

On the use of the *C* map in Patterson deconvolution procedures

Rocco Caliandro,^a Benedetta Carrozzini,^a Giovanni Luca Cascarano,^a Giuliana Comunale^b and Carmelo Giacovazzo^{a*}

^aIstituto di Cristallografia, CNR, Via G. Amendola 122/O, 70126, Bari, Italy, and ^bDipartimento di Chimica, Università della Basilicata, 85100, Potenza, Italy. Correspondence e-mail: carmelo.giacovazzo@ic.cnr.it

The cross-correlation function between the target and a model electron density, denoted as the *C* map, has been crystallographically characterized. In particular, a study of its interatomic vectors and of their relation with the Patterson vectors has been undertaken. Since the *C* map is not available during the phasing process, the *C'* map, its centric modification, is considered. It may be computed at any stage of the phasing process and shows properties that are very useful for the crystal structure determination process. It has been combined with the implication transformation method and with vector-superposition techniques for performing the Patterson deconvolution and obtaining an initial model for dual-space recycling. While Patterson methods are traditionally considered to be more efficient for structures containing heavy atoms, the *C* map extends their potential to light-atom structures (*i.e.* containing atoms not heavier than O).

© 2013 International Union of Crystallography
Printed in Singapore – all rights reserved

1. Notation

N, N_p : number of atoms in the unit cell of the target structure (the one we want to phase) and of the model structure, respectively.

F, E, F_p, E_p : structure factor and normalized structure factors of the target and of the model structure, respectively.

$\sum_N = \sum_{j=1}^N f_j^2, \sum_{N_p} = \sum_{j=1}^{N_p} f_j^2$: f_j is the scattering factor of the j th atom, thermal factor included.

$I_i(x)$: modified Bessel function of order i .

$D = \langle \cos(2\pi\mathbf{h}\Delta\mathbf{r}) \rangle$: the average is performed per resolution shell.

$\sigma_A = D(\sum_{N_p}/\sum_N)^{1/2}$.

$\mathbf{C} \equiv (\mathbf{R}, \mathbf{T})$: symmetry operator. \mathbf{R} and \mathbf{T} are the rotation and the translation matrix, respectively.

2. Introduction

In a recent paper (Carrozzini *et al.*, 2010) the cross-correlation function $C(\mathbf{r})$ was crystallographically characterized. We briefly recall its definition. Given two functions $f(x)$ and $g(x)$, their cross-correlation function is defined by

$$C(y) = f(x) \otimes g(x) = \int_{-\infty}^{+\infty} f^*(x)g(x+y) dx, \quad (1)$$

where the star indicates the complex conjugate. The cross correlation is both associative and distributive but not commutative: owing to the classical Wiener–Kinchin theorem, it satisfies the relations

$$T[f(x) \otimes g(x)] = [Tf(x)]^*[Tg(x)] \quad (2)$$

and

$$T[g(x) \otimes f(x)] = [Tg(x)]^*[Tf(x)]. \quad (3)$$

The right-hand sides of equations (2) and (3) are complex conjugates.

The above definition was applied to the following two functions: the electron density,

$$\rho(\mathbf{r}) = \sum_{j=1}^N \rho_j(\mathbf{r} - \mathbf{r}_j) = (1/V) \sum_{\mathbf{h}} F_{\mathbf{h}} \exp(-2\pi i \mathbf{h} \cdot \mathbf{r}), \quad (4)$$

(from now on referred to as the density of the target structure, because we are interested in knowing this), and the model density (presumed to be known),

$$\rho_p(\mathbf{r}) = \sum_{j=1}^{N_p} \rho_{p_j}(\mathbf{r} - \mathbf{r}'_j) = (1/V) \sum_{\mathbf{h}} F_{p\mathbf{h}} \exp(-2\pi i \mathbf{h} \cdot \mathbf{r}). \quad (5)$$

ρ_j and ρ_{p_j} are atomic electron densities, the first centered on \mathbf{r}_j , the second centered on $\mathbf{r}'_j = \mathbf{r}_j + \Delta\mathbf{r}_j$. If N_p/N is sufficiently high and the $\Delta\mathbf{r}_j$'s are sufficiently small, the two functions $\rho(\mathbf{r})$ and $\rho_p(\mathbf{r})$ are highly correlated with each other. The target and model structure are assumed to show the same space-group symmetry.

