Acta Crystallographica Section A
Foundations of Crystallography
ISSN 0108-7673

Received 15 November 2007
Accepted 10 January 2008

Minimally resolution biased electron-density maps

Angela Altomare,[a] Corrado Cuocci,[b] Carmelo Giacovazzo,[a,b]* Gihan Salah Kamel,[c] Anna Moliterni[a] and Rosanna Rizzi[a]

[a]IC, Bari, via Amendola 122/o, 70126 Bari, Italy, [b]Dipartimento Geomineralogico, Università di Bari, Campus Universitario, Via Orabona 4, 70125 Bari, Italy, and [c]Department of Physics, Faculty of Science, Helwan University, Cairo, Egypt. Correspondence e-mail: carmelo.giacovazzo@ic.cnr.it

Electron-density maps are calculated by Fourier syntheses with coefficients based on structure factors. Diffraction experiments provide intensities up to a limited resolution; as a consequence, the Fourier syntheses always show series-termination errors. The worse the resolution, the less accurate is the Fourier representation of the electron density. In general, each atomic peak is shifted from the correct position, shows a deformed (with respect to the true distribution of the electrons in the atomic domain) profile, and is surrounded by a series of negative and positive ripples of gradually decreasing amplitude. An algorithm is described which is able to reduce the resolution bias by relocating the peaks in more correct positions and by modifying the peak profile to better fit the real atomic electron densities. Some experimental tests are performed showing the usefulness of the procedure.

1. Notation

f_j scattering factor of the jth atom, thermal factor included
f_j^0 scattering factor of the jth atom at rest
N number of atoms in the unit cell
s $2\sin\theta/\lambda$
$F_{\mathbf{h}}$ structure factor with vectorial index $\mathbf{h}$
$E_{\mathbf{h}}$ normalized structure factor
B_j isotropic thermal parameter of the jth atom
RES experimental data resolution (in Å)

2. Introduction

Let

$$\rho(\mathbf{r}) = \sum_{j=1}^{N} \rho_j(\mathbf{r} - \mathbf{r}_j)$$

be the electron density of a crystal structure constituted by N atoms: $\mathbf{r}_j$ is the atomic position of the jth atom and $\rho_j(\mathbf{r})$ is its electron density. For X-rays, $\rho(\mathbf{r})$ is a non-negative definite function [*i.e.* $\rho(\mathbf{r}) \geq 0$ at any point of the unit cell]. When calculated *via* structure-factor amplitudes, $\rho(\mathbf{r})$ should coincide with

$$\rho'(\mathbf{r}) = \frac{1}{V}\sum_{\mathbf{h}} F_{\mathbf{h}} \exp(-2\pi i \mathbf{h}\cdot\mathbf{r})$$

provided the summation is extended to an infinite number of Miller indices and $\langle|F_{\mathbf{h}}|\rangle \to 0$ when RES $\to 0$. In practice, the summation is limited to the measured domain of the reciprocal space, represented by the shape function $\Phi(\mathbf{r}^*)$ [$\Phi(\mathbf{r}^*) = 1$ inside the measured domain, $\Phi(\mathbf{r}^*) = 0$ outside it]. Correspondingly, the electron-density map available in practice, say $\rho'(\mathbf{r})$, is based on the structure factors $F'_{\mathbf{h}} = F_{\mathbf{h}}\Phi(\mathbf{r}^*)$ rather than on $F_{\mathbf{h}}$, where

$$\rho'(\mathbf{r}) = \rho'(\mathbf{r}) \otimes T[\Phi(\mathbf{r}^*)] = \rho(\mathbf{r}) \otimes \zeta(\mathbf{r}) = \sum_{j=1}^{N} \rho_j(\mathbf{r} - \mathbf{r}_j) \otimes \zeta(\mathbf{r}). \quad (1)$$

$\zeta(\mathbf{r})$ is the Fourier transform of $\Phi(\mathbf{r}^*)$ and $\otimes$ represents the convolution operation.

The main features of $\rho'(\mathbf{r})$ are the following: (*a*) it can be negative in more or less extended regions of the unit cell; (*b*) the atomic peaks are broadened and are surrounded by a series of negative and positive ripples of gradually decreasing amplitude; (*c*) since ripples may overlap with other ripples and with atomic peaks, the maxima of the electron density $\rho'(\mathbf{r})$ are at $\mathbf{r}'_j$, not perfectly coinciding with the atomic positions $\mathbf{r}_j$; (*d*) the modification of the peak profiles and the displacements of the atomic peaks are resolution dependent; (*e*) $\rho'(\mathbf{r})$ will usually show more maxima and minima than $\rho(\mathbf{r})$; in favourable conditions, it may be approximated by the main N maxima in the map:

$$\rho'(\mathbf{r}) \approx \sum_{j=1}^{N} \rho'_j(\mathbf{r} - \mathbf{r}'_j), \quad (2)$$

where $\rho'_j(\mathbf{r})$, $j = 1, \ldots, N$, are suitable peak profiles.

Returning from $\rho'(\mathbf{r})$ to $\rho(\mathbf{r})$ *via* deconvolution procedures offers several advantages: peaks should move in more correct positions, false peaks due to the limited experimental data resolution may be eliminated, and peak broadening may be reduced.

It is a common belief that resolution bias in crystallographic electron-density maps is an unavoidable cost to pay: it is

generally considered an intrinsic unavoidable characteristic of the electron-density maps, generated by the physics of the diffraction experiment. This is rather unexpected if one considers that the resolution bias is precisely described by the basic diffraction theory. At non-atomic resolution, the bias is larger and any current attempt to interpret the electron-density maps in terms of molecular models is performed by using the prior information on the molecular geometry rather than by reducing the resolution bias. Correcting the bias should increase the efficiency of any cyclic procedure based on the electron-density modification, where the usefulness of the $(n+1)$th map is based on the quality of the phases derived from the inversion of the nth map. The current electron-density-modification procedures (Cowtan & Main, 1993; Abrahams, 1997; Cowtan, 1999; Terwilliger, 1999, 2003; Hunt & Deisenhofer, 2003) aim at fitting the expected characteristics of the electron-density function by suitable restraints in direct or reciprocal space. *E.g.* solvent flattening tries to fit the positivity of the map in direct space, histograms in one or more dimensions, in direct and/or reciprocal space, try to capture the stereochemical information and improve the phase estimates. None of these approaches is used to gain resolution and to relocate peaks.

Traditional Patterson search methods often suppress (in the reciprocal space) the ripples generated by the origin peak, by calculating $(|E|^2-1)$ Patterson maps. In contrast, the correction for the series-termination effects is usually not applied to observed electron-density maps. The only exception we know is reported in a recent paper by Burla *et al.* (2006), where an algorithm is described that allows the elimination of the bias in a region around the heavy-atom peak. The algorithm allowed the complexity limit for proteins solvable *ab initio* to extend from 2268 (value attained by Mooers & Matthews, 2006) to 6319 non-H atoms in the asymmetric unit (provided RES is better than 1.2 Å).

However, a perfect deconvolution procedure allowing the return from $\rho'(\mathbf{r})$ to $\rho(\mathbf{r})$ may be applied only if $\zeta(\mathbf{r})$ and the atomic positions $\mathbf{r}_j$ are *a priori* known. Luckily, the function $\zeta(\mathbf{r})$, *i.e.* the supplementary information which will be explicitly used by our algorithm, may be deduced from the experimental data; in contrast, we have experimental access only to the vectors $\mathbf{r}'_j$. In this paper, we describe a mathematical approach leading, from $\rho'(\mathbf{r})$, to a modified electron-density map [say $\rho'_{\rm mod}(\mathbf{r})$] which is expected to be closer to $\rho(\mathbf{r})$ than $\rho'(\mathbf{r})$.

We will first apply the approach to point-atom structures (see §§3–5). This case is approximately (*i.e.* the displacement factor blurs the exact theoretical shape of the scattering factor) met when neutron diffraction data are available. Point atoms are also simulated in X-ray or in electron crystallography when E maps are considered. Indeed, the normalization procedure replaces the usual atomic scattering factors f_j, monotonically decreasing with s, by their normalized scattering factors $\nu_j = f_j/(\sum_{j=1}^{N} f_j^2)^{1/2}$, which are nearly constant *versus* the diffraction angle. In this way, point-like atoms (the effects of the displacement factor is analysed in §5 and in Appendix *A*) replace electron clouds: the atomic peaks in the corresponding electron densities are sharper but the quality of the map is degraded by strong truncation effects (which are a function of RES).

In §5, the usual electron-density maps employed in X-ray crystallography will be considered.

In this pilot study, we will apply the theory to some selected cases to show that peaks in the $\rho'_{\rm mod}(\mathbf{r})$ maps are closer to the true atomic positions than peaks of the $\rho'(\mathbf{r})$ maps, and that the $\rho'_{\rm mod}$'s are more interpretable in terms of molecular models. The full potential of the theory will be revealed as soon as massive applications are made to X-ray, neutron and electron diffraction, single-crystal or powder data, small molecules as well as macromolecules, at different resolutions and when structure factors are still affected by phase errors. These applications are deferred to forthcoming papers.

