

Acta Crystallographica Section C

**Crystal Structure
Communications**

ISSN 0108-2701

Absolute-structure reports

Howard D. Flack*Acta Cryst.* (2013). **C69**, 803–807

Copyright © International Union of Crystallography

Author(s) of this paper may load this reprint on their own web site or institutional repository provided that this cover page is retained. Republication of this article or its storage in electronic databases other than as specified above is not permitted without prior permission in writing from the IUCr.

For further information see <http://journals.iucr.org/services/authorrights.html>



Acta Crystallographica Section C: Crystal Structure Communications specializes in the rapid dissemination of high-quality studies of crystal and molecular structures of interest in fields such as chemistry, biochemistry, mineralogy, pharmacology, physics and materials science. The numerical and text descriptions of each structure are submitted to the journal electronically as a Crystallographic Information File (CIF) and are checked and typeset automatically prior to peer review. The journal is well known for its high standards of structural reliability and presentation. *Section C* publishes approximately 1000 structures per year; readers have access to an archive that includes high-quality structural data for over 10000 compounds.

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

Absolute-structure reports

Howard D. Flack

Chimie minérale, analytique et appliquée, University of Geneva, Geneva,
Switzerland

Correspondence e-mail: crystal@flack.ch

Received 13 May 2013

Accepted 29 May 2013

All the 139 noncentrosymmetric crystal structures published in *Acta Crystallographica Section C* between January 2011 and November 2012 inclusive have been used as the basis of a detailed study of the reporting of absolute structure. These structure determinations cover a wide range of space groups, chemical composition and resonant-scattering contribution. Defining A and D as the average and difference of the intensities of Friedel opposites, their level of fit has been examined using $2AD$ and selected- D plots. It was found, regardless of the expected resonant-scattering contribution to Friedel opposites, that the Friedel-difference intensities are often dominated by random uncertainty and systematic error. An analysis of data collection strategy is provided. It is found that crystal-structure determinations resulting in a Flack parameter close to 0.5 may not necessarily be from crystals twinned by inversion. $\text{Friedif}_{\text{stat}}$ is shown to be a robust estimator of the resonant-scattering contribution to Friedel opposites, very little affected by the particular space group of a structure nor by the occupation of special positions. There is considerable confusion in the text of papers presenting achiral noncentrosymmetric crystal structures. Recommendations are provided for the optimal way of treating noncentrosymmetric crystal structures for which the experimenter has no interest in determining the absolute structure.

Keywords: absolute structure; Flack parameter; non-centrosymmetric structure.

1. Introduction

Acta Crystallographica Section C has recently published its second special virtual issue which is devoted to the topic of absolute structure (http://journals.iucr.org/special_issues/2012/absolutestructure). The current author wrote the editorial for this special issue (Flack, 2012). All 139 noncentrosymmetric crystal structures published in the special virtual issue of *Acta Crystallographica Section C* coming from publications between January 2011 and November 2012 inclusive were studied. The current note reports the findings of a systematic survey of these noncentrosymmetric crystal-structure deter-

minations and comments on how these may be better understood and improved.

2. Data mining

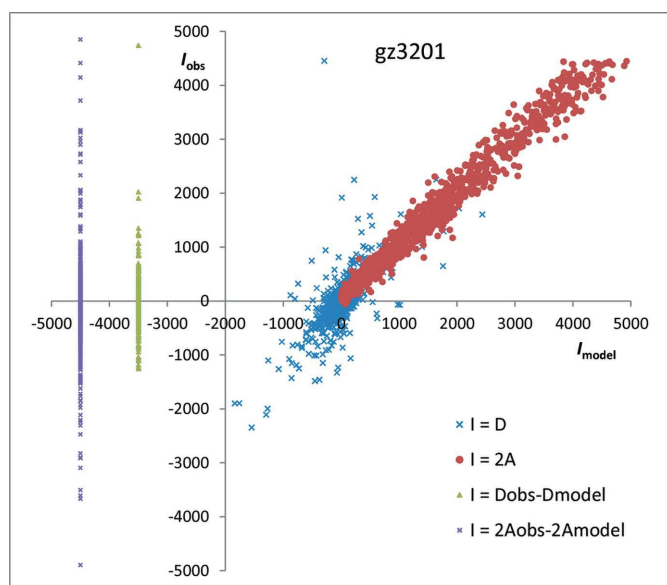
For each publication, the material comprises the paper itself and the all-important supplementary files containing the data measurement and structural information, and the observed and model diffraction intensities. The text of each paper was read with attention.