According to equation (1) we have

$$C(\mathbf{u}) = \rho(\mathbf{r}) \otimes \rho_p(\mathbf{r}) = \int_{\mathbf{s}} \rho(\mathbf{r})\rho_p(\mathbf{r} + \mathbf{u}) d\mathbf{r} \quad (6)$$

and, in accordance with relations (2) and (3),

$$C(\mathbf{u}) = (1/V) \sum_{\mathbf{h}} |F_{\mathbf{h}} F_{p\mathbf{h}}| \exp i(\varphi_{\mathbf{h}} - \varphi_{p\mathbf{h}}) \exp(-2\pi i \mathbf{h} \cdot \mathbf{u}), \quad (7)$$

where $F_{\mathbf{h}} = |F_{\mathbf{h}}| \exp(i\varphi_{\mathbf{h}})$ and $F_{p\mathbf{h}} = |F_{p\mathbf{h}}| \exp(i\varphi_{p\mathbf{h}})$ are the structure factors of $\rho(\mathbf{r})$ and $\rho_p(\mathbf{r})$, respectively. $|F_{\mathbf{h}}F_{p\mathbf{h}}| \exp i(\varphi_{\mathbf{h}} - \varphi_{p\mathbf{h}})$ is the Fourier coefficient of $C(\mathbf{u})$: since it is a complex number, $C(\mathbf{u})$ is acentric [$C(\mathbf{u})$ is centric only if both $\rho(\mathbf{r})$ and $\rho_p(\mathbf{r})$ are centric]. It was shown by Carrozzini *et al.* (2010) that the space group of the C function is the symmorphic variant of the space group of the target structure (e.g. $P222$ as opposed to $P2_12_12_1$).

If the model and target structures are correlated then:

(a) $\varphi_{\mathbf{h}} \simeq \varphi_{p\mathbf{h}}$. Then the C map will show a peak at the origin, the amplitude of which increases with the correlation:

$$C(\mathbf{0}) = (1/V) \sum_{\mathbf{h}} |F_{\mathbf{h}}F_{p\mathbf{h}}| \exp i(\varphi_{\mathbf{h}} - \varphi_{p\mathbf{h}}). \quad (8)$$

(b) If $N_p \rightarrow N$ and $\langle |\Delta \mathbf{r}_j| \rangle \rightarrow 0$, then $\varphi_{p\mathbf{h}} \rightarrow \varphi_{\mathbf{h}}$ and $C(\mathbf{u}) \rightarrow P(\mathbf{u})$, where $P(\mathbf{u})$ is the Patterson function (Patterson, 1934*a,b*).

The map $C(\mathbf{u})$ cannot be computed during the phasing process, essentially because the $\varphi_{\mathbf{h}}$'s are unknown. Fortunately, the approximating function $C'(\mathbf{u})$, given by

$$C'(\mathbf{u}) = (1/V) \sum_{\mathbf{h}} m_{\mathbf{h}} |F_{\mathbf{h}}F_{p\mathbf{h}}| \exp(-2\pi i \mathbf{h} \cdot \mathbf{u}), \quad (9)$$

is easily computable. In equation (9) $m = \langle \cos(\varphi - \varphi_p) \rangle = I_1(X)/I_0(X)$ and $X = 2\sigma_A |EE_p| / (1 - \sigma_A^2)$ (Sim, 1959; Srinivasan & Ramachandran, 1965; Read, 1986). $C'(\mathbf{u})$ is a useful approximation to $C(\mathbf{u})$ and shows remarkable properties: the Fourier coefficients of $C'(\mathbf{u})$ are real numbers, and consequently the space group of $C'(\mathbf{u})$ is centric, so coinciding with the Patterson space group (e.g., $Pmmm$ if the space group of the target is $P2_12_12_1$).

It was observed by Carrozzini *et al.* (2010) that since both $\rho(\mathbf{r})$ and $\rho_p(\mathbf{r})$ are non-negative definite functions, $C(\mathbf{u})$ and therefore $C'(\mathbf{u})$ are also non-negative definite. The map $C'(\mathbf{u})$ may therefore be suitably modified and Fourier inverted, as in the usual electron-density modification (EDM) procedures (Cowtan, 1999; Abrahams, 1997; Abrahams & Leslie, 1996; Refaat & Woolfson, 1993; Giacovazzo & Siliqi, 1997), so leading to better estimates of the invariants ($\varphi_{\mathbf{h}} - \varphi_{p\mathbf{h}}$). The first applications of the procedure showed that it is able to drive to convergence sets of phases that are far away from the correct values.

In this paper we suggest applying the C' map to the Patterson deconvolution process, with particular interest in its integration with implication transformation (Simpson *et al.*, 1965; Pavelčík *et al.*, 1992) and superposition methods (Buerger, 1959; Richardson & Jacobson, 1987; Sheldrick, 1992).

Implication transformation and Patterson superposition methods are traditionally applied to structures containing heavy atoms. It has also been shown (Pavelčík, 1988; Pavelčík & Pivovarcikova, 2002) that they may be successfully applied to solve small (e.g., less than 80 atoms in the asymmetric unit) light-atom structures (i.e., no atoms heavier than O). No attempt has been made so far to solve medium-size (i.e., from 80 to 400 non-H atoms in the asymmetric unit) light-atom structures. In this paper we:

(a) characterize the interatomic peak distribution in a C' map and its relationship with Patterson interatomic peaks;

(b) describe a new procedure for *ab initio* phasing, involving the C' map rather than the Patterson map; and

(c) apply the new procedure to a large set of small- and medium-size structures, so demonstrating the greater efficiency of the new method.

