

# HUG and SQUEEZE: using *CRYSTALS* to incorporate resonant scattering in the SQUEEZE structure-factor contributions to determine absolute structure

Richard I. Cooper,<sup>a\*</sup> Howard D. Flack<sup>b‡§</sup> and David J. Watkin<sup>a</sup>

Received 17 August 2017

Accepted 25 September 2017

Edited by P. Raithby, University of Bath, UK

‡ Deceased 2 February 2017.

§ RIC & DJW are grateful to HDF for his inspiration and encouragement. We have tried to preserve his contributions to this manuscript unaltered where possible; any errors or omissions are our own.

**Keywords:** disorder; resonant scattering; absolute structure.

<sup>a</sup>Chemical Crystallography, University of Oxford, 12 Manseld Road, Oxford, Oxfordshire OX1 3TA, England, and <sup>b</sup>Chimie minerale, analytique et appliquee, University of Geneva, Geneva, Switzerland. \*Correspondence e-mail: richard.cooper@chem.ox.ac.uk

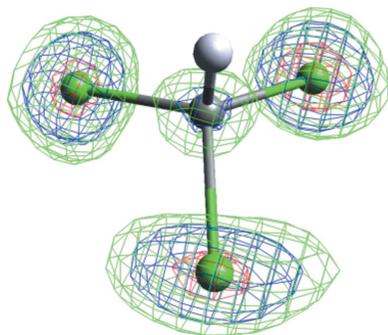
The resonant-scattering contributions to single-crystal X-ray diffraction data enable the absolute structure of crystalline materials to be determined. Crystal structures can be determined even if they contain considerably disordered regions because a correction is available *via* a discrete Fourier transform of the residual electron density to approximate the X-ray scattering from the disordered region. However, the corrected model cannot normally account for *resonant scattering* from atoms in the disordered region. Straightforward determination of absolute structure from crystals where the strongly resonantly scattering atoms are not resolved has therefore not been possible. Using an approximate resonant-scattering correction to the X-ray scattering from the disordered regions, we have developed and tested a procedure (HUG) to recover the absolute structure using conventional Flack *x* refinement or other post-refinement determination methods. Results show that in favourable cases the HUG method works well and the absolute structure can be correctly determined. It offers no useful improvement in cases where the original correction for the disordered region scattering density is problematic, for example, when a large fraction of the scattering density in the crystal is disordered, or when voids are not occupied equally by the disordered species. Crucially, however, if the approach does not work for a given structure, the statistics for the absolute structure measures are not improved, meaning it is unlikely to lead to misassignment of absolute structure.

## 1. Background

The refinement of crystal structures usually requires the structural parameters to be adjusted by the method of least squares to minimize the differences between  $|F_o|$  and  $|F_c|$  (or  $I_o$  and  $I_c$ , where  $I$  are squared structure amplitudes,  $|F^2|$ ).

Cases exist where this procedure is complicated by the fact that part of the structure cannot easily be modelled by clearly defined individual atoms. This situation may exist in extended lattice structures with voids which contain independent molecules (host and guest structures), or in discrete molecule structures in which the lattice is stabilized by the inclusion of solvent molecules or ions necessary to preserve charge balance.

If these subsidiary molecules are not spatially constrained by the surrounding lattice, they may have freedom to move even in the solid state, or have the possibility of occupying alternative positions and orientations. Often this ambiguity can be modelled by large anisotropic atomic displacement factors (ADPs) or by the superposition of displaced partially occupied images of the molecule. In unfavourable cases, the average scattering density in the cavity cannot reasonably be modelled by independent atoms. This situation has been



addressed by replacing the atomic model of the contents of the cavity by the discrete Fourier transform of the electron density in the cavity computed from the observed structure amplitudes and phases obtained from the atomic model of the resolved part of the structure (van der Sluis & Spek, 1990; Spek, 2015).

### 1.1. The SQUEEZE procedure

The structure factor can be computed either as the Fourier transform of the continuous periodic electron density in the crystal:

$$F_{\mathbf{h}} = \int_V \rho(\mathbf{x}) e^{2\pi i(\mathbf{h}\mathbf{x})} dV, \quad (1)$$

or as the summation of the contributions from individual ‘atoms’:

$$F_{\mathbf{h}} = \sum_j f_j e^{2\pi i(\mathbf{h}\mathbf{x}_j)}. \quad (2)$$

SQUEEZE defines a region of the unit cell,  $V$ , in which the disordered part of the crystal structure is located. No atomic model of  $V$  is available. The content of  $V$  is represented by a real electron density,  $\rho_V(\mathbf{x})$ :

$$\rho_V(\mathbf{x}) = \begin{cases} \rho(\mathbf{x}), & \text{if } \mathbf{x} \in V \\ 0, & \text{otherwise} \end{cases}. \quad (3)$$

The SQUEEZE procedure then uses the hybrid structure-factor expression:

$$F_{\mathbf{h}} = \sum_j f_j e^{2\pi i(\mathbf{h}\mathbf{x}_j)} + \int_V \rho_V(\mathbf{x}) e^{2\pi i(\mathbf{h}\mathbf{x})} dV. \quad (4)$$

The first term is a summation over the resolved atoms as in equation 2. The integral in the second term is evaluated for  $\mathbf{x} \in V$ , which contains unresolved electron density. The resulting expression accounts for scattering from both resolved atoms and unresolved electron density.

The integral can be replaced by a summation over a suitable resolution grid of electron density:

$$F_{\mathbf{h}} = \sum_j f_j e^{2\pi i(\mathbf{h}\mathbf{x}_j)} + \sum_x \rho_V(\mathbf{x}) e^{2\pi i(\mathbf{h}\mathbf{x})}. \quad (5)$$

### 1.2. Resonant scattering

If a material is in a noncentrosymmetric space group and it contains one or more atoms with significant resonant scattering for the wavelength in use, it may be possible to determine the absolute structure of the sample using Flack’s interpretation of the observed Bijvoet differences (Flack, 1983). Representing  $|F|_{\mathbf{h}}^2$  by  $I^+$  and  $|F|_{-\mathbf{h}}^2$  by  $I^-$ :

$$I_o^+ \approx I_c^+ = (1 - x) I_s^+ + x I_s^-, \quad (6)$$

where the subscript  $s$  indicates a quantity computed from the atomic model with the Flack parameter  $x$  set to zero (*i.e.* a nontwinned *single* crystal),  $c$  a quantity computed from an inversion-twinned model (*i.e.* Flack parameter not necessarily zero) and  $o$  an observed quantity (Cooper *et al.*, 2016). The

Flack( $x$ ) may be determined either during the least-squares refinement, or by post-refinement methods (Parsons *et al.*, 2013). The success of this procedure depends upon the quality of the data and upon the absolute structure resolving power of the material, conveniently estimated by Friedif (Flack & Shmueli, 2007). The magnitude of Friedif is increased in the presence of atoms with large resonant scattering factors, even if these atoms are not part of an enantiomerically pure host material. This means that the possibility of reliably determining the absolute structure of an all-light-atom structure can be increased if the material crystallizes with a suitable molecule of solvation.