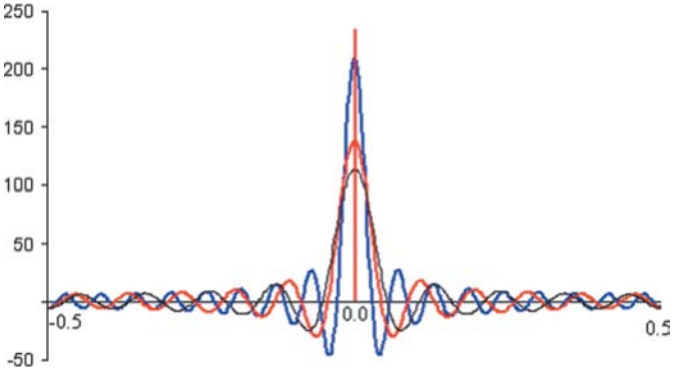

Figure 1
$\rho'(x)$ for a one-point-atom structure. The point atom is a delta function located at the origin [$\rho(x) = 6\delta(x)$] and represented by a red vertical bar. Blue line: RES = 1.2 Å; red line: RES = 1.8 Å; black line: RES = 2.2 Å.

3. Peak deconvolution for a 'one-point-atom' and for a 'two-point-atom' structure: one-dimensional case, atoms at rest

As a didactical example, let us first consider a one-dimensional unit cell with period a, containing only a 'one-point-atom', located at the origin: its electron density may be written as $\rho(x) = c\delta(x)$, where δ is the Dirac delta function and c is a suitable constant. If we calculate the structure factors up to RES (*i.e.* $F_{\mathbf{h}} = c$ for any $\mathbf{h}$) and, from them, the electron density $\rho'(x)$ by Fourier inversion, we should obtain

$$\rho'(x) = c\delta(x) \otimes \zeta(x) = c\zeta(x).$$

$\zeta(x)$ is equal to $[\sin(2h_{\rm max}+1)\pi x]/\sin \pi x$, the function theoretically expected to describe the effect of the resolution in the one-dimensional case, $h_{\rm max} = {\rm int}(a\,{\rm RES}^{-1})$ is the maximum index used in the electron-density calculation. The maximum value of $\zeta(x)$ is $(2h_{\rm max}+1)$, attained at the origin.

Suppose that we have observed the electron density $\rho'(x) = c\sin(2h_{\rm max}+1)\pi x/\sin \pi x$. Can we return back to the correct density $\rho(x)$ given $\rho'(x)$ and the prior knowledge that $\rho(x)$ is a delta function? In Fig. 1, we plot $\rho'(x)$ when $a = 20$ Å, $c = 6$ and RES = 1.2, 1.8 and 2.2 Å. While both $\rho'(x)$ and $\rho(x)$ have their maximum at the origin, relevant differences can be found between them: the main peak of ρ' is no longer a delta

function, positive and negative ripples, with gradually decreasing amplitude, surround it. The lower the resolution, the smaller are the ripple intensities and their frequency.

In this simple example, recovering $\rho(x)$ given $\rho'(x)$ is quite immediate. If the experimental fits with the function $c\sin(2h_{\max}+1)\pi x/\sin\pi x$, one can immediately deduce that $\rho(x)$ is a delta function centred on the origin.

It may be worthwhile noting that recovering $\rho(x)$ does not imply any peak shift; indeed, the main maximum in $\rho'(x)$ is located in the origin as is the delta function. That does not occur if a 'two-point-atom' structure is considered. Suppose the electron density $\rho'(x)$ is given by

$$\begin{aligned}\rho'(x) &= [c_1\delta(x-x_1)+c_2\delta(x-x_2)]\otimes\zeta(x)\\ &= c_1\zeta(x-x_1)+c_2\zeta(x-x_2).\end{aligned} \tag{3}$$

In Fig. 2(*a*), for a = 20 Å, $c_1 = c_2 = 6$, $x_1 = 0.0$, $x_2 = 0.15$, RES = 1.8 Å, we show the delta functions by two red vertical bars, $c_1\zeta(x-x_1)$ by a blue line, $c_2\zeta(x-x_2)$ by a green line. In Fig. 2(*b*), $\rho'(x)$ is plotted by a black line and is overlapped with

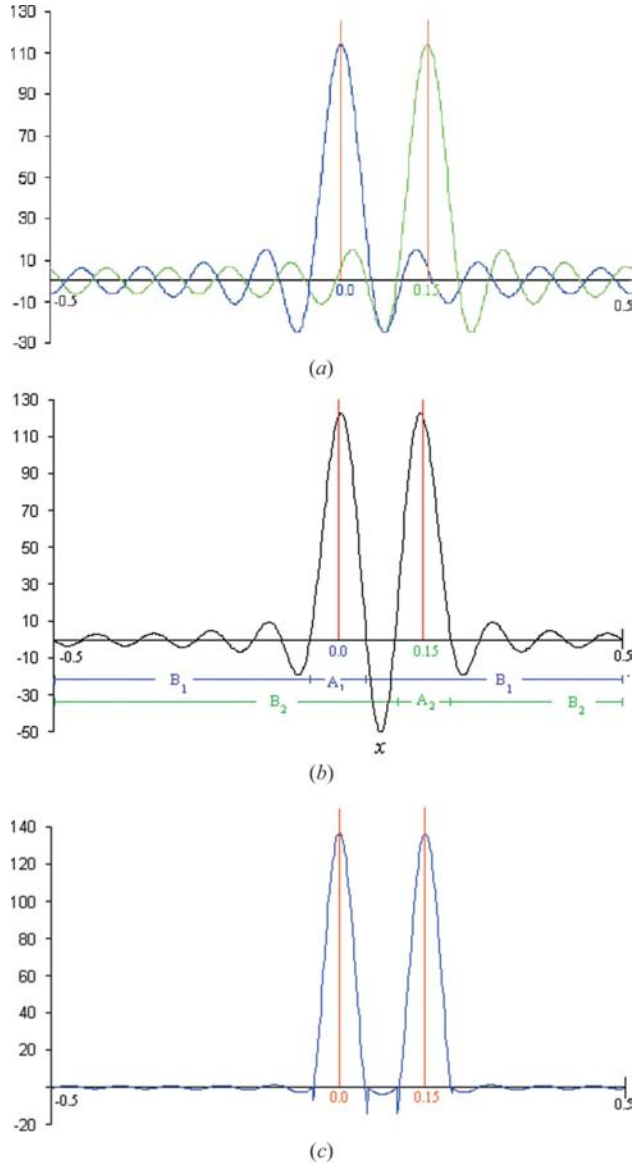

Figure 2
Two-point-atom structure: the two point atoms are located at $x_1 = 0.0, x_2 = 0.15$ and are represented by $6\delta(x-x_1)$ and $6\delta(x-x_2)$, respectively (red vertical bars). RES = 1.8 Å. (*a*) Blue line: $6\zeta(x-x_1)$; green line: $6\zeta(x-x_2)$. (*b*) $\rho'(x) = c_1\zeta(x-x_1)+c_2\zeta(x-x_2)$ (black line). The two main maxima of $\rho'(x)$ are at $x'_1 = 0.06$ and $x'_2 = 0.145$. (*c*) $\rho''(x)$ (blue line) and $\rho''_{\rm mod}(x)$ (red vertical bars).

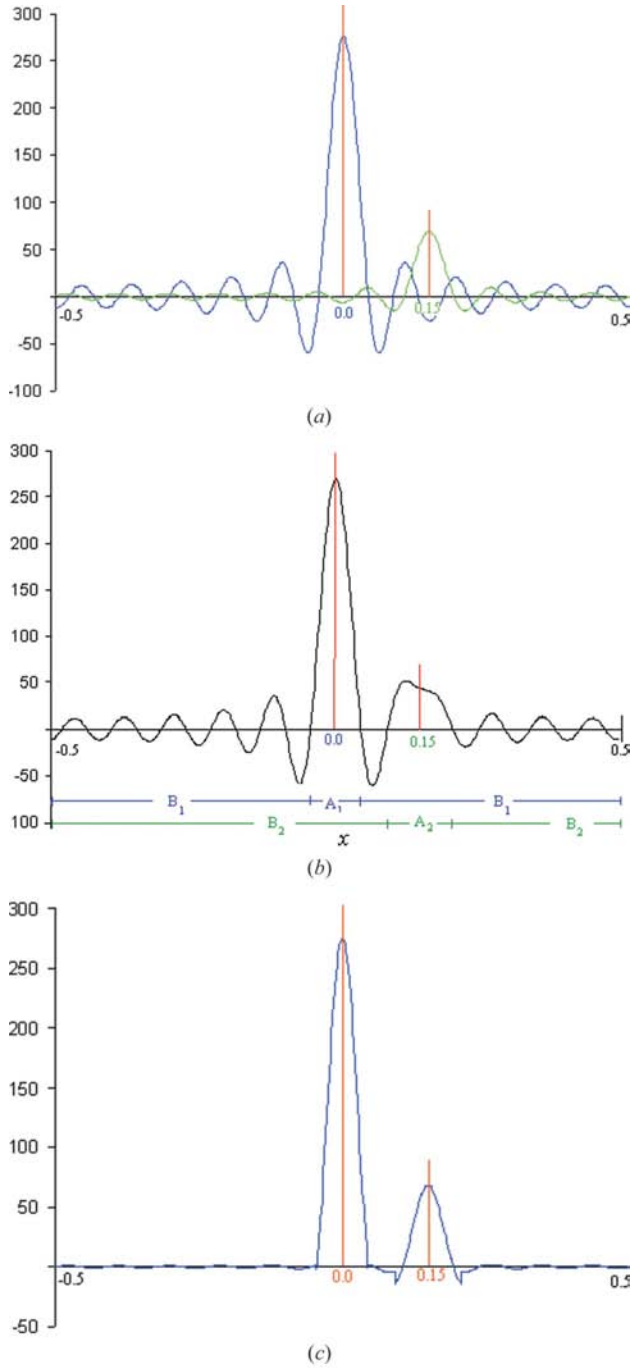

Figure 3
Two-point-atom structure: the two point atoms are located at $x_1 = 0.0$, $x_2 = 0.15$ and are represented by $12\delta(x-x_1)$ and $3\delta(x-x_2)$, respectively (red vertical bars). RES = 1.8 Å. (*a*) Blue line: $12\zeta(x-x_1)$; green line: $3\zeta(x-x_2)$. (*b*) $\rho'(x) = c_1\zeta(x-x_1)+c_2\zeta(x-x_2)$ (black line). The two main maxima of $\rho'(x)$ are at $x'_1 = 0.0$ and $x'_2 = 0.125$. (*c*) $\rho''(x)$ (blue line) and $\rho''_{\rm mod}(x)$ (red vertical bars).

the two vertical bars. The two main maxima of $\rho'(x)$ are at $x'_1 = 0.006$ and $x'_2 = 0.145$. The position misfit is equal to 0.12 Å for the first and 0.10 Å for the second atom. The experimental interatomic distance is therefore equal to 2.78 Å, against the true value of 3 Å. The misfit for the first atom depends on the specific interaction between the first atom peak profile and the ripples of the second atom: similar considerations hold for the second atom. In general, if a positive (negative) ripple is just under the peak, the position misfit vanishes but the amplitude of the peak increases (decreases). If the ripple is misplaced with respect to the peak profile, a position misfit occurs.