We will not consider in this paper the application of the C' map to proteins, because it requires the use of some supplementary filtering techniques and will be described in a separate paper. However, as a general consequence of the results described here, Patterson deconvolution techniques, possibly integrated with the use of the C' map, should be considered as the most versatile tool for the solution of the crystallographic phase problem. This conclusion is supported by the following considerations: they succeed even in the case of light-atom structures (as shown in this paper), in the case of powder data (Burla *et al.*, 2007), and in the case of large proteins containing heavy atoms (up to 7900 atoms in the asymmetric unit with a data resolution of 1.65 Å, or also at 1.92 Å resolution for a protein with about 1300 non-H atoms in the asymmetric unit) (Caliandro *et al.*, 2008).

3. Patterson deconvolution and implication transformations

To better clarify the role of C' as powerful substitute for the P map, we recap the most advanced procedures aiming at deconvolving the Patterson function by using implication transformation and superposition methods. The typical procedure may be essentially summarized in three steps:

(i) Given the symmetry operator C_s , the related implication transformation $I_s(\mathbf{r})$ is calculated: it is a function of the atomic position \mathbf{r} defined over the corresponding Harker section.

$$I_s(\mathbf{r}) = P(\mathbf{r} - C_s \mathbf{r}) / n_s, \quad (10)$$

where P is the Patterson function and n_s is the number of symmetry operators that give rise to the same Harker section (Harker, 1936).

(ii) The symmetry minimum function is calculated, given by

$$\text{SMF}(\mathbf{r}) = \min_{s=1}^{\overline{m}} [I_s(\mathbf{r})], \quad (11)$$

where min indicates that SMF assumes in \mathbf{r} the minimum among the values of the \overline{m} independent functions $I_s(\mathbf{r})$.

(iii) The largest SMF peaks are used in turn to calculate the minimum superposition function

$$S(\mathbf{r}) = \min[P(\mathbf{r} - \mathbf{r}_q), \text{SMF}(\mathbf{r})]. \quad (12)$$

Sometimes more than one superposition vector is used, according to

$$S(\mathbf{r}) = \min[P(\mathbf{r} - \mathbf{r}_q), \dots, P(\mathbf{r} - \mathbf{r}_n), \text{SMF}(\mathbf{r})]. \quad (13)$$

From now on we will refer to the vectors $\mathbf{r}_q, \dots, \mathbf{r}_n$ as to *pivot vectors*, because they regulate the map overlap. As stated in §2, the above techniques have been recently revisited, with a

dramatic increase in efficiency (Caliandro *et al.*, 2008). The key to the improvements is in the following two additional steps:

(iv) Efficient filtering algorithms are applied to break down the additional crystallographic symmetry present in the function SMF and the residual Patterson symmetry in the $S(\mathbf{r})$ map.

(v) Subsequent cycles of electron density modification–difference electron density modification (EDM–DEDM) are automatically applied to the current maps to obtain higher-quality model maps.

In the next sections we will describe a simple algebra showing the potential advantages obtainable by replacing the Patterson by the function C' in step (iii) of the algorithm described above. The new algorithm for Patterson deconvolution will be described in §6 and its practical applications in §7.

4. C , C' and P maps

Let us analyze, in the C , C' and P maps, those features of the interatomic vectors that are particularly relevant for the application of implication transformation and superposition methods. In the algebraic calculations below we will emphasize the role of the heavy atoms, because they are often part of the model structure. The relations we will obtain are, however, quite general (the reader can set the number of heavy atoms equal to zero in our expressions if the structure is composed only of light atoms). We will suppose that:

(a) the target unit cell contains N atoms, NH of which are heavy atoms;

(b) the positions of the heavy atoms are known, with negligible errors, and constitute the model;

(c) \mathbf{r}_{Hi} , $i = 1, \dots, NH$ are the heavy-atom positions and \mathbf{r}_{lv} are the light-atom positions. We order the light atoms in such a way that they follow the heavy-atom list (*i.e.*, l goes from $NH + 1$ to N).

We want to obtain, under the above hypotheses, a more complete model of the target structure by replacing, in equations (12) and (13), $P(\mathbf{r} - \mathbf{r}_q)$ by $C(\mathbf{r} - \mathbf{r}_q)$ or $C'(\mathbf{r} - \mathbf{r}_q)$.