In the case that the solvent molecule is highly disordered, it may not be possible to model it with discrete atoms, so that the only way to complete the analysis is to SQUEEZE the solvent region, which in the standard implementation makes no allowance for a resonant contribution from the solvent to the computed structure amplitudes. This means that a conventionally SQUEEZED solvent cannot be used to help in the determination of absolute structure, as demonstrated at the start of §3 below.

## 2. Methods

Fourier transformation of  $\rho_V(\mathbf{x})$  leads to its structure factor,  $F(\rho_V(\mathbf{x}))_{\mathbf{h}}$ , which is added to the structure factor of the atomic model for the ordered part of the unit cell.

The structure factor is a complex number having both magnitude and phase, and may be written as  $F_{\mathbf{h}} = A_{\mathbf{h}} + iB_{\mathbf{h}}$ , where  $A_{\mathbf{h}}$  and  $B_{\mathbf{h}}$  are the real and imaginary parts of  $F_{\mathbf{h}}$ , respectively, and  $i = \sqrt{-1}$ .

### 2.1. The HUG procedure – enhancing SQUEEZE to include resonant scattering

*If the disordered volume contains atoms with strong resonant scattering, how might one proceed to incorporate this resonant scattering contribution to  $\rho_V(\mathbf{x})$ ?*

**Method 1** Construct a model in which the resonant scattering contribution of  $V$  is distributed uniformly over  $V$ . Then  $\rho_V(\mathbf{x})$  can be modified to become the complex  $\rho'_V(\mathbf{x})$ :

$$\rho'_V(\mathbf{x}) = \rho_V(\mathbf{x}) + c + id, \quad (7)$$

in which  $c$  and  $d$  are constants to be chosen or determined in some way.

The inconvenience of this simple model is that the resonant-scattering contribution is distributed widely over  $V$  and its contribution in reciprocal space will diminish more rapidly as a function of  $\sin\theta/\lambda$  than with an atomic model. Such consideration leads to:

**Method 2** Construct a model in which the resonant scattering contribution of  $V$  is assumed to be proportional to  $\rho_V(\mathbf{x})$  at each point  $\mathbf{x}$ . Then  $\rho_V(\mathbf{x})$  can be modified to become the complex  $\rho'_V(\mathbf{x})$ :

$$\rho'_V(\mathbf{x}) = \rho_V(\mathbf{x}) + c\rho_V(\mathbf{x}) + id\rho_V(\mathbf{x}) = (1 + c + id)\rho_V(\mathbf{x}). \quad (8)$$

In this way, the major part of the resonant-scattering contribution will be located at the positions of high electron density

Table 1

Light-atom structures containing molecules of solvation or counter-ions.

Flack( $x$ ) is the Flack parameter determined as part of the normal refinement by classical least squares (Flack, 1983). Bijvoet( $d$ ) is the Flack parameter determined by a post-refinement analysis of Bijvoet differences (Cooper *et al.*, 2016).  $Z_h$  is the electron count for the host structure,  $Z_s$  for the solvent or counter-ion. All of the examples except awisac02 were fully characterized in the original work. In this latter material, the dichloromethane was modelled with SQUEEZE. The original  $R$  factor (marked \*) was obtained with the standard SQUEEZE procedure.

Code	Formula	Friedif whole	Friedif no solvent	$Z_h$	$Z_s$	$(Z_h+Z_s)/Z_h$	Original $R$ (%)	Flack( $x$ )	Bijvoet( $d$ )
sgd240	$C_{16}H_{22}NO, C_{10}H_{15}O_4S, C_2H_3N$	73.8	76.0	256	22	1.09	6.4	0.01 (4)	0.05 (4)
fg3257	$2(C_{17}H_{28}N_2O_5), C_2H_6OS$	277.0	33.2	368	34	1.13	2.7	0.02 (1)	0.02 (0)
ky3014	$C_{12}H_{24}O_6, C_6H_6NO_2, 2(H_2O), ClO_4$	88.9	7.0	229	49	1.21	8.2	-0.01 (8)	0.00 (3)
sgd464	$C_{20}H_{27}NO_4, CHCl_3$	124.7	6.1	186	58	1.31	11.1	-0.14 (12)	-0.08 (5)
sgd475	$C_7H_3F_2O_2, C_8H_{16}NO_2, CHCl_3$	530.8	48.1	166	58	1.35	3.1	0.01 (1)	0.01 (1)
sk3422_III	$C_2H_7N_4O, F_{0.12}H_{1.89}O_3P$	327.6	24.7	55	41	1.75	1.9	0.04 (28)	0.05 (1)
awisac02	$C_{36}H_{48}O_4, CH_2Cl_2$	111.2	5.3	296	42	1.14	6.4*	n/a	n/a

in  $\rho_V(\mathbf{x})$ . The advantage of this model is that it is ‘more atomic’ than that of **Method 1**, a positive attribute intended to imply exactly the same as the ‘large  $f$ ’, usually associated with heavier elements, and whose ghosts would leave more miasma<sup>1</sup> in the difference density.

If  $F_h$  is the Fourier transform of  $\rho(\mathbf{x})$ , then the Fourier transform,  $F'_h$ , of  $\rho'_V(\mathbf{x})$  is given by:

$$F'_h = (1 + c)F_h + idF_h \\ = [(1 + c)A_h - dB_h] + i[(1 + c)B_h + dA_h], \quad (9)$$

where  $c$  and  $d$  set the ratios of the real and imaginary parts of the resonant-scattering contribution from the electron density in the disordered region of the crystal.

A reasonable first approximation for  $c$  and  $d$  is to assume that the regions of high electron density in the unresolved volume are those of highest resonant scattering. One may even take the step of assuming that the resonant-scattering contribution is proportional to electron density, so that:

$$c = \frac{\sum f'_{sol}}{\sum f_{sol}} \quad \text{and} \quad d = \frac{\sum f''_{sol}}{\sum f_{sol}}, \quad (10)$$

where the summations are over the expected atoms in the solvent. Note that equation 10 might use  $f_{sol}^0$  instead of  $f_{sol}$  in the denominator, to avoid overcorrection for the resonant signal at higher  $\sin\theta/\lambda$ , however, trial-and-error has shown equation 10 to be the more effective formulation.

The resonant scattering contribution in the solvent region is thus  $(c + id)\rho_V(x)$  which can be used as a modifier to correct the structure-factor components  $A$  and  $B$  of the region  $V$  for resonant scattering. By inspection of equation 9, we obtain

$$A_{hug} = (1 + c)A_{sqz} - d.B_{sqz} \\ B_{hug} = (1 + c)B_{sqz} + d.A_{sqz}, \quad (11)$$

where the subscripts  $sqz$  indicate the complex contribution to the structure factor returned by SQUEEZE due to unresolved electron density in the volume  $V$ , and  $hug$  indicates the same contribution corrected for resonant scattering. Refinement is undertaken in the usual way, except that the  $A$  and  $B$  parts of

the structure factor computed from the resolved atoms are supplemented by the addition of  $A_{hug}$  and  $B_{hug}$ , respectively.

$$A_{calc} = A_{res} + A_{hug} \\ B_{calc} = B_{res} + B_{hug}, \quad (12)$$

where the subscript *res* indicates structure-factor components for the resolved part of the structure.