In Fig. 3(*a*), the same conditions chosen for Fig. 2 are used, but for $c_1 = 12$ and $c_2 = 3$. The two main maxima of $\rho'(x)$ are now [see Fig. 3(*b*)] at $x'_1 = 0.0$ and $x'_2 = 0.125$. The position misfit is equal to 0.0 Å for the heavier atom, 0.5 Å for the lighter; the experimental interatomic distance is equal to 2.50 Å, against the true value of 3 Å. The figure suggests that each peak generates ripples with amplitudes proportional to the amplitude of the main peak; accordingly, heavy atoms can produce greater deformation of the surrounding electron density than light atoms.

Recovering $\rho(x)$ given $\rho'(x)$ and the prior knowledge that $\rho(x)$ is the sum of two delta functions is now a problem a bit more complicated than for the 'one-point-atom' structure. Indeed, it is necessary to find the best fitting between $\rho'(x)$ and the sum of the two functions $c_1\zeta(x - x_1)$ and $c_2\zeta(x - x_2)$, by varying the parameters c_1, c_2, x_1 and x_2. Least-squares refinement is an ideal tool for solving the problem. Since the system is not linear, a starting model is necessary for securing convergence. We can approximate it by the expression

$$\rho'(x) \approx c'_1\zeta(x - x'_1) + c'_2\zeta(x - x'_2),$$

where

$$c'_1 = \rho'(x'_1)/\zeta(0) \quad \text{and} \quad c'_2 = \rho'(x'_2)/\zeta(0).$$

The function ζ and the parameters c'_1, c'_2, x'_1, x'_2 are experimentally available.

While structures with two delta functions can be easily managed *via* least-squares refinement, the problem becomes more complicated for '*n*-point-atom' structures with $n \geq 3$. Luckily, equation (3) suggests a simpler way for reducing the resolution bias, based on a more general idea of peak. In practice, the traditional concept of a Gaussian-like peak has to be replaced by a two-component function, extending over the full unit cell, and constituted by a *main peak* and by the corresponding ripples. Let $\rho'_j(x - x'_j)$ be such a function, $[A]_j = (x'_j - d, x'_j + d)$, $j = 1, 2$, be the full width of its main peak (see Figs. 2*b* and 3*b*), and $[B]_j$ the rest of the unit cell. According to equation (3), $\rho'_j(x - x'_j)$ may be split into two functions, $\rho'_{[A]_j}(x - x'_j)$ defined in $[A]_j$ and $\rho'_{[B]_j}(x - x'_j)$ defined in $[B]_j$. In accordance with this point of view, the experimental main peak $\rho'_{[A]_j}(x - x'_j)$ is the sum of the unknown function $c_j\zeta(x - x_j)$ and of the ripples of the other peak falling in the domain $[A]_j$. In its turn, $\rho'_{[B]_j}(x - x'_j)$ represents the ripples related to the *j*th main peak. In the formulas

$$\rho'_{[A]_1}(x - x'_1) \approx c_1\zeta(x - x_1) + \rho'_{[B]_2}(x - x'_2)$$
$$\rho'_{[A]_2}(x - x'_2) \approx c_2\zeta(x - x_2) + \rho'_{[B]_1}(x - x'_1),$$

the two equation establish the desired relation between the x'_j-based and the x_j-based models. Then more useful peak profiles may be recovered:

$$c_1\zeta(x - x_1) \approx \rho'_{[A]_1}(x - x'_1) - \rho'_{[B]_2}(x - x'_2) \qquad (4a)$$
$$c_2\zeta(x - x_2) \approx \rho'_{[A]_2}(x - x'_2) - \rho'_{[B]_1}(x - x'_1). \qquad (4b)$$

To apply equations (4), in practice we need to estimate $\rho'_{[B]_j}(x - x'_j)$ from the experimental data. Since both $\zeta(x)$ and x'_j are given by the experiment, we split $\zeta(x - x'_j)$ into two components:

$$\zeta(x - x'_j) = \zeta_{[A]_j}(x - x'_j) + \zeta_{[B]_j}(x - x'_j),$$

where

$$\zeta_{[A]_j}(x - x'_j) = \zeta(x - x'_j) \text{ in } [A]_j, \quad \zeta_{[A]_j}(x - x'_j) = 0 \text{ in } [B]_j,$$
$$\zeta_{[B]_j}(x - x'_j) = 0 \text{ in } [A]_j, \quad \zeta_{[B]_j}(x - x'_j) = \zeta(x - x'_j) \text{ in } [B]_j.$$

Then we can establish the relationship

$$\rho'_{[B]_j}(x - x'_j) \approx c'_j\zeta_{[B]_j}(x - x'_j),$$

with c'_j close to c_j.

In accordance with equations (4), the deconvolution algorithm may be described in two steps.

(I) The ripple contribution is subtracted from $\rho'(x)$, so giving rise to

$$\rho''(x) = \rho(x) - \rho'_{[B_1]}(x - x'_1) - \rho'_{[B_2]}(x - x'_2).$$

Suppose that the main maxima of the above function are centred on the new values x''_1 and x''_2, let

$$c''_1 = \rho''(x''_1)/\zeta(0) \quad \text{and} \quad c''_2 = \rho''(x''_2)/\zeta(0).$$

(II) In accordance with our prior knowledge, $\rho'_{\rm mod}$ is assumed to be

$$\rho'_{\rm mod} = c''_1\delta(x - x''_1) + c''_2\delta(x - x''_2).$$

$\rho''(x)$ and $\rho'_{\rm mod}(x)$ are plotted in Figs. 2(*c*) and 3(*c*). We obtain $x''_1 = 0.0$ and $x''_2 = 0.15$ for the case plotted in Fig. 2(*c*): the position misfit is equal to 0 Å for both point atoms (against the values of 0.12 and 0.10 Å, respectively, found before the application of the algorithm). For the case plotted in Fig. 3(*c*), $x''_1 = 0.0$ and $x''_2 = 0.15$: again no position misfit is detected for both atoms.

The values $c''_1 = c''_2 = 5.95$ for Fig. 2 and $c''_1 = 11.93$ and $c''_2 = 2.97$ for Fig. 3 are very close to c_1 and c_2. Owing to the approximations involved in the algorithm, some residual ripples remain in $\rho'_{\rm mod}(x)$ (see Figs. 2*c* and 3*c*): $\rho'_{\rm mod}(x)$ is however a more useful representation of $\rho(x)$ than $\rho'(x)$.

4. Peak deconvolution for three-dimensional point-atom structures: atoms at rest

Let us now consider a three-dimensional structure constituted by point atoms at rest located in the positions $\mathbf{r}_j$; its electron density is

$$\rho(\mathbf{r}) = \sum_{j=1}^{N} c_j \delta(\mathbf{r} - \mathbf{r}_j). \quad (5)$$

The experimental electron density will be

$$\rho'(\mathbf{r}) = \sum_{j=1}^{N} c_j \delta(\mathbf{r} - \mathbf{r}_j) \otimes \zeta(\mathbf{r}) = \sum_{j=1}^{N} c_j \zeta(\mathbf{r} - \mathbf{r}_j), \quad (6)$$

where $\zeta(\mathbf{r})$ may be obtained as follows:

(a) a point atom with weight $c = 1$ is put at the origin;

(b) it is assumed $F_{\mathbf{h}} = 1$ for any $\mathbf{h}$ belonging to the observed set of reflections;

(c) $\zeta(\mathbf{r}) = (1/V)\sum_{\mathbf{h}} F_{\mathbf{h}} \exp(-2\pi i \mathbf{h}\cdot\mathbf{r})$ is calculated where the summation goes over the set of observed reflections, Friedel-related included.

If Φ is a sphere, $\zeta(\mathbf{r})$ will have a spherical symmetry: the expected function is

$$\zeta(r) = \tfrac{4}{3}\pi r_{\max}^3 y, \quad (7)$$

where

$$y = 3\frac{\sin(2\pi r_{\max}^* r) - (2\pi r_{\max}^* r)\cos(2\pi r_{\max}^* r)}{(2\pi r_{\max}^* r)^3}$$

and $r_{\max}^* = \mathrm{RES}^{-1}$.