For ideal diffraction data, Patterson peak positions will be the union of the following three sets:

$$\{\mathbf{r}_{Hi} - \mathbf{r}_{Hj}, i, j = 1, \dots, NH\} \quad (14a)$$

$$\{\pm(\mathbf{r}_{Hi} - \mathbf{r}_{lv}), i = 1, \dots, NH, v = NH + 1, \dots, N\} \quad (14b)$$

$$\{\mathbf{r}_{lv} - \mathbf{r}_{l\mu}, v, \mu = NH + 1, \dots, N\}. \quad (14c)$$

In this order they correspond to heavy–heavy, heavy–light and light–light atom distances.

Let us now shift the Patterson map by the pivot vector \mathbf{r}_{Hq} : we should obtain a noisy image of the structure. The following peak sets arise:

$$\{\mathbf{r}_{Hi} - \mathbf{r}_{Hj} + \mathbf{r}_{Hq}, i, j = 1, \dots, NH\} \quad (15a)$$

$$\{\pm(\mathbf{r}_{Hi} - \mathbf{r}_{lv}) + \mathbf{r}_{Hq}, i = 1, \dots, NH, v = NH + 1, \dots, N\} \quad (15b)$$

$$\{\mathbf{r}_{lv} - \mathbf{r}_{l\mu} + \mathbf{r}_{Hq}, v, \mu = NH + 1, \dots, N\}. \quad (15c)$$

Emphasizing the case $j = q$ for the subset (15a) and the case $i = q$ for the subset (15b) allows us to rewrite the peaks (15) as

$$\{\mathbf{r}_{Hi}, i = 1, \dots, NH; \mathbf{r}_{Hi} - \mathbf{r}_{Hj} + \mathbf{r}_{Hq}, i, j = 1, \dots, NH, j \neq q\} \quad (16a)$$

$$\{\mathbf{r}_{lv}, v = NH + 1, \dots, N; (\mathbf{r}_{lv} - \mathbf{r}_{Hi}) + \mathbf{r}_{Hq}, i = 1, \dots, NH, i \neq q, v = NH + 1, \dots, N; (\mathbf{r}_{Hi} - \mathbf{r}_{lv}) + \mathbf{r}_{Hq}, i = 1, \dots, NH, v = NH + 1, \dots, N\} \quad (16b)$$

$$\{\mathbf{r}_{lv} - \mathbf{r}_{l\mu} + \mathbf{r}_{Hq}, v, \mu = NH + 1, \dots, N\}. \quad (16c)$$

If the peaks (16) are overlapped with the SMF map, the set (16a) would provide the heavy-atom substructure plus noise, the set (16b) would generate all the light-atom positions plus noise, and the set (16c) would produce only noise.

Let us now describe what we should obtain if the C map is shifted by the pivot vector \mathbf{r}_{Hq} . The interatomic vectors present in the C map will be the union of two sets: according to the chosen enantiomorph [which depends on whether in equation (7) we use $\exp i(\varphi_h - \varphi_{ph})$ or $\exp i(\varphi_{ph} - \varphi_h)$] the following sets arise (see Appendix A):

$$\{\mathbf{r}_{Hi} - \mathbf{r}_{Hj}, i, j = 1, \dots, NH\} \cup \{\mathbf{r}_{lv} - \mathbf{r}_{Hi}, j = 1, \dots, NH, v = NH + 1, \dots, N\} \quad (17a)$$

or

$$\{\mathbf{r}_{Hi} - \mathbf{r}_{Hj}, i, j = 1, \dots, NH\} \cup \{\mathbf{r}_{Hi} - \mathbf{r}_{lv}, j = 1, \dots, NH, v = NH + 1, \dots, N\}. \quad (17b)$$

The reader may notice that, in accordance with Appendix A, the heavy-atom–heavy-atom distances (more generally speaking, the distances between the atoms included in the model) constitute a centric set, even if the target space group is acentric, while the distances between heavy and light atoms form an acentric set. In our examples we will comply with (17a).

If the heavy-atom position \mathbf{r}_{Hq} is added to the interatomic vectors in equations (17), we obtain the following vectorial sets:

$$\{\mathbf{r}_{Hi}, i = 1, \dots, NH; \mathbf{r}_{Hi} - \mathbf{r}_{Hj} + \mathbf{r}_{Hq}, i, j = 1, \dots, NH, j \neq q\} \quad (18a)$$

$$\{\mathbf{r}_{lv}, v = NH + 1, \dots, N; (\mathbf{r}_{lv} - \mathbf{r}_{Hi}) + \mathbf{r}_{Hq}, i = 1, \dots, NH, i \neq q, v = NH + 1, \dots, N\}. \quad (18b)$$

We observe: (i) the set (18a) provides the heavy-atom substructure plus the same amount of noise included in the set (16a); (ii) the set (18b) generates all the light-atom positions, but less noise than in the set (16b); (iii) the set (16c), corresponding to light–light interatomic vectors, is absent when the C map is used.