## 2.2. HUGging in CRYSTALS

Since its inception, *CRYSTALS* has had a facility for storing the precomputed  $A$  and  $B$  parts for a reflection so that they can be added into the  $A$  and  $B$  parts computed from an atomic model (Carruthers, 1977). The original use was to facilitate the development of a poorly resolved part of a structure. The  $A$  and  $B$  parts of the well-resolved atoms were computed once and stored in the database. Structure-factor contributions were then computed from the atoms in experimental models of the disorder and added to the stored parts. This gave significant time-savings when the well-resolved part of the structure contained a large number of atoms compared with the disordered part (Watkin *et al.*, 1985). With the publication of the SQUEEZE program, this procedure could be reversed. For more than 20 years, an interface between SQUEEZE and *CRYSTALS* has enabled the  $A$  and  $B$  parts of the database to hold contributions to the structure factor computed from the discrete Fourier transform of electron density in parts of the unit cell not modelled by independent atoms. This procedure has the virtue that during refinement, the values of  $F_{obs}$  (or  $I_{obs}$ ) are not modified. The enhanced strategy (HUG<sup>2</sup>) represented by equation 12 has been implemented in *CRYSTALS* (Versions after 24/02/2017) by an external module which uses a proposed molecular formula for the solvent to correct the standard output from unmodified SQUEEZE before passing the modified  $A$  and  $B$  parts into *CRYSTALS*.

The concept was evaluated by processing several structures with well-resolved solvent molecules. The  $A$  and  $B$  parts for an atomic model of the solvent were first computed and stored in the *CRYSTALS* database and then used together with the

<sup>2</sup> Just a few weeks before he died, Howard Flack drastically reorganized a draft of this manuscript and changed its prosaic title to ‘HUG and SQUEEZE’. One could not help remembering the CAMEL JOCKEY, with or without the humps (Watkin & Schwarzenbach, 2017).

Table 2

Structures from the literature were re-refined in *CRYSTALS* (against  $F^2$  with *SHELX*-type weights) and the absolute-structure parameters determined.

The solvent/counter-ion was excluded and the structures re-refined using the modified  $A_{hug}$  and  $B_{hug}$  parts of the structure factor computed by SQUEEZE.  $Flack(x)$  is the Flack parameter determined as part of the normal refinement by classical least squares (Flack, 1983).  $Bijvoet(d)$  is the Flack parameter determined by a post-refinement analysis of Bijvoet differences (Cooper *et al.*, 2016).  $R(Av_s)$  and  $R(D_s)$  are defined in equations 13 and 14. The atomic models for structures marked  $\alpha$  were modelled with disordered solvent or counter-ion. The structure marked  $\beta$  was squeezed in the original work. The HUGged  $R$  factor and  $Bijvoet(d)$  marked  $\dagger$  were obtained when the  $d$  parameter was multiplied by 1.5.

Code	HUGged $R$ (%)	HUGged $Flack(x)$	HUGged $Bijvoet(d)$	Space group	Electrons found (expected)	$R(Av_s)$	$R(D_s)$
sgd240	7.1	0.01 (10)	0.27 (3)	$P2_12_12_1$	28 (22)	0.06	0.75
fg3257	4.1	-0.30 (4)	-0.30 (1)	$P2_12_12_1$	47 (58)	0.05	0.33
ky3014 $\alpha$	8.0	0.05 (8)	0.27 (5)	$Pna2_1$	45 (49)	0.10	0.91
sgd464	9.3	0.00 (9)	-0.02 (5)	$P2_1$	54 (58)	0.07	0.44
sgd475 $\alpha$	4.0	0.00 (2)	0.00 (1)	$C2$	61 (58)	0.12	0.40
sk3422_III	13.8	-0.2 (3)	0.06 (4)	$Cc$	41 (41)	0.51	0.91
awisac02 $\beta$	5.9 $\dagger$	0.1 (2)	0.33 (3)	$P2_1$	48 (42)	n/a	n/a

main structure in a normal refinement. These refinements were compared with a refinement in which the  $A$  and  $B$  parts were from a solvent which was SQUEEZED and HUGged. The agreement between the HUGged and atomic structure amplitudes can be estimated by

$$R(Av_s) = \frac{\sum |Av_{hug} - Av_{atom}|}{\sum |Av_{atom}|} \quad (13)$$

$$R(D_s) = \frac{\sum |D_{hug} - D_{atom}|}{\sum |D_{atom}|}, \quad (14)$$

where  $Av_{atom}$  and  $D_{atom}$  are the average and difference of structure-factor magnitudes of a Bijvoet pair computed from a fully atomic model, and  $Av_{hug}$  and  $D_{hug}$  are equivalent values computed from a HUGged model. The agreement can be visualized in plots of  $D_{hug}$  versus  $D_{atom}$ .

### 3. Results

Table 1 lists solvated structures selected from the recent literature where the molecules of interest contained only light

atoms and the solvents were reasonably well defined: sgd240 (Chernega *et al.*, 2009), fg3257 (Bojarska *et al.*, 2012), ky3014 (Shi *et al.*, 2012), sgd464 (Davies *et al.*, 2013a), sgd475 (Davies *et al.*, 2013b), sk3422 (Fábry *et al.*, 2012) and awisac02 (Qian *et al.*, 2016). The absolute structures were confirmed by re-refining the atomic model in *CRYSTALS*. The solvents were then excluded from the structure-factor calculation and modelled using the standard SQUEEZE procedure. The outputs from SQUEEZE were HUGged as explained above, and the structures re-refined. The applicability of the procedure was assessed by comparing the  $Flack(x)$  (Flack, 1983) and  $Bijvoet(d)$  (Cooper *et al.*, 2016) parameters determined from the atomic model and the HUGged model, and by plotting  $D_s$ , the computed Bijvoet difference, for one model against the other.

The effect of modelling regions of the crystal structure containing strong resonant scatterers with scattering from an electron-density map with no resonant scattering effects may be demonstrated by a comparison of the absolute structure statistics for sgd464: the  $Flack(x)$  and  $Bijvoet(d)$  parameters for a complete atomic model are -0.14 (12) and -0.08 (5)