As it is of particular interest to determine the degree of fit between the observed intensities and those calculated from the solved structure, the previous work of Flack *et al.* (2011) and Parsons *et al.* (2012) was followed by plotting $2A_{\text{obs}}$ against $2A_{\text{model}}$ and D_{obs} against D_{model} on the same $2AD$ graph. $A(hkl)$ and $D(hkl)$ are, respectively, the average and the difference of the intensities of Friedel opposites (hkl and $\bar{h}\bar{k}\bar{l}$); $A(hkl) = \frac{1}{2}[|F(hkl)|^2 + |F(\bar{h}\bar{k}\bar{l})|^2]$ and $D(hkl) = |F(hkl)|^2 - |F(\bar{h}\bar{k}\bar{l})|^2$. Subscript 'obs' refers to the observed values and 'model' refers to a crystal structure model allowing for twinning by inversion. The $2AD$ graphs are constructed to cover the full range of the values of D_{obs} in a data set, with both axes extending over the same range of values. A good fit between $2A_{\text{obs}}$ and $2A_{\text{model}}$, and D_{obs} and D_{model} shows up as a distribution of data points spread around the straight line of slope 1 passing through the origin. In order to prepare the $2AD$ plots and the various statistical descriptors, the intensity data of each structure were separated into three classes of reflections: (p) pairs of Friedel opposites of acentric reflections, (a) unpaired acentric reflections and (c) centric reflections. This classification is dependent on the point group of the crystal structure and relevant values are presented in Table 1 of Shmueli & Flack (2009). Plots of selected D values were also prepared. These contain $D_{\text{obs}}, D_{\text{single}}$ data points, where 'single' refers to a structure model for a single crystal without twinning by inversion. The criteria for the selection of data points for these plots are described in general terms in Parsons *et al.* (2013). The range of D_{obs} and D_{single} values was limited to 20–30% larger than $|D_{\text{single}}|_{\text{max}}$. Data points were also limited by their A_{obs} value, by taking only those points for which either $A_{\text{obs}} > 10\% A_{\text{max}}$ or $A_{\text{obs}} > 1\% A_{\text{max}}$. In these selected D plots, for the ideal case of no random uncertainty and no systematic error in the D_{obs} values, a single crystal in the orientation of the model will show a straight line of points passing through the origin with a slope of +1. If the model is inverted, the slope will be -1 . A crystal twinned by inversion in a proportion of 50:50 will show a line of slope 0. Random uncertainty and systematic error in D_{obs} will cause scatter of the data points about these ideal situations.

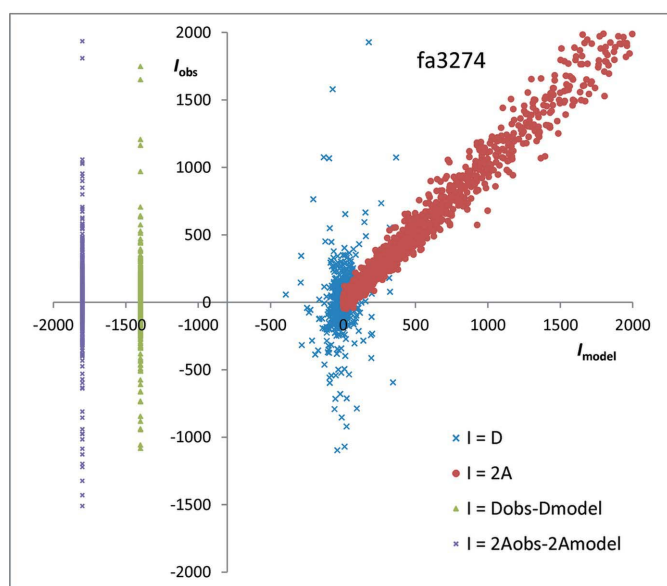
Four Friedif values were calculated for those data sets containing the necessary information. Friedif quantifies the ratio of the root-mean-square Friedel difference to the Friedel average, and is dependent on the contribution of resonant scattering to the diffraction intensities at the wavelength of measurement. $\text{Friedif}_{\text{stat}}$ is a theoretical value based on statistical considerations and is calculated from the chemical composition of the compound and the wavelength of the X-