As a consequence, the integration of the C map in the superposition method may offer a model structure less noisy than the Patterson map. Unfortunately, the C map is unknown during the phasing process, but we may replace it by its approximation: the C' map. If m is a good approximation of $\cos(\varphi - \varphi_p)$ for a sufficiently large set of reflections (for the

moment we will assume that this condition is satisfied), then the C' peaks will be located at

$$\{\mathbf{r}_{Hi} - \mathbf{r}_{Hj}, i, j = 1, \dots, NH\} \cup \{\pm(\mathbf{r}_{lv} - \mathbf{r}_{Hi}), j = 1, \dots, NH, v = NH + 1, \dots, N\}. \quad (19)$$

By adding the heavy-atom position \mathbf{r}_{Hq} to the interatomic vectors in (19) we obtain

$$\{\mathbf{r}_{Hi}, i = 1, \dots, NH, \mathbf{r}_{Hi} - \mathbf{r}_{Hj} + \mathbf{r}_{Hq}, i, j = 1, \dots, NH, j \neq q\} \quad (20a)$$

$$\{\mathbf{r}_{lv}, v = NH + 1, \dots, N; \mathbf{r}_{lv} - \mathbf{r}_{Hi} + \mathbf{r}_{Hq}, i = 1, \dots, NH, i \neq q, v = NH + 1, \dots, N; \mathbf{r}_{Hi} - \mathbf{r}_{lv} + \mathbf{r}_{Hq}, i = 1, \dots, NH, v = NH + 1, \dots, N\}. \quad (20b)$$

Let us compare sets (20), obtained by the use of the C' map, with sets (16) obtained by using the P map:

(i) (20a) and (20b) coincide with (16a) and (16b), respectively. Therefore both set (20) and set (16) provide the atomic positions of the full target structure, with the same amount of noise.

(ii) The C' map has no noise term corresponding to (16c). This property may be very important when the target structure contains a large number of light atoms, as frequently occurs for medium-size structures and for proteins.

This superior characteristic of the C' map requires that one condition is satisfied: the good quality of the model. To clarify this point, let us summarize the results of our algebraic analysis and combine them with the conclusions described in Appendix A. The C map may be calculated only if the cosines $\cos(\varphi - \varphi_p)$ are known: it is acentric if the target structure is acentric, with symmetry defined by the symmorphic variant of the space group of the target. Since the target phases are usually unknown, $\cos(\varphi - \varphi_p)$ may be replaced by its statistical estimate m . The corresponding map is now the C' map, centrosymmetric, with the same space group as the Patterson map. The C peaks (ideally) are also peaks of the C' map.

The C' map, ideally, will only show interatomic vectors relating the N_p atoms to the N atoms of the target, regardless of the value of N_p . The reader should not think that this property is guaranteed by the condition that m is a good approximation of $\cos(\varphi - \varphi_p)$: this conclusion is clearly demonstrated in Appendix A. As a consequence, the property will also hold when N_p is very small, for example when N_p is equal to one or two (see the practical applications described in §7). However, if N_p is very small, one wrong atomic position in the model structure may lead to wrong calculated amplitudes $|F_p|$ and therefore to wrong C' maps. The assumption that the quality of the model is a basic condition for the success of the C' -based procedure is therefore demonstrated.

There is another question which deserves to be discussed: are the C or the C' maps new variants (weighted) of the Patterson functions? The answer may be summarized as follows:

(i) The C map involves phases of the target and of the model structures, and therefore cannot be considered a variant

of the Patterson function. Indeed, the C -map symmetry does not coincide with the Patterson symmetry.

(ii) The C' map is part of the Patterson map, showing only model–target vectors. The weight m privileges the structure factors with large observed and calculated amplitudes. From a formal point of view, the C' function may be considered as a special Patterson function with weight $m|F_p|/|F|$, but this equivalence does not grasp the essence of the C' map. Indeed, while the Patterson function is calculated from the observations, the C' -map calculation needs a model, and therefore corresponding amplitudes and phases.

(iii) Some similarity exists between C' and the sum function (Buerger, 1959), where a weight is necessary: such a weight, however, is a complex number which is calculated from the translational vectors used for the image superposition.

The above considerations describe the ideal properties of the C' map. In practice, m is only a statistical estimate of $\cos(\varphi - \varphi_p)$: therefore the light-atom–light-atom peaks may be present in the C' map, but their intensity is expected to be much smaller compared with the corresponding peaks in the Patterson map.

A further property makes the use of the C' map more interesting: while the Patterson map is invariant during the phasing process, the C' map changes with the current model: we will show that this characteristic is crucial for a successful crystal structure determination.