The maxima of $\rho'(\mathbf{r})$ will be located at positions $\mathbf{r}'_j$, usually not coincident with the positions $\mathbf{r}_j$. We want to find an electron density $\rho'_{\mathrm{mod}}(\mathbf{r})$ which better approximates $\rho(\mathbf{r})$, given $\rho'(\mathbf{r})$ and the prior knowledge that $\rho(\mathbf{r})$ is a sum of N delta functions. In accordance with §3, an approximate solution of the problem may be found by replacing the traditional concept of a three-dimensional Gaussian-like peak by a two-component function: *i.e.* $\rho'_j(\mathbf{r} - \mathbf{r}'_j)$ extends over the full unit cell and is constituted by a main peak and by the corresponding ripples. Let $[A]_j$ be the domain containing all the unit-cell points belonging to the jth main peak and $[B]_j$ the rest of the unit cell (see Fig. 4). According to §3, $\rho'_j(\mathbf{r} - \mathbf{r}'_j)$ is split into two functions, $\rho'_{[A]_j}(\mathbf{r} - \mathbf{r}'_j)$ defined in $[A]_j$, representing the jth main peak, and $\rho'_{[B]_j}(\mathbf{r} - \mathbf{r}'_j)$, defined in $[B]_j$ and describing its ripples. *I.e.*

$$\rho'_j(\mathbf{r} - \mathbf{r}'_j) = \rho'_{[A]_j}(\mathbf{r} - \mathbf{r}'_j) + \rho'_{[B]_j}(\mathbf{r} - \mathbf{r}'_j).$$

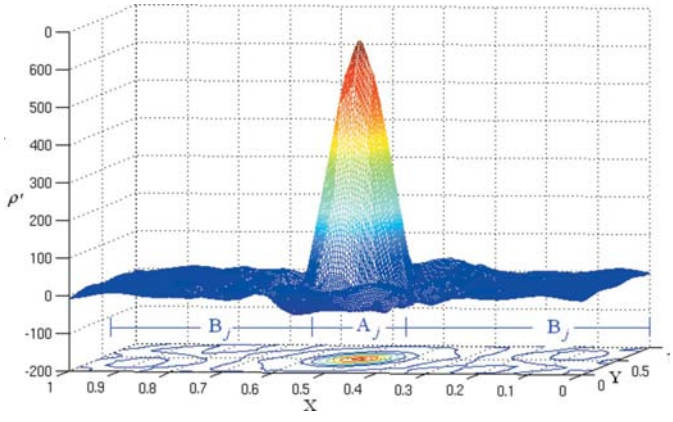

Figure 4
Point-atom structure: $\rho'(\mathbf{r})$ close to the jth peak located at (0.5, 0.5, 0.0). RES = 1.54 Å.

Accordingly, $\zeta(\mathbf{r} - \mathbf{r}'_j)$ is split into two components, *e.g.*

$$\zeta(\mathbf{r} - \mathbf{r}'_j) = \zeta_{[A]_j}(\mathbf{r} - \mathbf{r}'_j) + \zeta_{[B]_j}(\mathbf{r} - \mathbf{r}'_j),$$

where

$$\zeta_{[A]_j}(\mathbf{r} - \mathbf{r}'_j) = \zeta(\mathbf{r} - \mathbf{r}'_j) \text{ in } [A]_j, \quad \zeta_{[A]_j}(\mathbf{r} - \mathbf{r}'_j) = 0 \text{ in } [B]_j,$$
$$\zeta_{[B]_j}(\mathbf{r} - \mathbf{r}'_j) = 0 \text{ in } [A]_j, \quad \zeta_{[B]_j}(\mathbf{r} - \mathbf{r}'_j) = \zeta(\mathbf{r} - \mathbf{r}'_j) \text{ in } [B]_j,$$

from which we can establish the relationship

$$\rho'_{[B]_j}(\mathbf{r} - \mathbf{r}'_j) \approx c'_j \zeta_{[B]_j}(\mathbf{r} - \mathbf{r}'_j). \quad (8)$$

$c'_j = \rho'(\mathbf{r}'_j)/\zeta(0)$ is expected to be close to c_j. Equation (8) estimates $\rho'_{[B]_j}(\mathbf{r} - \mathbf{r}'_j)$ in terms of experimental quantities.

We can now consider the experimental main peak $\rho'_{[A]_j}(\mathbf{r} - \mathbf{r}'_j)$ as the sum of the unknown function $c_j\zeta(\mathbf{r} - \mathbf{r}_j)$ and of the ripples of the other peaks falling in the domain $[A]_j$:

$$\rho'_{[A]_j}(\mathbf{r} - \mathbf{r}'_j) \approx c_j\zeta(\mathbf{r} - \mathbf{r}_j) + \sum_{i\neq j=1}^{N} \rho'_{[B]_i}(\mathbf{r} - \mathbf{r}'_i).$$

Since $\rho'_{[B]_j}(\mathbf{r} - \mathbf{r}'_j)$ vanishes in $[A]_j$, we can rewrite the above equation in the form

$$\rho'_{[A]_j}(\mathbf{r} - \mathbf{r}'_j) \approx c_j\zeta(\mathbf{r} - \mathbf{r}_j) + \sum_{i=1}^{N} \rho'_{[B]_i}(\mathbf{r} - \mathbf{r}'_i),$$

from which

$$c_j\zeta(\mathbf{r} - \mathbf{r}_j) \approx \rho'_{[A]_j}(\mathbf{r} - \mathbf{r}'_j) - \sum_{i=1}^{N} \rho'_{[B]_i}(\mathbf{r} - \mathbf{r}'_i)$$

may be derived. The algorithm may now be described in two steps.

(I) the overall ripple contribution is subtracted from $\rho'(\mathbf{r})$, so giving rise to the electron-density map

$$\rho''(\mathbf{r}) = \rho'(\mathbf{r}) - \sum_{j=1}^{N} \rho'_{[B]_j}(\mathbf{r} - \mathbf{r}'_j) \approx \rho'(\mathbf{r}) - c'_j\zeta_{[B]_j}(\mathbf{r} - \mathbf{r}'_j).$$

(II) Let the main maxima of $\rho(\mathbf{r})$ be centred on the values $\mathbf{r}''_j$. Then ρ'_{mod} is assumed to be

$$\rho'_{\mathrm{mod}} = \sum_{j=1}^{N} c''_j \delta(\mathbf{r} - \mathbf{r}''_j),$$

where $c''_j = \rho(\mathbf{r}''_j)/\zeta(0)$.

To check the usefulness of the above algorithm, we simulated in $P1$ a simple point-atoms crystal structure constituted by a benzene ring in a unit cell with parameters a = 3.848, b = 11.005, c = 12.727 Å, α = 106, β = 95, $\gamma = 80^\circ$, RES = 1.54 Å. All the atoms were located in the plane z = 0: in Figs. 5(a), (b), we show the true atomic positions by red balls. In the same figures, the positions of the largest six maxima of $\rho'(\mathbf{r})$ are shown by blue balls: Fig. 5(a) is plotted perpendicular to the z axis, Fig. 5(b) is in the ring plane. If we apply the two-step algorithm described above, we obtain the new positions $\mathbf{r}''_j$, drawn in Fig. 6(a) (perpendicular to $\mathbf{z}$) and in Fig. 6(b) (in the ring plane) by blue balls. One $\rho'(\mathbf{r})$ peak is far from the correct position and is remarkably out from the benzene plane: its position is well recovered in ρ'_{mod}, which is also able to maintain the planarity of the ring.

In Table 1, we give the intensities of the six largest peaks in $\rho'(\mathbf{r})$ and their distances from the corresponding true atomic positions. In the last columns of Table 1, we show the corresponding figures for the $\rho''(\mathbf{r})$ map. In the last row, the corresponding average distances are given. The improvement is quite remarkable.

5. The peak deconvolution procedure for a three-dimensional crystal structure

Consider first a point-atom structure, each atom of which is affected by thermal vibration. Its electron density may be represented by

$$\rho(\mathbf{r}) = \sum_{j=1}^{N} c_j \delta(\mathbf{r} - \mathbf{r}_j) \otimes t_j(\mathbf{r}) = \sum_{j=1}^{N} c_j t_j(\mathbf{r} - \mathbf{r}_j).$$

$t_j(\mathbf{r})$ represents the thermal vibration of the jth atom and is in general anisotropic: in our calculations, it will be approximated by a simpler Gaussian function with spherical symmetry. Then the structure factor may be written as

$$F_{\mathbf{h}} = \sum_{j=1}^{N} c_{Tj} \exp(2\pi i \mathbf{h} \cdot \mathbf{r}_j),$$

where

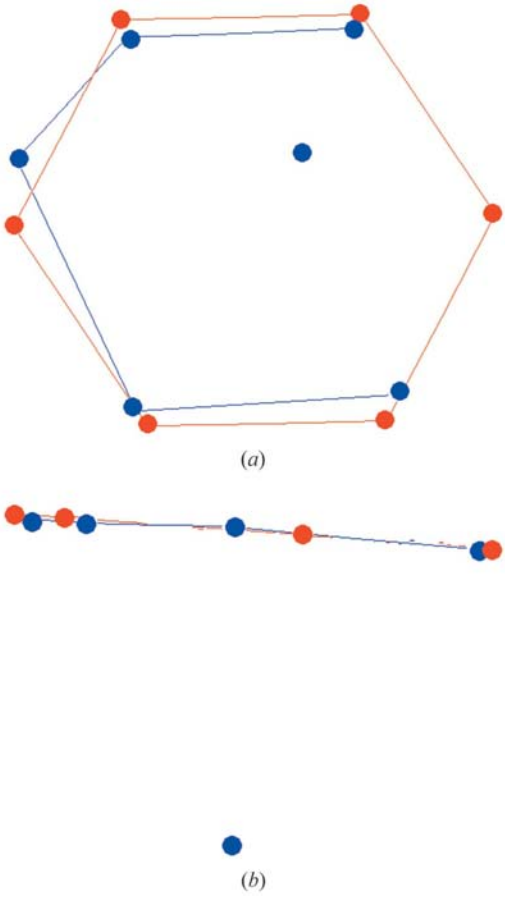

Figure 5
Red balls: atomic positions of the simulated structure. Blue balls: the largest peak positions in $\rho'(\mathbf{r})$; (*a*) view perpendicular to the z axis, (*b*) view in plane of the ring.