radiation. It is defined in Flack & Shmueli (2007), Shmueli *et al.* (2008), Shmueli & Flack (2009) and Flack *et al.* (2011). $Friedif_{obs}$ is derived from the A_{obs} and D_{obs} values. As A and D have different dependencies on $\sin(\theta)/\lambda$, a specialized normalization technique needs to be applied (Parsons *et al.*, 2012). $Friedif_{model}$ and $Friedif_{single}$ are derived from the calculated $|F|^2$ in the data sets. The subscripts ‘model’ and ‘single’ are defined above.

Table 1 in the *Supplementary materials* gives relevant statistics on the structures. The spreadsheet from which Table 1 was prepared is available at <http://crystal.flack.ch/ACVirtual2012.xlsx>.



(a)



(b)

Figure 1

2AD plots for gz3201 in (a) and fa3274 in (b). Both gz3201 and fa3274 have $Friedif_{stat} \approx 470$. gz3201 shows a good fit of D_{obs} to D_{model} , whereas for fa3274 the D_{obs} are dominated by random uncertainties and systematic errors.

Each paper is identified by its *Acta Cryst. C* Co-editor code, comprising two alphabetic characters followed by a 4-digit integer. If there is only one noncentrosymmetric structure presented in a paper, the structure itself is also identified by its Co-editor code. For papers containing several noncentrosymmetric structures, each Co-editor code is augmented by the structure label used in the paper, *e.g.* em3048IIIa. Table 1 presents the following information: co-editor code, space group, Z' (the number of molecules in the asymmetric unit), Flack $x(u)$ parameter (Flack, 1983), $R_{|F|^2 > 2u(|F|^2)}$, R_A (a conventional R factor on all A), R_D (a conventional R factor on all D), R_{Aweak} (a conventional R factor on reflections with $A_{obs} < |D_{obs}|_{max}$), N_p (number of Friedel pairs of acentric reflections), N_a (number of unpaired acentric reflections), N_c (number of centric reflections), $Friedif_{stat}$, $Friedif_{obs}$, $Friedif_{model}$, $Friedif_{single}$, category and literature reference with its clickable doi. The category describes the fit of D_{obs} to D_{model} in three classes: in category 1 there is a good fit, in category 3 there is no fit, and category 2 describes intermediate cases. The 2AD and selected- D plots are also available in the *Supplementary materials*.

3. Analysis

The 139 crystal structures in this survey cover a wide range of space groups, chemical composition and resonant-scattering contribution. $Friedif_{stat}$ values range from 3 to 1386. There are five structures in the triclinic family (*i.e.* space group $P1$), plenty in the monoclinic and orthorhombic families, seven in the tetragonal, six in the hexagonal and one in the cubic families.

For some compounds, the intensity data were not suitable for producing 2AD and selected- D plots. In the main these were compounds with a low value of $Friedif_{stat}$ for which the Friedel opposites had been merged and averaged before the final refinement. Due to the overlap of Friedel opposites in a powder diffraction experiment, it is not possible to produce 2AD and selected D plots from these; there were four such cases. The intensity data of eg3082 (Haghjoo *et al.*, 2012) had a data manipulation or software error. The data had clearly been merged and averaged in point group 2 rather than the m appropriate to its space group, Cc . For the remaining 97 crystal structures, 2AD and selected D plots were prepared. The reader is strongly advised to examine these 2AD and selected D plots to come to one's own opinion on what has been achieved in these crystal structure analyses as concerns absolute-structure determination.

The $2A_{obs}, 2A_{model}$ data points of all compounds are distributed about a straight line of slope 1 passing near to the origin. There is thus a good or excellent fit between $2A_{obs}$ and $2A_{models}$.

Many of the determinations, especially those with a low value of $Friedif_{stat}$, show that the D_{obs}, D_{model} data points do not follow any discernible straight line of slope 1 passing through the origin, but are distributed about the straight line with $D_{model} = 0$. There is thus no indication of any observed resonant-scattering contribution to the Friedel differences.