A simple graphical demonstration of the above described algebra is shown in Fig. 1: it is very didactical, though not realistic, but it is useful for checking the properties of the vectorial sets. In Fig. 1(a) the target electron density is shown

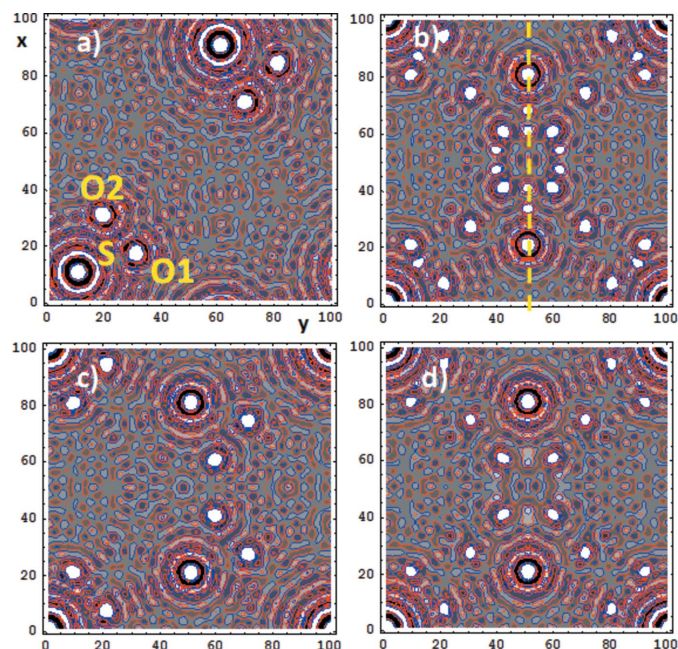


Figure 1
(a) A three-atom two-dimensional target structure: plane group pg ; (b) the corresponding Patterson map, plane group $p2mm$; (c) the C map when the S atom is the only atom in the model structure. Its symmetry group is pm , but the Patterson S-atom substructure satisfies the $p2mm$ symmetry; (d) the corresponding C' map, with $p2mm$ symmetry.

at 1 Å data resolution: a two-dimensional unit cell is used, with $a = b = 20$ Å, plane group pg , S is at (0.1, 0.1), and O atoms are at (0.165, 0.302) and (0.302, 0.185). In Fig. 1(b) the Patterson map is depicted, plane group $p2mm$. In Fig. 1(c) the C map is shown when S is the only atom in the model: the map is acentric, space group pm , while the component substructure corresponding to S–S distances is centric, with plane group $p2mm$ (since the model substructure contains only one symmetry-independent atom, the resulting vector substructure consists of two Harker peaks). The peaks corresponding to light-atom–light-atom distances are absent. The C' map is shown in Fig. 1(d): its plane group is $p2mm$, and the intensities corresponding to O–O distances are very weak.

An additional two-dimensional example (simulating a centric arrangement of the atoms in three dimensions) is shown in Fig. S1:¹ a two-dimensional unit cell is used, with $a = b = 20$ Å, plane group $p2$, with S at (0.1, 0.1), and O atoms at (0.1, 0.3) and (0.3, 0.1). In this case P , C and C' maps show the same $p2$ symmetry: C and C' are more closely related, and show very faint peaks corresponding to O–O distances.

In §2 we anticipated that the new phasing procedure, exploiting the properties of the C' map, may succeed even when applied to light-atom structures. We therefore need to generalize to light-atom structures the algebraic results described above, so far tailored for structures containing heavy atoms. For this we will suppose that the current model is composed of NH light atoms (instead of by NH heavy atoms). Then:

(a) in a C map the vectors between model atoms will show the Laue symmetry, the vectors between model and non-model atoms should satisfy the symmorphic variant of the target space group and, finally, the vectors between non-model atoms and non-model atoms should be absent; and

(b) in a C' map the vectors between model atoms and the vectors between model and non-model atoms should satisfy the Laue symmetry, and the vectors between non-model atoms and non-model atoms should be weak or absent.

5. The C' map deconvolution

The standard Patterson superposition is usually carried out by equations (11)–(13). In our applications (13) was not effective: the reasons are the following. Multiple map superposition, as symbolically represented in equation (13), has been attempted by several workers: the risk, frequently met in the practical applications, is that the resulting $S(\mathbf{r})$ map becomes poorer and poorer when the number of overlapped maps increases. Indeed, as an effect of resolution bias and of the vector overlap in the Patterson map, Patterson peaks are shifted from their ideal positions and the symmetry minimum function may vanish even in sites corresponding to heavy-atom positions in the electron-density map. A straightforward implementation of the superposition technique, based on the

shifted C' map rather than on the shifted Patterson map, would lead to

$$S(\mathbf{r}) = \min[C'(\mathbf{r} - \mathbf{r}_q), \text{SMF}(\mathbf{r})]. \quad (21)$$

To more clearly compare the expected features of (12) and (21) we observe:

(i) adding \mathbf{r}_q to the set of interatomic vectors (19) provides the set (20): target positional vectors and noise are obtained, as described in §4;

(ii) if equation (12) is applied, the noise reduction may not be equally efficient. Indeed, the P map always contains, particularly for large structures, a huge amount of noise corresponding to light-atom distances.