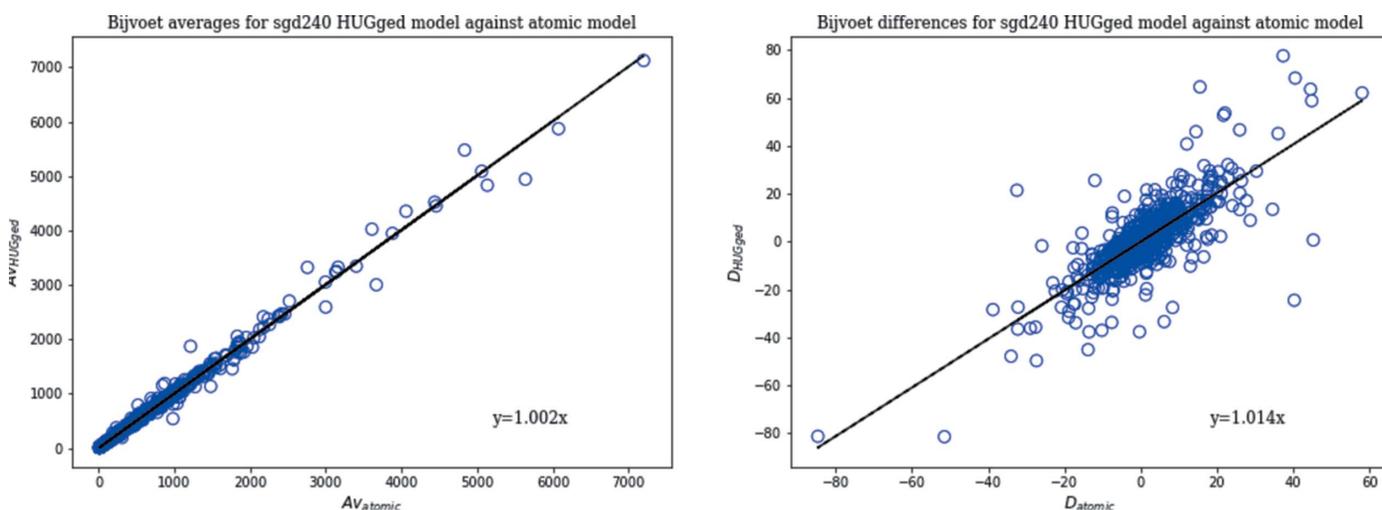


Figure 1

Averages ( $Av_s$ ) and differences ( $D_s$ ) of Bijvoet pairs computed from the HUGged and atomic models of structure sgd240. The gradients of both plots (1.00 and 1.01) show that the HUGged model is a fair approximation to the atomic model;  $R(Av_s) = 0.06$  and  $R(D_s) = 0.80$ . The abscissa is values computed from the fully resolved atomic model and the ordinate values are computed from the HUGged model.

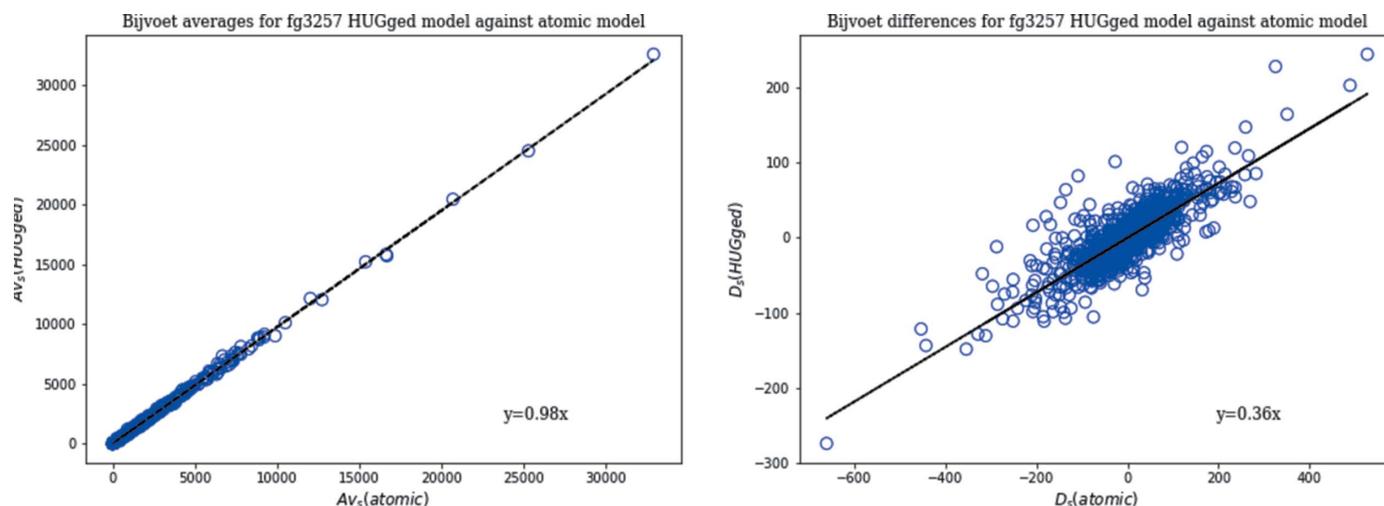


Figure 2

Averages and differences of Bijvoet pairs for structure fg3257. The gradient of the averages (0.98) suggests that the *A* and *B* parts from SQUEEZE are reasonably well scaled and the gradient of the differences (0.36) suggests that the *d* factor used to generate the resonant differences should be multiplied by about 0.3.

respectively, and the *R*1 value is 11.1%. After the chloroform molecule is removed and SQUEEZE applied, the remaining structure can be refined to an *R*1 value of 9.39%, but the Flack(*x*) and Bijvoet(*d*) parameters are now 4.7 (10) and 0.7 (2) (Table 2).

### 3.1. Structure sgd240

This is an organic material of known absolute configuration (from the starting materials) containing a sulfoxide group and a well-behaved acetonitrile of solvation. The data were measured with Mo *K* $\alpha$  radiation. The solvent contains 22 electrons, SQUEEZE returns 114 electrons/cell in the voids, and  $-3$  electrons/cell outside the voids. The principal normal

and resonant-scattering atom is sulfur in the main molecule. The resonant scattering from the solvent is marginal so that the unmodified SQUEEZE refinement is essentially the same as the modified (Fig. 1).

### 3.2. Structure fg3257

This is an organic material containing C, H, N and O atoms, with two independent molecules and one dimethyl sulfoxide solvent in the asymmetric unit. The absolute structure was determined from the X-ray diffraction data. Modifying the SQUEEZE output using constants determined by equation 10 gives a Bijvoet(*d*) parameter of  $-0.30$  (1). Multiplying *c* and *d* by factors of 1.5 and 2.0 gave Bijvoet(*d*) values of  $-0.02$  (1) and  $0.12$  (0), respectively. The Flack(*x*) parameter increased from  $-0.30$  (4) with unscaled *d* values to  $-0.00$  (2) and

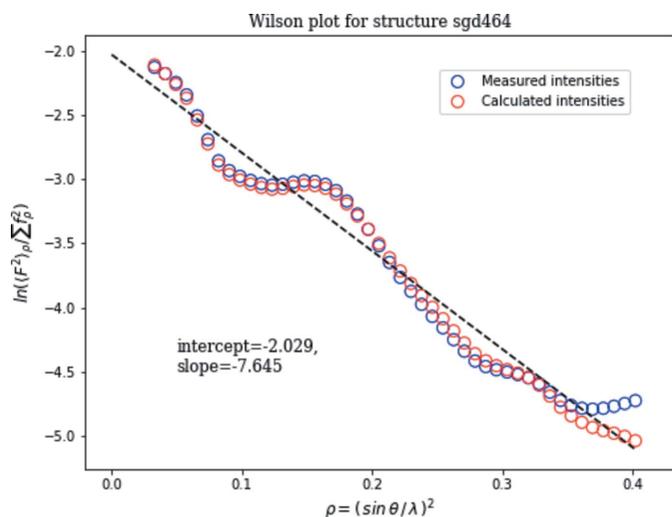


Figure 3

Wilson Plot for structure sgd464. The abscissa is  $\ln((F_{obs}^2)_\rho / \sum f_p^2)$  and the ordinate  $\rho = (\sin\theta/\lambda)^2$ . Blue circles are calculated from measured diffraction intensities and red circles are calculated from the final atomic model (for comparison). The up-turn at about  $\rho = 0.35$  is characteristic of data being recorded beyond the real diffraction limit and therefore consisting mainly of noise.