This effect is interpreted as being due to the intensity data being dominated by random uncertainties and systematic errors. In the structure analyses studied here, the first sign of any visible resonant-scattering effect in the observed intensity data occurs at $Friedif_{stat} \simeq 70$. There is thus a vast area for improvement in data measurement and correction techniques.

With values of $Friedif_{stat}$ larger than 400, crystal structures are approached with high symmetry (*i.e.* the tetragonal, hexagonal and cubic families) for which, in general, the $D_{obs}D_{model}$ data points very clearly follow a straight line of slope 1 passing through the origin. The $2AD$ plot of *gz3201* (Falvello *et al.*, 2011) presented in Fig. 1(a) is for a compound with $Friedif_{stat} \simeq 470$ (incidentally, one observes that the $2A_{obs}2A_{model}$ data points do not appear to pass through the origin). Clearly the data points for *gz3201* follow a straight line of slope 1 passing through the origin. However, a high value of $Friedif_{stat}$ does not guarantee an ideal distribution of the $D_{obs}D_{model}$ data points, as can be seen in Fig. 1(b) showing the $2AD$ plot of *fa3274* (de Candia *et al.*, 2012) which is a compound also with $Friedif_{stat} \simeq 470$. For *fa3274*, the data points are clustered around the straight line $D_{model} = 0$, indicating no observable resonant-scattering signal. Moreover, compounds with lower values of $Friedif_{stat}$ can give excellent $2AD$ plots, see for example compound *sk3422III* (Fábry *et al.*, 2012) which has a medium value of $Friedif_{stat}$ of 328. The problem of poor $2AD$ plots is entirely one of data measurement and treatment. For *gz3201*, Falvello *et al.* (2011) have undertaken their experimentation with success whereas for *fa3274*, de Candia *et al.* (2012) have produced intensity data for which the D_{obs} are dominated by random uncertainty and systematic error.

The lowest value of R_D , a conventional R factor on the D values, in these data sets is 28.5% for structure *lg3086I* (Guesmi & Driss, 2012). This is the best that has been seen to date.

The compounds *fa3259* (Waltenberger *et al.*, 2011; $Friedif_{stat} = 36$) and *bm3104III* (Frampton *et al.*, 2011; $Friedif_{stat} = 22$) are examples of careful work from the user community in pharmacy. Their compounds contain only light atoms and the interest was to determine the absolute configuration of the molecule through the determination of the absolute structure of the crystal. The authors are well informed and have followed all of the standard recommendations to achieve a reliable result. Cu $K\alpha$ radiation was chosen, a full sphere of intensity data was measured, the data were measured at low temperature, the absolute-structure determination was undertaken by determining the usual absolute-structure parameter (Flack, 1983) and other advanced indicators relying on Bayesian statistics (Hooft *et al.*, 2008). It seems entirely justified to say that they have followed all of the recommendations of the crystallographic community and obtained a satisfactory result. Sadly, the $2AD$ and selected D plots of these two crystal structure determinations, available in the *Supplementary materials*, tell a different story. The $D_{obs}D_{model}$ data points are grouped around the straight line of $D_{model} = 0$ and the range of $|D_{obs}|$ is much larger than that of $|D_{model}|$. This behaviour is indicative of D_{obs} values which are

dominated by random uncertainty and systematic error and for which the resonant-scattering signal is either very weak or non-existent.

It would seem helpful to make some general remarks on authors' comprehension of the subject area of absolute structure and absolute configuration. There is a deep rift between the understanding and comments on the subject between:

(i) chiral (necessarily noncentrosymmetric) crystal structures and the deduction of the absolute configuration of molecules: this is of a high level,

(ii) achiral noncentrosymmetric crystal structures where there are many infidelities in the texts (*e.g.* the authors state that they have determined the absolute configuration of the molecules in such a crystal structure). This is a problem that writers of textbooks on crystallography and organizers of schools on crystallography need to address.