Some graphical examples (Fig. 2 and Figs. S2, S3) can illustrate the potential superiority of the C' -map-based deconvolution: we will use the ideal structures depicted in Figs. 1(a) and S1(a).

In Fig. 2(a) we show the SMF map corresponding to the two-dimensional structure illustrated in Fig. 1(a). According to the planar group symmetry pg , the SMF peaks are of columnar type. The four strongest columns correspond to the S–S Harker peaks: the number four is due to the origin ambiguity along the x axis. We freely choose for the x coordinate of the S atom that corresponding to the column highlighted by the yellow arrow: the y coordinate is free, and we choose, for simplicity, the true y coordinate of the sulfur. In Fig. 2(b) we show the Patterson map shifted by $\mathbf{r}_q = \mathbf{r}_s$; the minimum superposition functions (12) and (21) are shown, respectively, in Figs. 2(c) and 2(d). It can be noted that in both

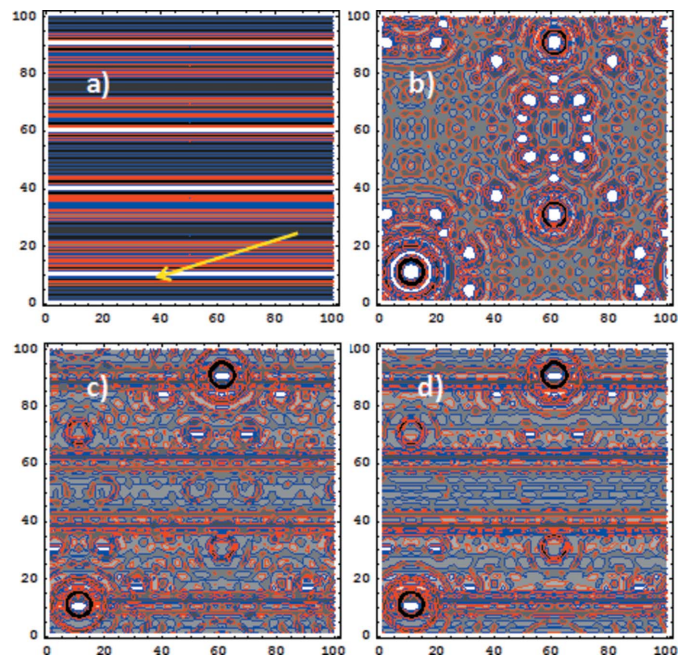


Figure 2
(a) SMF map corresponding to the structure illustrated in Fig. 1(a); (b) Patterson map shifted by \mathbf{r}_s (\mathbf{r}_s is the sulfur position, chosen along the columnar peak highlighted by the yellow arrow). For simplicity, the y coordinate of the sulfur has been arbitrary fixed to the true one; (c, d) symmetry minimum function obtained according to equations (12) and (21), respectively.

¹ Supplementary material for this paper is available from the IUCr electronic archives (Reference: PC5015). Services for accessing this material are described at the back of the journal.

cases the full structure is obtained but some false peaks are present, and using (12) or (21) leads to equivalent results.

In Fig. S2 we show the results obtained by applying the Patterson deconvolution to the target structure depicted in Fig. S1(a). Since the strongest Patterson peak, say the S–S peak, is at (0.2, 0.2), the SMF map will show a strong peak at (0.1, 0.1) (see Fig. S2a), which is used as a pivot vector $\mathbf{r}_q = \mathbf{r}_s$ in map superposition. In Fig. S2(b) we show the Patterson map shifted by \mathbf{r}_s and in Figs. S2(c) and S2(d) the minimum superposition functions (12) and (21), respectively: the full structure is again obtained with some false peaks, but the use of (21) provides a cleaner map.

In more detail, we observe:

(a) The $S(\mathbf{r})$ maps obtained *via* equation (12) (say Figs. 2c and S2c) and equation (21) (say Figs. 2d and S2d) have very sharp peak domains, a symptom of partial peak overlap at the atom sites (even in this ideal condition).

(b) None of the $S(\mathbf{r})$ maps is completely deconvoluted, because in all the cases a double superposition is calculated: a triple one would be necessary to obtain a single image of the structure.

(c) The larger efficiency of the C' -map-based deconvolution with respect to that based on the Patterson map is expected to increase with the structure complexity, where the disturbance from light-atom–light-atom vectors is larger: this is confirmed by the experimental tests described in §6.