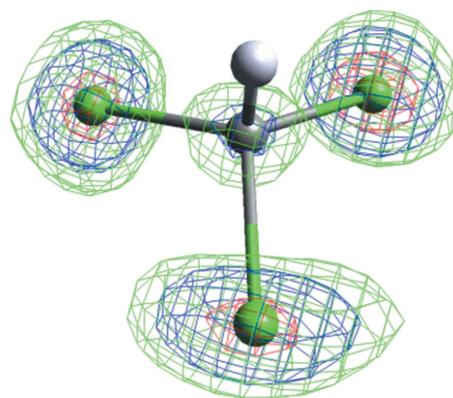


Figure 4

The *A* and *B* parts from SQUEEZE for structure sgd464 were modified to include resonant scattering from the solvent and were used without contributions from the host molecule to compute calculated structure factors and phases, which were then used to compute a Fourier synthesis. The electron density (arbitrary contour levels where red > blue > green) shows a good representation of the solvent. The chloroform molecule has been overlaid as a guide to the eye. Chloroform contains 58 electrons and the unmodified SQUEEZE map contains 54 electrons per void.

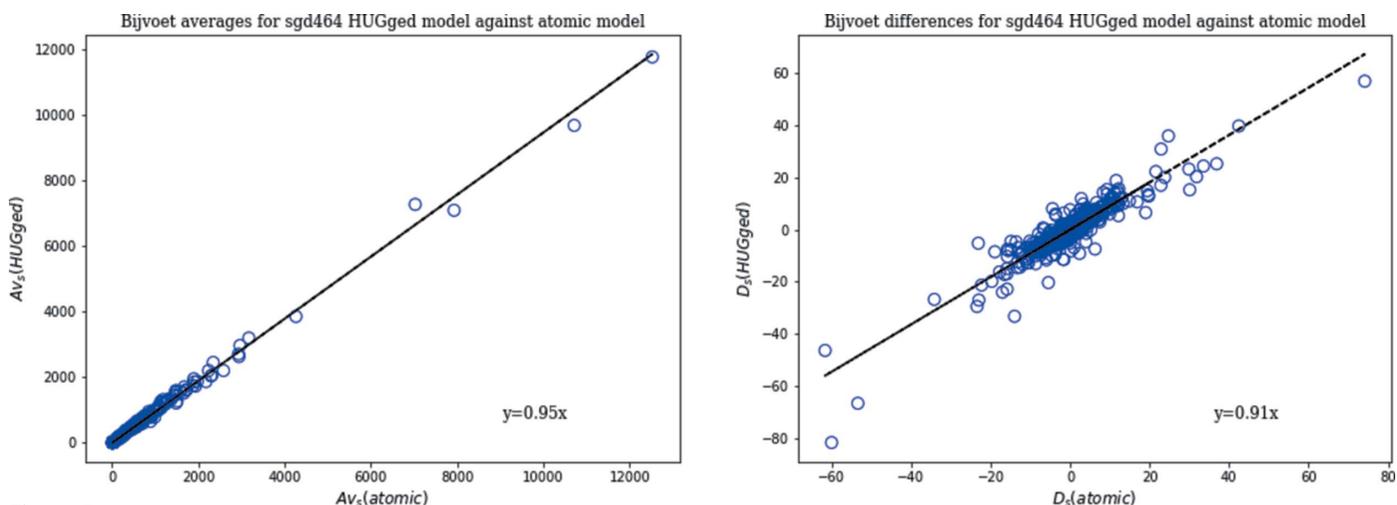


Figure 5

Averages ( $Av_s$ ) and differences ( $D_s$ ) of Bijvoet pairs for structure sgd464 computed from an atomic model of the solvent and a HUGged model. The gradients of both plots (0.95 and 0.91) show that the HUGged model is a fair approximation to the atomic model;  $R(Av_s) = 0.07$  and  $R(D_s) = 0.44$ .

–0.13 (2) with the increasing scaling factors. The near-unity value of the ratio  $(Z_h + Z_s)/Z_h$  (1.13) suggests that the scaling of the  $A$  and  $B$  parts from SQUEEZE is almost correct, confirmed by the gradient (0.98) of the plot of the averaged Bijvoet pairs of the HUGged model versus the atomic model (Fig. 2); the need for scaling of  $d$  is demonstrated in the right-hand plot.

### 3.3. Structure sgd464

This is an organic material containing C, H, N and O atoms, with a diastereomeric ratio (dr) > 99:1 and chloroform of solvation. The data were measured with Mo  $K\alpha$  radiation.

The Wilson plot (Fig. 3) showed an anomaly for  $\rho > 0.35$  so the structure was re-refined excluding the high-angle data. A difference density map showed small ( $< 0.5 \text{ e } \text{\AA}^{-3}$ ) local maxima near chlorine. The solvent was excluded from the structure, and the main molecule was re-refined with HUGged

contributions. A Fourier map (Fig. 4) computed using only  $A_{hug}$  and  $B_{hug}$  clearly recovered the solvent with elongated distributions near chlorine in an otherwise featureless map.

The HUGged structure refined to a lower  $R$  factor and gave a Bijvoet( $d$ ) estimate of Flack's parameter similar to that from the fully atomic model refinement (Fig. 5).

### 3.4. Structure sgd475

This is an organic material containing C, H, F, O and N atoms (dr > 99:1), with disordered chloroform of solvation and measured with Cu  $K\alpha$  radiation. The difference map for the atomic modelled structure shows residual density. The normal probability plot for the weighted residuals  $w(F_o^2 - F_c^2)^2$ , had a slope of 1.36 and many outliers. The *SHELX*-type weighting coefficients (Cruickshank, 1961) were 0.049 and 0.873, and the  $D_s/\sigma(D_o)$  plot (Watkin & Cooper, 2016) was unusually skewed. The conventional  $R$  factor for the HUGged model is