A short analysis of the data collection strategies was also undertaken. For each data set, counts of the number of paired acentric, unpaired acentric and centric reflections were available. The proportion of centric reflections is out of the control of the experimenter and is dictated by the symmetry of the crystal and the cell dimensions. However, the ratio of unpaired acentric to all acentric reflections (paired and unpaired) is under experimental control and is of great importance for undertaking an analysis of the structure determination in terms of A and D , since the latter values cannot be calculated for unpaired acentric reflections. One requires, for as many as possible of the acentric reflections, that both members of each Friedel pair be measured. It was found that 42% of the data sets had a ratio of unpaired acentric to all acentric reflections of 1% or less, for 65% this proportion was 5% or less. However, for 19% of the data sets, this proportion was 10% or more, rising to the record level of 93% in one case.

Crystals twinned by inversion in a 50:50 ratio have a Flack parameter of 0.5. Ideally, in the absence of random uncertainty and systematic error, their D_{obs} against D_{single} plot should show a line of data points with $D_{obs} = 0$. Consequently, the 12 structure analyses [*gz3204* (Arlin *et al.*, 2012), *ku3043* (Cora *et al.*, 2011), *bi3018* (Zhai & Xu, 2011), *sk3390* (Simmons *et al.*, 2011), *gg3263* (Szalda *et al.*, 2012), *fg3255* (Ojala *et al.*, 2012), *wq3017* (Zhong & Qian, 2012), *tp3005* (Zhang *et al.*, 2012), *bm3112* (Cordes *et al.*, 2011), *fn3089* (Hendsbee *et al.*, 2011), *qs3001* (Pan *et al.*, 2011) and *ku3054* (Lukashuk *et al.*, 2011); $Friedif_{stat}$ values in the range 6–1386] were examined, as these reported a value of the Flack parameter close to 0.5, *i.e.* $0.25 < x < 0.75$. It is particularly in the selected D plot that one would hope to find a distribution of data points related to the ideal behaviour. However, in the data sets studied, with the possible exception of *wq3017* and *fn3089*, this is not the case. The plots look more like D_{obs} values dominated by random uncertainty and systematic error. The evidence that the crystals are really twinned by inversion in a proportion close to 50:50 is very far from convincing. As stated in Parsons *et al.* (2012), the expected value of the Flack parameter for a data set of random D_{obs} values is 0.5. It is perhaps for this reason that the 12 sets examined here have a value close to 0.5.

The $\text{Friedif}_{\text{stat}}$ value derived and used by Flack & Shmueli (2007), Shmueli *et al.* (2008), Shmueli & Flack (2009) and Flack *et al.* (2011) requires a fixed chemical composition of the crystal and a known wavelength of the X-radiation, and assumes a very large number of independent reflections, space group $P1$ and no atoms in special positions. Centrosymmetric substructures may be taken into account in some circumstances. $\text{Friedif}_{\text{single}}$ is taken from the model structure-factor amplitudes of the crystal structure which perhaps includes some atoms in special positions, a space group that is not necessarily $P1$ and a resolution of the data (*i.e.* the number of reflections) which is limited. In practice, $\text{Friedif}_{\text{stat}}$ is used as an estimate of $\text{Friedif}_{\text{single}}$ prior to structure solution. In Fig. 2, the plot of $\text{Friedif}_{\text{single}}$ against $\text{Friedif}_{\text{stat}}$ is shown for the data sets studied here. One sees that $\text{Friedif}_{\text{stat}}$ is indeed a robust estimate of $\text{Friedif}_{\text{single}}$. Fig. 3 shows the plot of $\text{Friedif}_{\text{obs}}$ against $\text{Friedif}_{\text{stat}}$. It tells a completely different story. Frequently the $\text{Friedif}_{\text{obs}}$ value is much larger than that of $\text{Friedif}_{\text{stat}}$, indicating that $\text{Friedif}_{\text{obs}}$ is dominated by random uncertainty and systematic error.

4. Concluding remarks

In short, for the evaluation of absolute-structure determination, over the years there has been a heavy reliance on the statistics (values and standard uncertainties) of derived parameters or something equivalent, and little study of the fit of the model to the observed quantities. There has been little questioning of the quality of crystals, data collection and correction methodologies. The onus is on the crystallographers to reveal to advanced users how their data measurement and treatment may be improved.