Table 1
For each jth atom of the six-atom simulated structure, d is the distance between the true atom position and the corresponding peak in $\rho'(\mathbf{r})$, int is its peak intensity; $d_{\rm mod}$ and $\rm int_{mod}$ are the corresponding values in $\rho'_{\rm mod}(\mathbf{r})$.

In the last row, the corresponding average distances are given.

j	d (Å)	int	$d_{\rm mod}$ (Å)	$\rm int_{mod}$
1	0.109	529	0.281	609
2	0.106	525	0.217	607
3	0.165	499	0.071	576
4	0.114	497	0.214	566
5	0.404	465	0.076	535
6	1.105	114	0.285	522
	$\langle d \rangle$ = 0.33		$\langle d \rangle$ = 0.19	

$$c_{Tj} = c_j \exp(-B_j s^2)$$

decreases with s. The experimental electron density $\rho'(\mathbf{r})$ will be

$$\rho'(\mathbf{r}) = \sum_{j=1}^{N} c_j t_j(\mathbf{r} - \mathbf{r}_j) \otimes \zeta(\mathbf{r}) = \sum_{j=1}^{N} c_j q_j(\mathbf{r} - \mathbf{r}_j),$$

where

$$q_j(\mathbf{r}) = t_j(\mathbf{r}) \otimes \zeta(\mathbf{r}).$$

The maxima of $\rho'(\mathbf{r})$ will be located at positions $\mathbf{r}'_j$ not coincident with the positions $\mathbf{r}_j$.

It may be observed that the thermal motion transforms $\rho(\mathbf{r})$ from a sum of δ functions into an electron-density sum of extended functions t_j. Therefore, this case may be considered as a special case of a more general situation, *i.e.* when $\rho(\mathbf{r})$ is the sum of extended atomic electron densities. Accordingly, we will not treat this case further; we only add, in Appendix *A*,

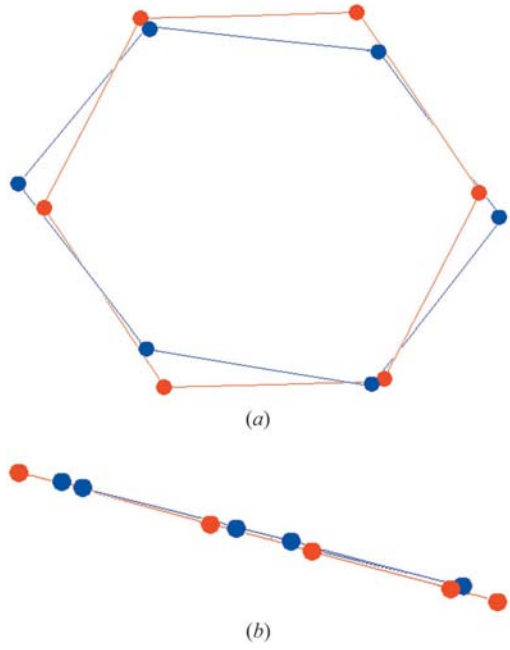

Figure 6
Red balls: atomic positions of the simulated structure. Blue balls: the largest peak positions in $\rho''(\mathbf{r})$; (*a*) view perpendicular to the z axis; (*b*) view in the plane of the ring.

some considerations concerning the applicability of our algorithm to usual point-atom structures described by X-ray E maps or to neutron diffraction data.

Let us now consider a three-dimensional crystal structure constituted by atoms located at $\mathbf{r}_j$, $j = 1, \ldots, N$. Its electron density is

$$\rho(\mathbf{r}) = \sum_{j=1}^{N} \rho_j(\mathbf{r} - \mathbf{r}_j)$$

and the observed electron density is

$$\rho'(\mathbf{r}) = \sum_{j=1}^{N} \rho_j(\mathbf{r} - \mathbf{r}_j) \otimes \zeta(\mathbf{r}) = \sum_{j=1}^{N} q_j(\mathbf{r} - \mathbf{r}_j). \quad (9)$$

The position of the jth main peak in $\rho'(\mathbf{r})$, say $\rho'_j(\mathbf{r} - \mathbf{r}'_j)$, is influenced by the superposition of the functions

$$q_i(\mathbf{r} - \mathbf{r}_i), \quad i = 1, \ldots, N, \ i \neq j. \quad (10)$$

In accordance with §§3–4, returning back to $\rho_j(\mathbf{r} - \mathbf{r}_j)$ requires the elimination of the ripples of the main peaks and a suitable modification of the mean peak profiles. That may be attempted as soon as the following points are clarified.

(a) In the case of real atoms, each peak of $\rho'(\mathbf{r})$ arises from the convolution of two broad functions, the first of which (say ρ_j) is atom dependent, the second (say ζ) is resolution dependent. As a consequence, each main peak will generally be wider than both the corresponding ρ_j and the main peak of $\zeta(\mathbf{r})$. Furthermore, each main peak may show, at a given resolution, a proper broadness, depending on the atomic species; in this case, the practical application of the deconvolution procedure would require $[A]$ and $[B]$ domains of different sizes, each size specific for each atomic species (in contrast, for point atoms at rest the $[A]$ size was constant).

(b) The location of the ripples depends on the resolution; if it was influenced also by the thermal factor of the peak-related atom, then the functions $\zeta(\mathbf{r})$ and $q_j(\mathbf{r})$ would show, at a given data resolution, a different peak location. In this case, the deconvolution procedure should face additional difficulties; indeed, the atomic species associated with a main peak and the corresponding atomic thermal factor often are not *a priori* known. Correspondingly, the specific function $q_j(\mathbf{r})$ to associate to the jth atom would remain ambiguous.

For assessing point (a), we plot in Fig. 7, for a = 20 Å and RES = 1.8 Å, the electron density of C, Mn, Br and Hg atoms when located at the origin and at rest. The Fourier coefficients of such a simple $\rho'(\mathbf{r})$ function are

$$F_{\mathbf{h}} = f^0(s).$$

It may be observed that the domain $[A]$ does not substantially vary with the atomic species, the ripple intensity is proportional to Z, the ripple frequency and location remain practically constant.

For assessing point (b), we plot in Fig. 8(a), for a = 20 Å and RES = 1.8 Å, the electron density of the C atom for different values of the thermal factor B. We observe: (i) the main peak intensity is inversely related to the B value; (ii) the size of the domain $[A]$ increases with B; (iii) the ripple intensities are inversely related to B but their frequency and location are roughly constant (a mathematical basis for this practical invariance is given in Appendix B); (iv) when the thermal

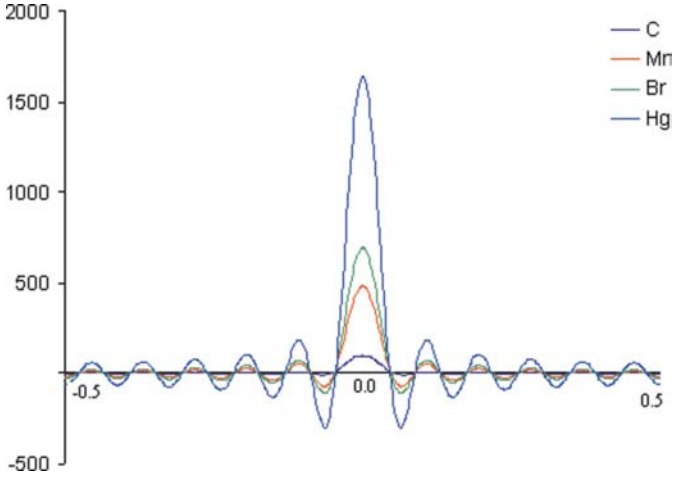

Figure 7
One-atom structure. $\rho'(x)$ when C, Mn, Br and Hg atoms, at rest, are located at the origin. RES = 1.8 Å.