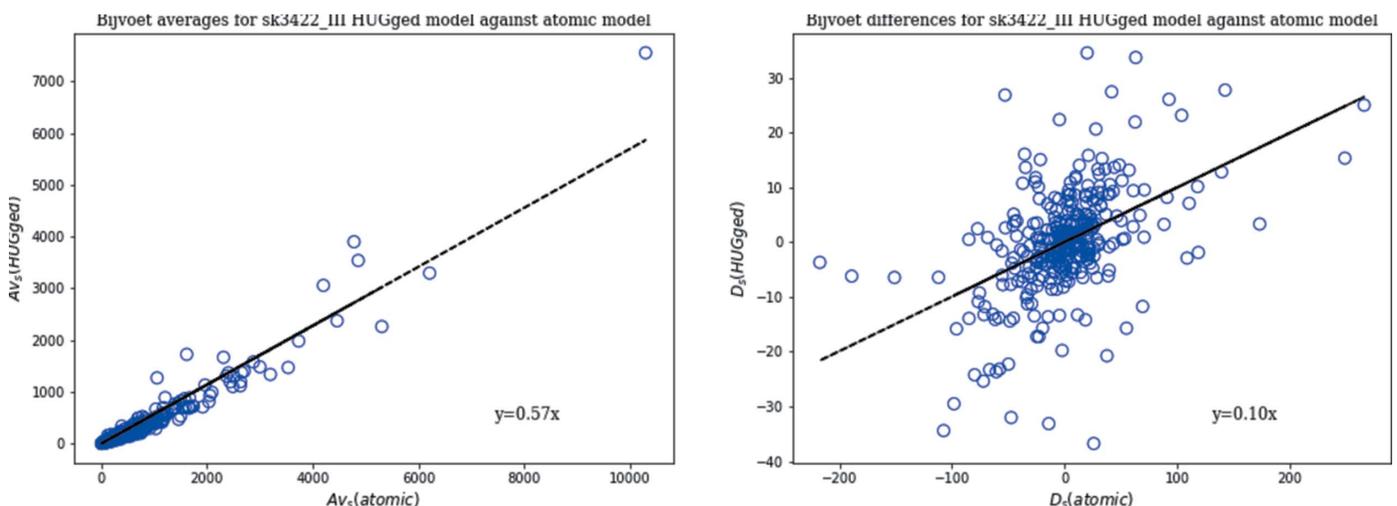


Figure 6

Averages ( $Av$ ) and differences ( $D$ ) of Bijvoet pairs for structure sk3422\_III, where the anion contribution was computed from an atomic model of the solvent and a HUGged model.

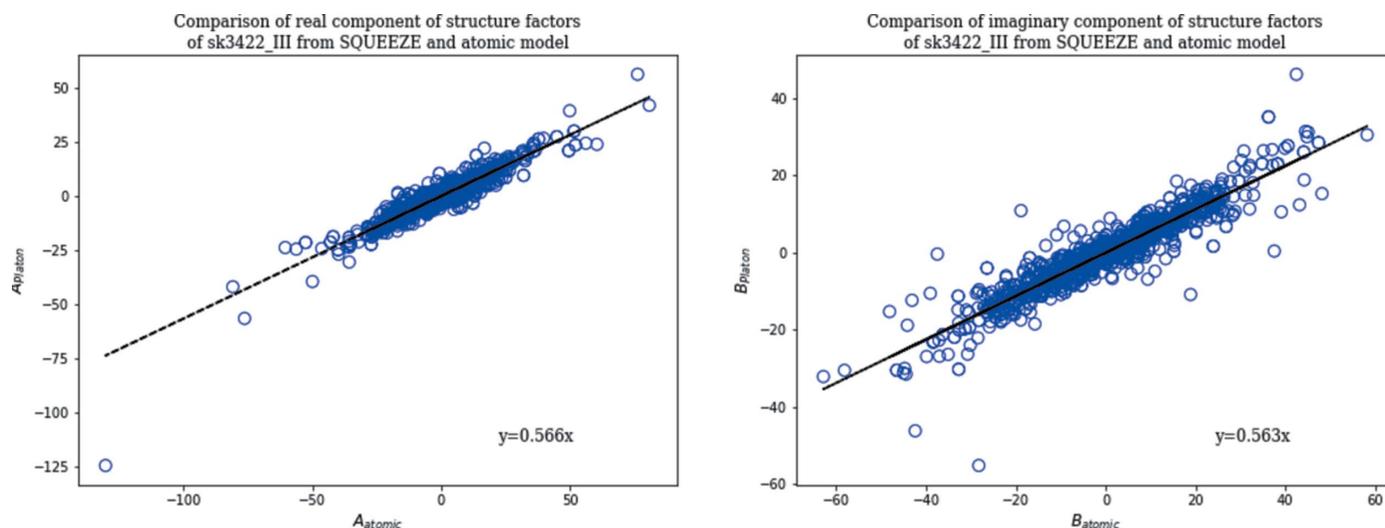


Figure 7

Plots of the  $A$  and  $B$  parts of the structure factor for the anion computed by SQUEEZE and from an atomic model. The contributions to the structure factor computed from the Fourier transformation of the difference electron density are underestimated by a factor of approximately one half. The data for the atomic model do not include any corrections for resonant scattering.

higher than that for the atomic model, but the Bijvoet( $d$ ) estimate of absolute structure is still reliably determined. The unmodified SQUEEZE map contains 61 electrons per void;  $R(Av_s) = 0.12$  and  $R(D_s) = 0.40$ .

### 3.5. Structure sk3422\_III

This is an organic salt consisting of C, H, N and O atoms, containing a cation and a disordered mixture of hydrogen phosphite ( $\text{H}_2\text{O}_3\text{P}^-$ ) and hydrogen fluorophosphate ( $\text{HF}\text{O}_3\text{P}^-$ ) anions in an 88:12 ratio. Although the Wilson plot looked normal, the  $N(z)$  plot contained bumpy deviations from the theoretical acentric curve. Refinement of the atomic

model gave a conventional  $R$  factor of 1.87% (*SHELX*-type weighing parameters of 0.032, 0.000). The refinement of the HUGged model was more problematic (conventional  $R = 13.8\%$ ). A *SHELX*-type weighting scheme could not be determined automatically and the parameters (0.40, 0.00) were set manually to get a roughly flat distribution of residuals. Not unsurprisingly, the absolute structure analysis of the HUGged data was also unsatisfactory. One possibility for these difficulties may have been failures in the interface between *PLATON* and *CRYSTALS*, but this is unlikely because the  $R1$  value computed from the SQUEEZEd data in *CRYSTALS* was 17%, comparable with a value of 17% computed by *PLATON* and 15% computed by *SHELXL* (Version 2014/7; Sheldrick, 2015).

The average of the Bijvoet pairs determined by HUGged SQUEEZE was approximately one-half of that determined

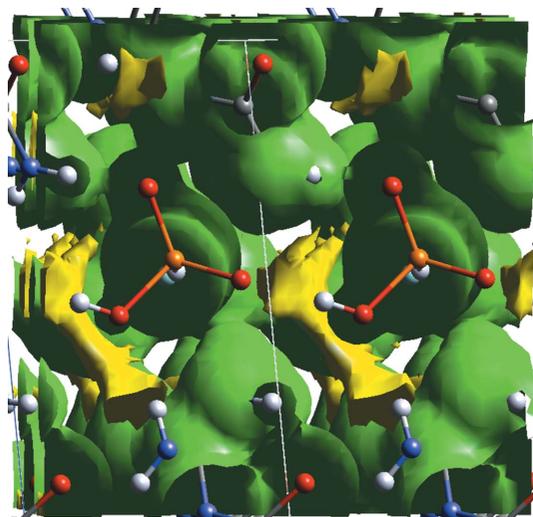


Figure 8

The space-filling contents of the unit cell of the fully resolved atomic model for structure sk3422\_III, viewed parallel to the fluorophosphate-hydrogen phosphite sheets. The green volumes are spheres centred on each atom, which are surrounded by transparent spheres, and any remaining volumes not enclosed by these invisible spheres are coloured yellow.