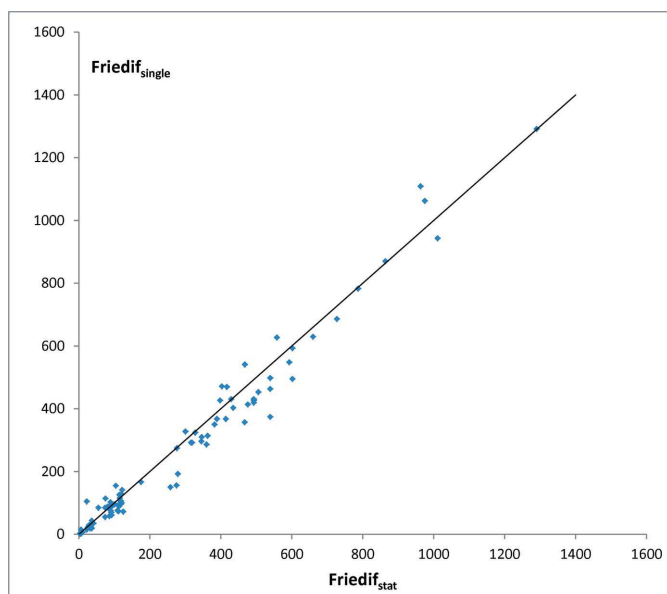


Figure 2
Plot of $\text{Friedif}_{\text{single}}$ against $\text{Friedif}_{\text{stat}}$ for the studied data sets. $\text{Friedif}_{\text{stat}}$ is shown to be a robust estimate of $\text{Friedif}_{\text{single}}$.

The necessary procedure to undertake absolute-structure determination whereby one obtains a value of the Flack parameter with as small a standard uncertainty as possible was outlined in §5 of Parsons *et al.* (2013). However, no suggestions were given for the optimal treatment of a noncentrosymmetric crystal structure for which the experimenter has no interest in determining the absolute structure. The following points need to be taken into account for such a case: (a) the experimenter cannot be required to make measurements in which all Friedel opposites are measured; (b) from the results presented above, some experiments yield good results for systems with low $\text{Friedif}_{\text{stat}}$ and some yield poor results for systems with large $\text{Friedif}_{\text{stat}}$, so the resonant-scattering contribution may or may not be visible in the intensity data; (c) it is inappropriate to require that Friedel opposites be averaged as the data set inevitably contains both pairs of Friedel opposites of acentric reflections and unpaired acentric reflections: on averaging the former provide a subset of the data corresponding to a Flack parameter of 0.5 and the latter a disjoint subset corresponding to the Flack parameter of the crystal. None of the existing software enables the two classes of reflections to be identified for the least-squares refinement.

Consequently, the most appropriate procedure for the optimal treatment of a noncentrosymmetric crystal structure for which the experimenter has no interest in determining the absolute structure is: (i) do not average Friedel opposites and do not classify the reflections into the paired acentric,

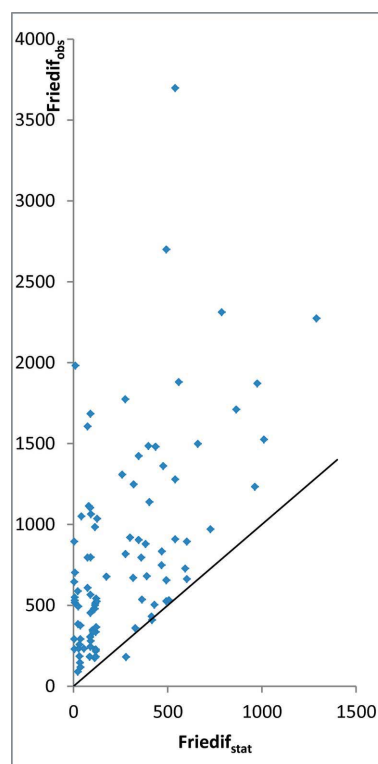


Figure 3
Plot of $\text{Friedif}_{\text{obs}}$ against $\text{Friedif}_{\text{stat}}$ for the studied data sets. $\text{Friedif}_{\text{obs}}$ is shown to be frequently dominated by random uncertainty and systematic error.

unpaired acentric and centric classes, and (ii) refine the Flack parameter by full-matrix least squares.

