhu, H.-Y. & Jiang, S.-D. (2007). *Acta Cryst.* E**63**, o2833.