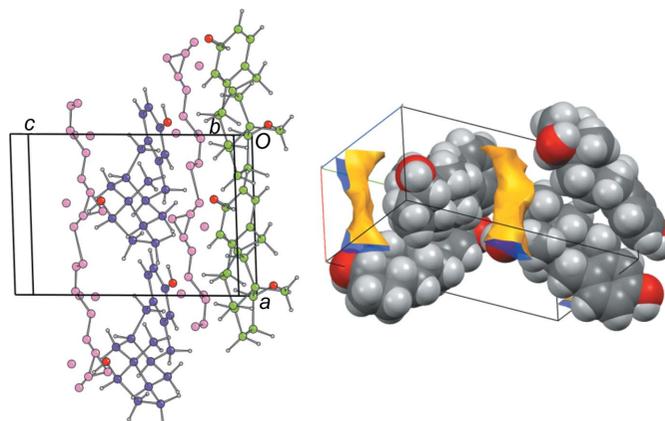
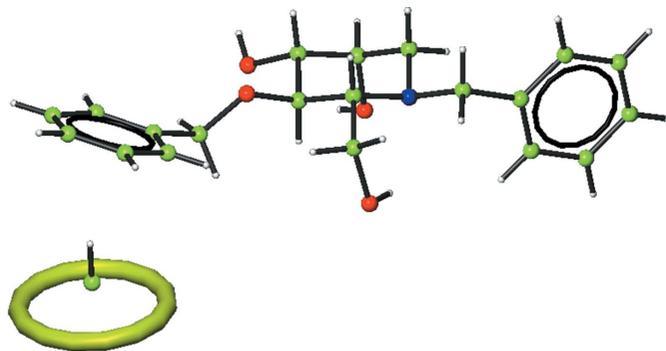


Figure 9

Left: *CAMERON* packing diagram (Pearce, 1995) of the ten strongest peaks in the *PLATON* SQUEEZE enhanced difference map (pink). The green and purple atoms are in the two independent steroid derivative molecules. The difference peaks form a continuous chain lying in a cylindrical void running through the crystal. Right: The void represents the solvent-accessible surface for a probe of radius 1.2 Å.

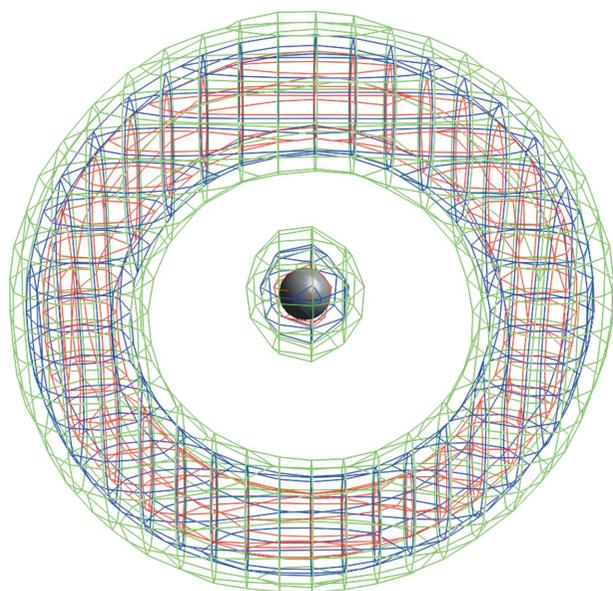


**Figure 10**

The modified structure of *sgd464*, showing the chloroform solvent at the bottom left of the image, with three individual Cl atoms replaced by a convolution of the scattering factors of three Cl atoms and a ring. The position, orientation and size of the ring were determined by fitting through the original Cl-atom positions.

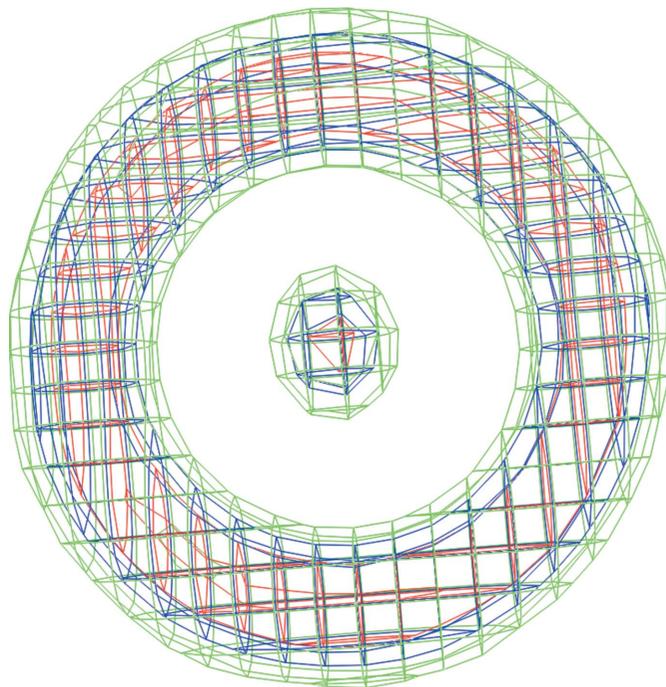
from an atomic model for the anion (Fig. 6). This led us to suspect that the root problem was the quality of the difference density maps. Plots of the *A* and *B* parts for the anion alone computed by SQUEEZE or an atomic model showed the same discrepancy (Fig. 7).

In *sgd240*, the ratio of the electron count of the atoms in the main moiety to that in the solvent was 11.6:1. In *sk3422\_III*, the ratio is only 1.3:1. The column headed  $(Z_h + Z_s)/Z_h$  in Table 1 shows that almost half of the total scattering is due to the anion, which could have an influence in the scaling, but clearly there is some other unidentified factor influencing the poor performance of *sk4322\_III*. The crystal structure consists of layers of hydrogen-bonded chains of fluorophosphonate–hydrogen phosphite sandwiched between layers of the organic cations. Fig. 8 shows the space-filling contents of the unit cell (MCE Version 2005 2.3.01) (Rohlíček & Hušák, 2007). The



**Figure 11**

For *sgd464*, the Fourier synthesis calculated from simulated data shows that the electron density is distributed as intended. Mesh contours are set at arbitrary positive values.



**Figure 12**

For *sgd464*, the Fourier synthesis computed from structure factors including the SQUEEZEd *A* and *B* parts is a close approximation to the original simulated data. Mesh contours are set at arbitrary positive values.

yellow regions represent inaccessible volumes in the structure, and it may be these which contribute to the problems.