The editor of *Acta Crystallographica Section C*, A. Linden, is thanked for providing the list of noncentrosymmetric crystal structures published therein in 2011 and 2012. D. J. Watkin and S. Parsons are thanked for reading and commenting on a draft version of the manuscript.

Supplementary data for this paper are available from the IUCr electronic archives (Reference: LN3158). Services for accessing these data are described at the back of the journal.

References

- Arlin, J.-B., Kennedy, A. R. & Shankland, K. (2012). *Acta Cryst.* **C68**, m29–m33.
- Cora, I., Czugler, M., Dódy, I. & Rečnik, A. (2011). *Acta Cryst.* **C67**, i33–i35.
- Cordes, D. B., Hua, G., Slawin, A. M. Z. & Woollins, J. D. (2011). *Acta Cryst.* **C67**, o509–o514.
- De Candia, A. G., Molnar, M., Slep, L. D. & Baggio, R. (2012). *Acta Cryst.* **C68**, m121–m126.
- Fábry, J., Fridrichová, M., Dušek, M., Fejfarová, K. & Krupková, R. (2012). *Acta Cryst.* **C68**, o76–o83.
- Falvello, L. R., Forcén-Vázquez, E., Mayoral, I., Tomás, M. & Palacio, F. (2011). *Acta Cryst.* **C67**, m359–m363.
- Flack, H. D. (1983). *Acta Cryst.* **A39**, 876–881.
- Flack, H. D. (2012). *Acta Cryst.* **C68**, e13–e14.
- Flack, H. D., Sadki, M., Thompson, A. L. & Watkin, D. J. (2011). *Acta Cryst.* **A67**, 21–34.
- Flack, H. D. & Shmueli, U. (2007). *Acta Cryst.* **A63**, 257–265.
- Frampton, C. S., MacNicol, D. D. & Wilson, D. R. (2011). *Acta Cryst.* **C67**, o188–o191.
- Guesmi, A. & Driss, A. (2012). *Acta Cryst.* **C68**, i55–i59.
- Haghjoo, F., Barnes, N. A., Pritchard, R., Godfrey, S. M. & Ratcliffe, S. (2012). *Acta Cryst.* **C68**, o216–o219.
- Hendsbee, A. D., Masuda, J. D. & Piórko, A. (2011). *Acta Cryst.* **C67**, m391–m394.
- Hooft, R. W. W., Straver, L. H. & Spek, A. L. (2008). *J. Appl. Cryst.* **41**, 96–103.
- Lukashuk, L. V., Lysenko, A. B., Krautscheid, H. & Domasevitch, K. V. (2011). *Acta Cryst.* **C67**, m378–m383.
- Ojala, W. H., Arola, T. M., Brigino, A. M., Leavell, J. D. & Ojala, C. R. (2012). *Acta Cryst.* **C68**, o270–o278.
- Pan, M.-Y., Wu, X.-H., Luo, D.-B., Huang, W. & He, Y. (2011). *Acta Cryst.* **C67**, o175–o178.
- Parsons, S., Flack, H. D. & Wagner, T. (2013). *Acta Cryst.* **B69**, 249–259.
- Parsons, S., Pattison, P. & Flack, H. D. (2012). *Acta Cryst.* **A68**, 736–749.
- Shmueli, U. & Flack, H. D. (2009). *Acta Cryst.* **A65**, 322–325.
- Shmueli, U., Schiltz, M. & Flack, H. D. (2008). *Acta Cryst.* **A64**, 476–483.
- Simmons, T. R., Pickett, C. J. & Wright, J. A. (2011). *Acta Cryst.* **C67**, o1–o5.
- Szalda, D. J., Ramig, K., Lavinda, O., Koren, Z. C. & Massa, L. (2012). *Acta Cryst.* **C68**, o160–o163.
- Waltenberger, B., Rollinger, J. M., Griesser, U. J., Stuppner, H. & Gelbrich, T. (2011). *Acta Cryst.* **C67**, o409–o412.
- Zhai, J. & Xu, L. (2011). *Acta Cryst.* **C67**, m202–m204.
- Zhang, W., Oliver, A. G. & Serianni, A. S. (2012). *Acta Cryst.* **C68**, o7–o11.
- Zhong, K.-L. & Qian, M.-Y. (2012). *Acta Cryst.* **C68**, m265–m268.