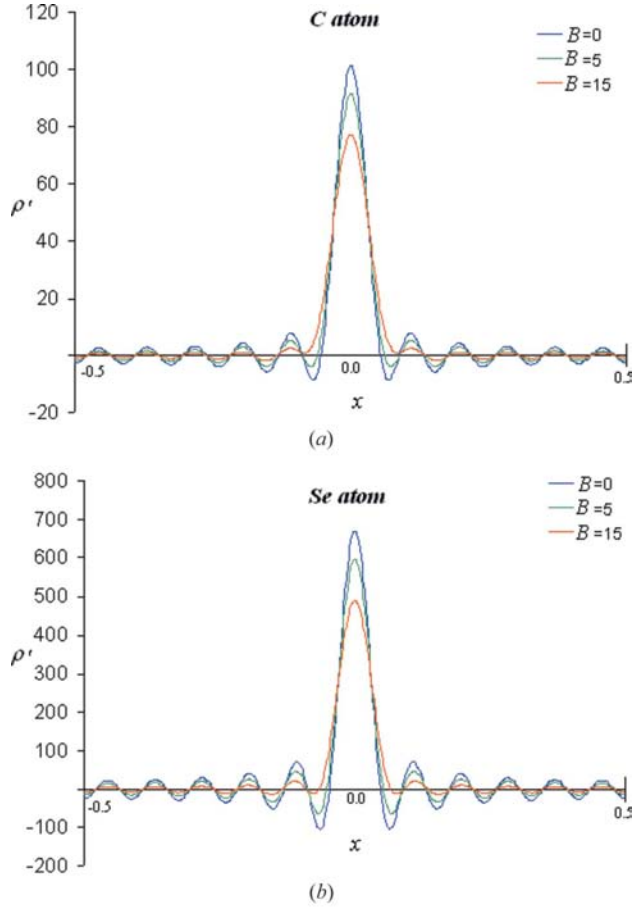

Figure 8
One-atom structure. $\rho'(x)$ at RES = 1.8 Å: (a) when the C atom is located at the origin with different values of the thermal factor B (blue: B = 0 Å^2; green: B = 5 Å^2; red: B = 15 Å^2); (b) the corresponding curves when the Se atom is located on the origin.

factor becomes larger than a given threshold, the limits of the main maximum do not coincide with the zero points of the function q. A similar behaviour may be found for atoms with lower and larger Z values: the B threshold for which the main maximum does not coincide with the zero point of the function q is directly related to Z (see Fig. 8*a*).

In accordance with the above results, the definitions of the $[A]_j$ and $[B]_j$ domains must be generalized: indeed, the definition according to which $[A]_j$ contains all the points belonging to the main jth peak is no longer self-evident. A practically working definition may be the following: for each jth peak $\rho'_j(\mathbf{r}-\mathbf{r}'_j)$, centred on $\mathbf{r}'_j$, a specific domain $[A]_j$ is defined, which is the collection of the points $\mathbf{r}$ for which $|\mathbf{r}-\mathbf{r}'_j| < d_m$, where d_m is the distance of the first minimum of $\rho'_j(\mathbf{r}-\mathbf{r}'_j)$ from $\mathbf{r}'_j$ (see Fig. 8*a* for C atoms). The domain $[B]_j$ is the rest of the unit cell.

If the $\rho'_j(\mathbf{r}-\mathbf{r}'_j)$ zero points are at a distance $d_0 < d_m$ from $\mathbf{r}'_j$ (see Fig. 8*b* for Se atoms), then $[A]_j$ is defined as the collection of points $\mathbf{r}$ for which $|\mathbf{r}-\mathbf{r}'_j| \leq d_0$. The domain $[B]_j$ is the rest of the unit cell.

The above definition may be immediately applied to isotropic main peaks: it may also be applied to anisotropic peaks if we replace $\rho'_j(\mathbf{r}-\mathbf{r}'_j)$ by $\langle\rho'_j(\mathbf{r}-\mathbf{r}'_j)\rangle$, where the average is calculated over spherical shells around the maximum of each main peak.

In accordance with §§3–4, we can represent $\rho'_j(\mathbf{r}-\mathbf{r}'_j)$ as the sum of two functions:

$$\rho'_j(\mathbf{r}-\mathbf{r}'_j) = \rho'_{[A]_j}(\mathbf{r}-\mathbf{r}'_j) + \rho'_{[B]_j}(\mathbf{r}-\mathbf{r}'_j), \tag{11}$$

where $\rho'_{[A]_j}(\mathbf{r}-\mathbf{r}'_j)$ corresponds to the jth main peak and $\rho'_{[B]_j}(\mathbf{r}-\mathbf{r}'_j)$ describes the related ripples. By analogy, we will divide $q_j(\mathbf{r}-\mathbf{r}_j)$ into $q_{[A]_j}(\mathbf{r}-\mathbf{r}_j)$ and $q_{[B]_j}(\mathbf{r}-\mathbf{r}_j)$. In addition, the function $\zeta(\mathbf{r})$, experimentally available by using a unit point atom located at the origin, is decomposed into the sum of two functions:

$$\zeta(\mathbf{r}) = \zeta_{[A]_\zeta}(\mathbf{r}) + \zeta_{[B]_\zeta}(\mathbf{r}),$$

where

$$\zeta_{[A]_\zeta}(\mathbf{r}) = \zeta(\mathbf{r}) \text{ for } \mathbf{r}\in[A]_\zeta, \quad \zeta_{[A]_\zeta}(\mathbf{r}) = 0 \text{ for } \mathbf{r}\in[B]_\zeta,$$
$$\zeta_{[B]_\zeta}(\mathbf{r}) = \zeta(\mathbf{r}) \text{ for } \mathbf{r}\in[B]_\zeta, \quad \zeta_{[B]_\zeta}(\mathbf{r}) = 0 \text{ for } \mathbf{r}\in[A]_\zeta.$$

$[A]_\zeta$ is the collection of points $\mathbf{r}$ for which $r \leq d_\zeta$, where d_ζ is the distance from the origin to the $\zeta(\mathbf{r})$ zero points. $[B]_\zeta$ is the rest of the unit cell. The above relationships allow us to approximate the experimental main peak $\rho'_{[A]}(\mathbf{r}-\mathbf{r}'_j)$ as the sum, in $[A]_j$, of the unknown function $q_{[A]_j}(\mathbf{r}-\mathbf{r}_j)$ and of the ripples of the other peaks falling in the domain $[A]_j$:

$$\rho'_{[A]_j}(\mathbf{r}-\mathbf{r}'_j) \approx q_{[A]_j}(\mathbf{r}-\mathbf{r}_j) + \sum_{i=1}^{N}\rho'_{[B]_i}(\mathbf{r}-\mathbf{r}'_i),$$

from which the unknown function

$$q_{[A]_j}(\mathbf{r}-\mathbf{r}_j) \approx \rho'_{[A]_j}(\mathbf{r}-\mathbf{r}'_j) - \sum_{i=1}^{N}\rho'_{[B]_i}(\mathbf{r}-\mathbf{r}'_i) \tag{12}$$

may be derived. While $\rho'_{[A]_j}(\mathbf{r}-\mathbf{r}'_j)$ is given by the experiment, the function $\rho'_{[B]_i}(\mathbf{r}-\mathbf{r}'_i)$ is still unknown for each i. Its estimation, necessary for the application of equation (12), may be obtained as follows.

(i) One atom (for example, of the most frequent species in the unit cell) is located in the origin. Its thermal factor may coincide with the overall thermal factor obtained by a Wilson plot.

(ii) The structure factors up to RES are calculated and, from them, the electron density is obtained.

(iii) Its origin peak $\psi(\mathbf{r})$ is divided into two parts, corresponding to the domains $[A]$ and $[B]$, respectively: $\psi(\mathbf{r}) = \psi_{[A]}(\mathbf{r}) + \psi_{[B]}(\mathbf{r})$.

(iv) Each function $\rho'_{[B]_i}(\mathbf{r}-\mathbf{r}'_i)$ is calculated according to the following formula:

$$\rho'_{[B]_i}(\mathbf{r}-\mathbf{r}'_i) = d_i\psi_{[B]}(\mathbf{r}-\mathbf{r}'_i),$$

where

$$d_i = \rho'_{[A]_i}(r'_i)/\psi_{[A]}(\mathbf{r}).$$

The algorithm for recovering $\rho'_{\rm mod}(\mathbf{r})$ is now described in three steps.

(*a*) We calculate

$$\rho''(\mathbf{r}) = \rho'(\mathbf{r}) - \sum_{j=1}^{N}\rho'_{[B]_j}(\mathbf{r}-\mathbf{r}'_j), \tag{13}$$

the main maxima of which are now centred at $\mathbf{r}''_j$.

(*b*) Each peak $\rho''_{[A]_j}(\mathbf{r}-\mathbf{r}''_j)$ may be considered a convolution, *i.e.*

$$\rho''_{[A]_j}(\mathbf{r}-\mathbf{r}''_j) \approx \rho_j(\mathbf{r}-\mathbf{r}''_j)\otimes\zeta_{[A]_\zeta}(\mathbf{r}). \tag{14}$$

In equation (14), we neglect the contribution of $\zeta_{[B]_\zeta}$ to the convolution. Let $c''_jG(\mathbf{r};\sigma''_j,\mathbf{r}''_j)$ be the Gaussian function best fitting $\rho''_{[A]_j}(\mathbf{r}-\mathbf{r}''_j)$, let $c_\zeta G(\mathbf{r};\sigma_\zeta,0)$ be the Gaussian function best fitting $\zeta_{[A]_\zeta}(\mathbf{r})$ in $[A]_\zeta$ $\{G(\mathbf{r};\sigma_\zeta,0)$ is set to zero in the domain $[B]_\zeta\}$ and let $c_jG(\mathbf{r};\sigma_j,\mathbf{r}''_j)$ be the Gaussian function best fitting $\rho_j(\mathbf{r}-\mathbf{r}''_j)$ in the domain $[A]_j$. While $\rho_j(\mathbf{r}-\mathbf{r}''_j)$, c_j and σ_j are at the moment unknown, the parameters c''_j, σ''_j, c_ζ and σ_ζ can be derived from the experimental data. Since the convolution of two Gaussian functions is a Gaussian function, we can establish the relation

$$c''_j = c_jc_\zeta, \quad \sigma''^2_j = \sigma^2_\zeta + \sigma^2_j.$$

Conversely, the unknown (and desired) parameters

$$c_j = c''_j/c_\zeta, \quad \sigma^2_j = \sigma''^2_j - \sigma^2_\zeta$$

may be obtained.