### 3.6. Structure *awisac02*

This steroid derivative has a known absolute configuration. There are two molecules in the asymmetric unit, which are conformationally almost identical except for one hydroxy H atom, but not related by any approximate symmetry element. The original authors could not locate an atomic model for the included solvent and SQUEEZEd the residual electron density, interpreting the electron count for the solvent-accessible volume as a disordered molecule of methylene dichloride. Examination of the peaks found in the *PLATON* difference synthesis showed that the residual density formed a continuous chain in a channel through the structure (Fig. 9).

HUGging the SQUEEZE output for a single molecule of  $\text{CH}_2\text{Cl}_2$  reduced the conventional *R* factor to 6.56%. However, the electron count in the cell voids from *PLATON* ( $95 e^-$ ) is more than that for two single molecules ( $84 e^-$ ) of  $\text{CH}_2\text{Cl}_2$ . Multiplying *c* and *d* (equation 10) by 1.5 (*i.e.* three molecules of the solvent per unit cell) reduced the *R* factor to 5.92. The refined Flack(*x*) of 0.1 (2) and Bijvoet(*d*) of 0.33 (3) are on the correct side of 0.5, but are not convincing. The Hooft *P*(2) probability does not compute and the *P*(3) probability is strongly in favour of a twinned material.

### 3.7. Simulated diffuse solvent

During the review of this manuscript, one referee was interested to know how HUGging would perform if the strong resonant scatterers in the solvent were highly disordered. This

situation could be studied by altering the model of compound sgd464, which contains a well-ordered chloroform of solvation. In order to simulate a very disordered solvent, the three Cl atoms were replaced by an annular distribution equivalent to the three Cl atoms (Schröder *et al.*, 2004) (Fig. 10).

The Flack( $x$ ) parameter was set at 0.02 and the  $U_{\text{iso}}$  value for the ring set at  $0.06 \text{ \AA}^2$ . The structure factors computed from this model were treated as (error free) observations, but retaining the estimated standard uncertainties of the original data; the  $R$  factor was 0.03%. The Fourier synthesis calculated from this simulated data had the distributed electron density shown in Fig. 11.

The whole chloroform residue, including the C and H atoms, was deleted and the data SQUEEZED. The Fourier synthesis computed from structure factors including the squeezed  $A$  and  $B$  parts was a close approximation to the original synthesized data (Fig. 12). Refinement of the structure including the SQUEEZED  $A$  and  $B$  parts gave an  $R$  factor of 1.6% and a Flack( $x$ ) parameter of 2.9. HUGging and reweighting the SQUEEZED data gave an  $R$  factor of 1.7% and a Flack( $x$ ) parameter of 0.02 (3), admirably close to the value of 0.02 used in simulating the data. The Bijvoet( $d$ ) parameter was 0.02 (1).

#### 4. Conclusions

These preliminary observations show that an approximation to the resonant scattering can be computed for a disordered solvent molecule or counter-ion which may be adequate for the determination of absolute structure. It seems that the greatest chance of success occurs when the solvent/counter-ion has significant resonant scattering, but its real scattering must not overwhelm that of the host molecules.

Since the HUG algorithm is only a post-processing of the output from SQUEEZE, the success of the method is critically dependent on the applicability of SQUEEZE. The computation of  $c$  and  $d$  (equation 10) has no knowledge of the distribution in the voids of the strong resonant scatterers so that the HUG procedure can only be expected to be indicative of the absolute structure. Except for awisac02 and the simulated data above (§3.7), in the cases examined here, the solvent had been modelled by discrete atoms so that the target results were known. This will not be the case in real-life applications, but it seems that if the absolute configuration of an enantiopure all

light-atom material is required, it makes sense to attempt to recrystallize it from a solvent containing strong resonant scatterers, even if there is a likelihood that these may be incorporated as disordered solvent.

#### Funding information

Funding for this research was provided by: EPSRC (grant No. EP/K013009/1 to RIC).

#### References

- Bojarska, J., Maniukiewicz, W., Sieroń, L., Fruziński, A., Kopczacki, P., Walczyński, K. & Remko, M. (2012). *Acta Cryst.* **C68**, o341–o343.
- Carruthers, J. R. (1977). In *Proceedings of the 4th European Crystallographic Meeting (ECM-4)*, Oxford, UK, 30 August–3 September 1977. Abstract Ob. 2.
- Chernega, A. N., Davies, S. G., Goodwin, C. J., Hepworth, D., Kurosawa, W., Roberts, P. M. & Thomson, J. E. (2009). *Org. Lett.* **11**, 3254–3257.
- Cooper, R. I., Watkin, D. J. & Flack, H. D. (2016). *Acta Cryst.* **C72**, 261–267.
- Cruikshank, D. W. J. (1961). In *Computing Methods and the Phase Problem*, edited by R. Pepinsky, J. M. Robertson & J. C. Speakman, Paper No. 6. Oxford: Pergamon Press.
- Davies, S. G., Figuccia, A. L. A., Fletcher, A. M., Roberts, P. M. & Thomson, J. E. (2013a). *Org. Lett.* **15**, 2042–2045.
- Davies, S. G., Fletcher, A. M., Roberts, P. M., Thomson, J. E. & Zammit, C. M. (2013b). *Chem. Commun.* **49**, 7037–7039.
- Fábry, J., Fridrichová, M., Dušek, M., Fejfarová, K. & Krupková, R. (2012). *Acta Cryst.* **C68**, o76–o83.
- Flack, H. D. (1983). *Acta Cryst.* **A39**, 876–881.
- Flack, H. D. & Shmueli, U. (2007). *Acta Cryst.* **A63**, 257–265.
- Parsons, S., Flack, H. D. & Wagner, T. (2013). *Acta Cryst.* **B69**, 249–259.
- Pearce, L. J. (1995). PhD thesis, University of Oxford, England.
- Qian, M., Engler-Chiurazzi, E. B., Lewis, S. E., Rath, N. P., Simpkins, J. W. & Covey, D. F. (2016). *Org. Biomol. Chem.* **14**, 9790–9805.
- Rohlíček, J. & Hušák, M. (2007). *J. Appl. Cryst.* **40**, 600–601.
- Schröder, L., Watkin, D. J., Cousson, A., Cooper, R. I. & Paulus, W. (2004). *J. Appl. Cryst.* **37**, 545–550.
- Sheldrick, G. M. (2015). *Acta Cryst.* **C71**, 3–8.
- Shi, P., Zhang, L. & Ye, Q. (2012). *Acta Cryst.* **C68**, o266–o269.
- Sluis, P. van der & Spek, A. L. (1990). *Acta Cryst.* **A46**, 194–201.
- Spek, A. L. (2015). *Acta Cryst.* **C71**, 9–18.
- Watkin, D. & Schwarzenbach, D. (2017). *J. Appl. Cryst.* **50**, 666–667.
- Watkin, D. J., Carruthers, J. R. & Betteridge, P. W. (1985). *CRYSTALS User Guide*. Chemical Crystallography Laboratory, University of Oxford, England.
- Watkin, D. J. & Cooper, R. I. (2016). *Acta Cryst.* **B72**, 661–683.