(*c*) We assume

$$\rho'_{{\rm mod}_j}(\mathbf{r}-\mathbf{r}''_j) \approx c_jG(\mathbf{r};\sigma_j,\mathbf{r}''_j)$$

and

$$\rho'_{\rm mod}(\mathbf{r}) \approx \sum_{j=1}^{N}c_jG(\mathbf{r};\sigma_j,\mathbf{r}''_j). \tag{15}$$

We expect that $\rho'_{\rm mod}(\mathbf{r})$ is a better approximation of $\rho(\mathbf{r})$ than $\rho'(\mathbf{r})$.

6. A minimally resolution biased Fourier synthesis

So far we have described an algorithm which modifies the experimental electron-density map to obtain a new map in which the resolution bias has been minimized. It may be worthwhile assessing if the modified map may be obtained *via* a Fourier synthesis using modified (RES-dependent) structure factors as coefficients. Let us rewrite equation (13) as follows:

$$\begin{aligned}\rho''(\mathbf{r}) &= \sum_{j=1}^{N} \rho_j''(\mathbf{r}-\mathbf{r}_j'') \\ &\approx \sum_{j=1}^{N} [\rho_j'(\mathbf{r}-\mathbf{r}_j') - \rho_{[B]_j}'(\mathbf{r}-\mathbf{r}_j')] \\ &\approx \sum_{j=1}^{N} [\rho_j'(\mathbf{r}-\mathbf{r}_j') - c_j'\zeta_{[B]_\zeta}(\mathbf{r}-\mathbf{r}_j')]. \end{aligned} \tag{16}$$

The inverse Fourier transform of equation (16), say

$$F_{\mathbf{h}}'' = F_{\mathbf{h}} - F_{[B]_{\zeta\mathbf{h}}},$$

should provide the coefficients of the minimally biased electron-density synthesis, where

$$F_{\mathbf{h}}'' = T^{-1}[\rho(\mathbf{r})], \tag{17a}$$

$$F_{\mathbf{h}} = T^{-1}[\rho'(\mathbf{r})] = T^{-1}\Big[\sum_{j=1}^{N} \rho_j'(\mathbf{r}-\mathbf{r}_j')\Big] = \sum_{j=1}^{N} f_j \exp(2\pi i \mathbf{h}\cdot\mathbf{r}_j) \tag{17b}$$

$$F_{[B]_\zeta} = T^{-1}[c_j'\zeta_{[B]_\zeta}(\mathbf{r}-\mathbf{r}_j')] = \sum_{j=1}^{N} c_j' f_{[B]_\zeta} \exp(2\pi i \mathbf{h}\cdot\mathbf{r}_j), \tag{17c}$$

$$f_{[B]_\zeta} = T^{-1}[\zeta_{[B]_\zeta}].$$

While equations (17*a*) and (17*b*) are not ambiguous (the first is a definition, the second is supported by our previous analysis), equation (17*c*) is doubtful. Indeed, $\zeta_{[B]_\zeta}$ has a profile quite different from the atomic profiles ρ_j so that it is not guaranteed that the function $\sum_{j=1}^{N} \rho_{[B]_j}'(\mathbf{r}-\mathbf{r}_j')$ may be reproduced by using pseudo-atoms in $\mathbf{r}_j$. Even in the most favourable case that it cancels the ripples generated by the atoms, it should introduce its own ripples. Since the problem requires a deeper analysis, we prefer to apply the algorithm described in §5 to the practical case examined in §7.

It may be worthwhile noticing that the application of our algorithm allows the extension of the resolution. The more $\rho_{\rm mod}'(\mathbf{r})$ approximates $\rho(\mathbf{r})$, the more the structure factors, calculated from $\rho_{\rm mod}'(\mathbf{r})$ beyond RES, will approximate the true ones: in the ideal case in which $\rho_{\rm mod}'(\mathbf{r}) \equiv \rho(\mathbf{r})$, extrapolation up to RES = 0 is allowed. The practical use of the extrapolation feature will be the object of our next study.

7. A first application to a real case

We decided to apply the algorithm to a particularly difficult case, in which the resolution bias adds its effects to non-negligible phase errors (*e.g.* arising from the application of direct methods) and to errors in the structure-factor moduli (*e.g.* the unavoidable consequence of the full-pattern-decomposition process in powder crystallography). We choose the crystal structure CAINE (Nowell *et al.*, 2002), $P\bar{1}$ space group, a = 7.40, b = 8.57, c = 13.69 Å, α = 106.2, β = 90.8, γ = 98.7°; chemical content in the asymmetric unit: $C_{15}H_{25}Cl_1N_2O_2$, RES = 1.46 Å; number of observed reflections in the observed 2θ range NREFL = 546. Only powder diffraction data are experimentally available and they were used in the applications described below. All the calculations were performed by a modified version of the program *EXPO2004* (Altomare *et al.*, 2004). The structure-factor moduli were estimated *via* the Le Bail method (Le Bail *et al.*, 1988); owing to various physical factors (*e.g.* peak overlapping, difficulty of defining the background *etc.*), the process ended with the value $R_{\rm mod}$ = 0.38, where $R_{\rm mod}$ is the classical crystallographic residual between true and Le Bail-estimated structure-factor moduli.

The phase problem was faced by applying direct methods: the best trial solution showed an average phase error for the NREFL reflections equal to $\langle\Delta\phi\rangle$ = 51°. The map $\rho'(\mathbf{r})$ is just the structure-factor Fourier transform. Peaks in $\rho'(\mathbf{r})$ were automatically located and their connectivity established by the standard *EXPO2004* routines. The map was very poor and was not interpretable, essentially because the peaks were far from their correct positions. Only NP = 10 peaks were at a distance less than 0.6 Å from the true atomic positions: their average shift was $\langle d\rangle$ = 0.26 Å. The results are shown in Fig. 9 where the symmetry-independent molecule, correctly oriented and positioned according to the published structure, is drawn in black and the NP peaks shown as green balls (their connectivity is given by green lines). Any attempt by *EXPO2004* to solve the structure by cycles of electron-density calculations and least-squares refinement failed.

The $\rho_{\rm mod}'(\mathbf{r})$ map was then calculated to remove the resolution bias in $\rho'(\mathbf{r})$. Now the standard peak-search routines of *EXPO2004* located NP = 18 peaks at a distance less than 0.6 Å from the true atomic positions: their average shift was $\langle d\rangle$ = 0.32 Å. Peak positions and connectivity are shown in red in Fig. 9. The impressive improvement of map quality is also testified by the average phase error $\langle\Delta\phi\rangle$ = 34°, obtained by Fourier inversion of $\rho_{\rm mod}'(\mathbf{r})$. The final effect is that the $\rho_{\rm mod}'(\mathbf{r})$ map is perfectly interpretable. At the end of the automatic procedure, *EXPO2004* completed the structure (all 20 atoms were correctly located with averaged distance $\langle d\rangle$ = 0.23 Å and phase error $\langle\Delta\phi\rangle$ = 24°).

Some final considerations are useful to describe the limits of the procedure.

(I) The large initial phase error ($\langle\Delta\phi\rangle$ = 51°) did not allow all the atoms to appear in the first electron-density map: indeed, of the $1.3N_{\rm asym}$ peaks involved in the procedure by the *EXPO2004* peak-search routines, only NP = 10 peaks were at a distance less than 0.6 Å from the correct atomic positions. $N_{\rm asym}$ is the number of non-H atoms in the asymmetric unit ($N_{\rm asym}$ = 20 in this case). The resolution bias correction moved 18 peaks to a distance less than 0.6 Å from the correct atomic positions, thus allowing the improvement of the molecular model. As a consequence, in a cyclic procedure, the phase error of the calculated structure factors diminished and therefore the next electron-density map improved. It may be concluded that the success of the algorithm does not need all

the peaks to be close to the correct atomic positions: good and wrong peaks can initially coexist without causing the failure of the procedure.

(II) The value of RES for our test structure is 1.46 Å. One of the conditions for the success of the algorithm is that atoms separated by distances in the interval 1.4–1.5 Å (typical C—C distances) give rise to electron-density peaks which may be split in single peaks. It is guessed that structures with RES = 1.6–1.7 Å are within the limits of the algorithm.

8. Conclusions

We have described an algorithm which curtails the resolution bias in electron-density maps. The algorithm has been successfully checked in some simulated and in one real case; dramatic improvements of the experimental maps have been obtained. We are not able to assess its full potential: extensive tests are necessary. It is expected that the method may be applied both during the phasing procedure for accelerating the phasing process and in the refinement steps to both small and macromolecules, to neutron as well as to electron and X-ray data, and for variable data resolution. It seems also able to extend the resolution. The eventual success of the next applications will make available a new powerful tool for modern crystallography, for which resolution is still a quite limiting factor.

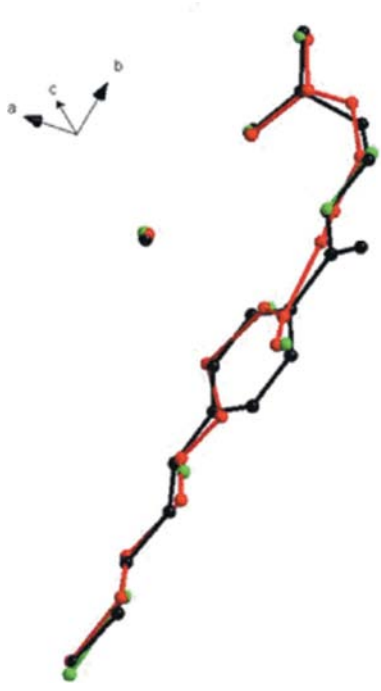

Figure 9
CAINE structure. In black: true atomic positions and connectivity. The Cl atom corresponds to the isolated atom. In green: the positions of the ten peaks in $\rho'(\mathbf{r})$ which are close to the true atomic positions. Only two pairs of atoms are connected by acceptable bond distances; the electron-density map is not interpretable. In red: the peak positions in $\rho'_{\rm mod}(\mathbf{r})$. The map is interpretable.

APPENDIX *A*

Useful approximations to point-atom structures are the so-called E maps: they are obtained by using the normalized structure factors

$$E_{\mathbf{h}} = F_{\mathbf{h}} \Big/ \left(\sum_{j=1}^{N} f_j^2 \right)^{1/2} \qquad (18)$$

as coefficients of the Fourier synthesis. The atomic scattering factor f_j includes the effect of the thermal vibration:

$$f_j = f_j^0 \exp(-B_j s^2/4).$$

The use of (18) replaces physical atoms by point atoms but does not completely eliminate the effects of the temperature factors. Indeed, in the absence of any prior information, it is usual to employ in the normalization process the overall thermal factor provided by the Wilson plot, *i.e.*

$$f_j = f_j^0 \exp(-\bar{B} s^2/4).$$

Then the structure factor may be modelled by

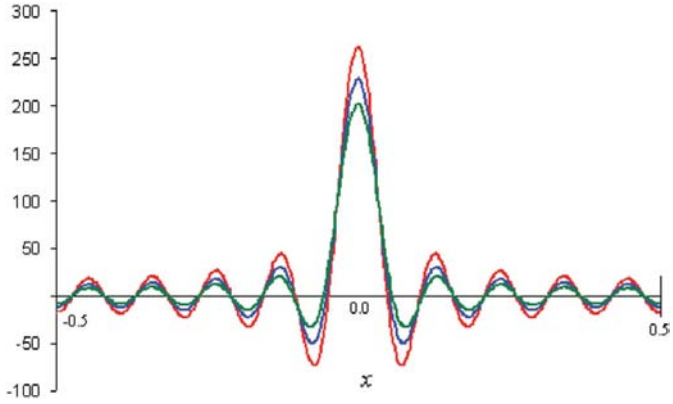

Figure 10
The function $c^0 \exp -[(B-\bar{B})s^2]\zeta(x)$ for $c^0 = 12$, a = 20 Å and RES = 1.8 Å. Red line: $B-\bar{B} = -7$. Blue line: $B-\bar{B} = 0$. Green line: $B-\bar{B} = 7$. In all the cases, the main peak may be approximated by a Gaussian function.

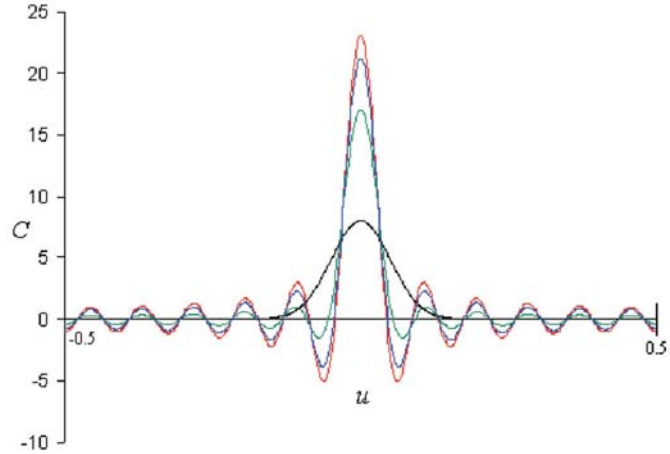

Figure 11
The function $C(u)$ is plotted: $\sigma = 0$ (red line); $\sigma = 0.05$ (blue line); $\sigma = 0.1$ (green line); $\sigma = 0.2$ (black line);

$$F_{\mathbf{h}} = \sum_{j=1}^{N} c_j \exp(2\pi i \mathbf{h}\cdot\mathbf{r}_j),$$

where

$$c_j = c_j^0 \exp -[(B_j - \bar{B})s^2/4] \qquad (19)$$

and

$$c_j^0 = \frac{f_j^0}{(\sum_{j=1}^{N} f_j^0)^{1/2}}.$$

While c_j^0 may be assumed to be nearly constant with s, c_j shows a different behaviour; it decreases or increases with s according to whether $B_j > \bar{B}$ or $B_j < \bar{B}$. The second case has no physical counterpart; indeed, the global scattering power of the point atom is unlimited, increasing with s. A necessary and sufficient condition for the applicability of our deconvolution technique is (see §5) that the function

$$\rho_j'(\mathbf{r} - \mathbf{r}_j) = c\zeta(\mathbf{r} - \mathbf{r}_j)$$

may be reliably approximated by a suitable Gaussian function $G_\zeta(\mathbf{r})$ in the domain $[A]$ also when c is defined by equation (19) with $B < \bar{B}$. In Fig. 10, we show (red line) $c\zeta(x)$, with c defined by equation (19), for $c^0 = 12$, $B - \bar{B} = -7$, $a = 20$ Å and RES = 1.8 Å. The peak may be approximated by a Gaussian function, sharper than that corresponding to $B = \bar{B}$ (blue line). The case $B > \bar{B}$ (*i.e.* $B - \bar{B} = 7$) is shown as a green line; the peak is broader than the first two.

APPENDIX *B*

In §5, we have obtained the following expression for the experimental electron density:

$$\rho'(\mathbf{r}) = \sum_{j=1}^{N} c_j \rho_j(\mathbf{r} - \mathbf{r}_j) \otimes \zeta(\mathbf{r}) = \sum_{j=1}^{N} c_j q_j(\mathbf{r} - \mathbf{r}_j).$$

The question is: if $\rho(\mathbf{r})$ is an atomic electron density, are the frequency and the location of the ripples defined by the function $C(u) = \rho(\mathbf{r}) \otimes \zeta(\mathbf{r})$ substantially different from those defined by the function $\zeta(\mathbf{r})$? In the main text, we answered the question by experimentally verifying that the frequency and location of the two functions are similar; we give here a mathematical support to the statement. For simplicity, we make our calculations in the one-dimensional space by calculating the convolution

$$C(u) = \rho(x) \otimes \zeta(x) = \int_{-\infty}^{+\infty} \rho(x)\zeta(u - x)\,\mathrm{d}x,$$

where

$$\rho(x) = c(2\pi\sigma^2)^{-1/2} \exp(-x^2/2\sigma^2),$$

$$\zeta(x) = \frac{\sin N\pi x}{\sin \pi x} = \sum_{h=-h_{\max}}^{h_{\max}} \cos 2\pi h x$$

and

$$N = (2h_{\max} + 1).$$

Then

$$C(u) = c\frac{1}{\sqrt{2\pi}\sigma} \sum_{h=-h_{\max}}^{h_{\max}} \cos 2\pi h u \int_{-\infty}^{+\infty} \exp\Big(-\frac{x^2}{2\sigma^2}\Big)\cos 2\pi h x\,\mathrm{d}x$$

$$= c \sum_{h=-h_{\max}}^{h_{\max}} w_h \cos 2\pi h u,$$

where

$$w_h = \exp(-2\pi^2 h^2 \sigma^2).$$

If $\rho(x)$ is a delta function then $\sigma^2 = 0$, $w_h = 1$ and $C \equiv \zeta$, as expected. If $\sigma^2 \neq 0$ then $w_h < 1$ for any $h \neq 0$; the ripple intensities are expected to decrease with σ. In contrast, the frequency and location of minima and maxima of the function $C(u)$ do not substantially change; indeed, the convolution between a Gaussian and an oscillating function reflects the distances between the centre of the Gaussian function (in our case the origin) and the minima and maxima of the oscillating function. Such distances do not substantially depend on the standard deviation of the Gaussian. The main differences between two $C(u)$ functions corresponding to two Gaussians with different σ values are expected to be close to the origin (more specifically, inside and close to the domain $[A]$) because a larger (smaller) $\rho(x)$ standard deviation implies a larger (smaller) standard deviation of the $C(u)$ origin peak (see Fig. 11). This region is not critical for the application of the algorithm described in §5.

References

Abrahams, J. P. (1997). *Acta Cryst.* D**53**, 371–376.
Altomare, A., Caliandro, R., Camalli, M., Cuocci, C., Giacovazzo, C., Moliterni, A. G. G. & Rizzi, R. (2004). *J. Appl. Cryst.* **37**, 1025–1028.
Burla, M. C., Caliandro, R., Carrozzini, B., Cascarano, G. L., De Caro, L., Giacovazzo, C., Polidori, G. & Siliqi, D. (2006). *J. Appl. Cryst.* **39**, 728–734.
Cowtan, K. (1999). *Acta Cryst.* D**55**, 1555–1567.
Cowtan, K. D. & Main, P. (1993). *Acta Cryst.* D**49**, 148–157.
Hunt, J. F. & Deisenhofer, J. (2003). *Acta Cryst.* D**59**, 214–224.
Le Bail, A., Duroy, H. & Fourquet, J. L. (1988). *Math. Res. Bull.* **23**, 447–452.
Mooers, B. H. M. & Matthews, B. W. (2006). *Acta Cryst.* D**62**, 165–176.
Nowell, H., Attfield, J. P., Cole, J. C., Cox, P. J., Shankland, K., Maginn, S. J. & Motherwell, W. D. S. (2002). *New J. Chem.* **26**, 469–472.
Terwilliger, T. C. (1999). *Acta Cryst.* D**55**, 1863–1871.
Terwilliger, T. C. (2003). *Acta Cryst.* D**59**, 1688–1701.