An alternative approach can be also followed, by replacing the SMF with the atomic minimum superposition (AMS) introduced by Simpson *et al.* (1965) and Pavelčík (1986). Let \mathbf{r}_q be the pivot vector and $\mathbf{C}_\tau \mathbf{r}_q$ a position symmetry equivalent to \mathbf{r}_q (*i.e.*, $\mathbf{C}_\tau \neq \mathbf{I}$ is a symmetry operator of the target space group). Then the AMS is calculated as

$$S(\mathbf{r}) = \min_{\tau=2}^m [P(\mathbf{r} - \mathbf{r}_q), P(\mathbf{r} - \mathbf{C}_\tau \mathbf{r}_q)], \quad (22)$$

which can be modified (by introducing the C' map) according to

$$S(\mathbf{r}) = \min_{\tau=2}^m [C'(\mathbf{r} - \mathbf{r}_q), C'(\mathbf{r} - \mathbf{C}_\tau \mathbf{r}_q)]. \quad (23)$$

The reader should observe that in equation (23): (i) in accordance with the reasons mentioned above, the two C' maps are not superimposed with the SMF map: this should only be used for finding the atomic positions to be employed as pivots; (ii) overlapping maps by using pivot vectors which are related by symmetry reduces the risk of obtaining poor $S(\mathbf{r})$ functions: such a risk is higher if two symmetry-independent pivot vectors are used. The deconvolutions performed according to (22) and (23) are reported in Figs. S3(a,c) and S3(b,d) for the pg and $p2$ examples, respectively. The resulting S maps are no longer constituted by sharp peaks: again, using (22) or (23) leads to equivalent results for the pg case, using (23) provides a cleaner map for the $p2$ case.

The use of (22) and (23) implies a large degree of map superposition for high-symmetry space groups, with the consequent risk of obtaining a final map that is too poor. To overcome this problem, we preferred to obtain a partial

decomposition and to use only two pivot vectors related by symmetry, according to

$$S(\mathbf{r}) = \min[P(\mathbf{r} - \mathbf{r}_q), P(\mathbf{r} - \mathbf{C}_2 \mathbf{r}_q)], \quad (24)$$

$$S(\mathbf{r}) = \min[C'(\mathbf{r} - \mathbf{r}_q), C'(\mathbf{r} - \mathbf{C}_2 \mathbf{r}_q)]. \quad (25)$$

As we shall see in the next section, equations (12), (21), (24) and (25) are tools of our deconvolution algorithm.

6. The new deconvolution algorithm

Using the C' map in superposition techniques requires a preliminary condition: a model structure should already be available. That may be achieved (staying with direct-space techniques) in a simple way. First, the Patterson map is computed, then the implication transformation $I_s(\mathbf{r})$ as defined by equation (10) is obtained, then the symmetry minimum function SMF(\mathbf{r}) as defined by equation (11) is derived, from which a starting model may be extracted. This corresponds to the block SMF in Fig. 3, where a schematic view of our algorithm and of all the phasing procedure is given. Owing to the dynamic nature of the C' map, the quality and the size of the model may change during the deconvolution process. To take into account this feature, we designed a three-step algorithm, in which the $S(\mathbf{r})$ map obtained at a given step is used to generate a new C' map to be used in the next step (see Fig. 3a).

Step A submits the coordinates and thermal factor of the pivot peak \mathbf{r}_q obtained in the block SMF to least-squares refinement (only if an atom heavier than Ca is present in the structure) and applies equation (12).

In step B the highest peak of the $S(\mathbf{r})$ map obtained at the end of step A, possibly submitted to least-squares refinement, is chosen as pivot peak \mathbf{r}_p : the corresponding peak is used as a model to build the C' map. Equation (21) is then applied to obtain a new $S(\mathbf{r})$ map, where interatomic vectors corresponding to distances between atoms not in the model have reduced intensity.

In step C two peaks are selected from the $S(\mathbf{r})$ map obtained at the end of step B: \mathbf{r}_{p1} , the highest one, and \mathbf{r}_{p2} , chosen as the highest peak at least 2 Å away from \mathbf{r}_{p1} (to avoid taking a ripple of \mathbf{r}_{p1}). \mathbf{r}_{p1} and \mathbf{r}_{p2} are used as a model for building a further C' map, which, shifted by \mathbf{r}_{p1} , is overlapped with the SMF map according to equation (21). The resulting $S(\mathbf{r})$ map is subjected to cycles of EDM in which filtering algorithms are applied to break down the residual Patterson symmetry. They have not been changed with respect to our previous implementation (Caliandro *et al.*, 2007): they were based on the properties of the FF and of the Patterson maps. The only modification introduced here is that now the C' map calculated in step C plays the role of the Patterson map in the old algorithm.

The algorithm described so far may also be applied to the space group $P1$, provided that the SMF map is replaced by the Patterson map in equations (12) and (21).

The above algorithm has been embedded in the *SIR2011* framework for *ab initio* crystal structure solution (Burla *et al.*, 2012), as sketched in Fig. 3(b). The trials progressively

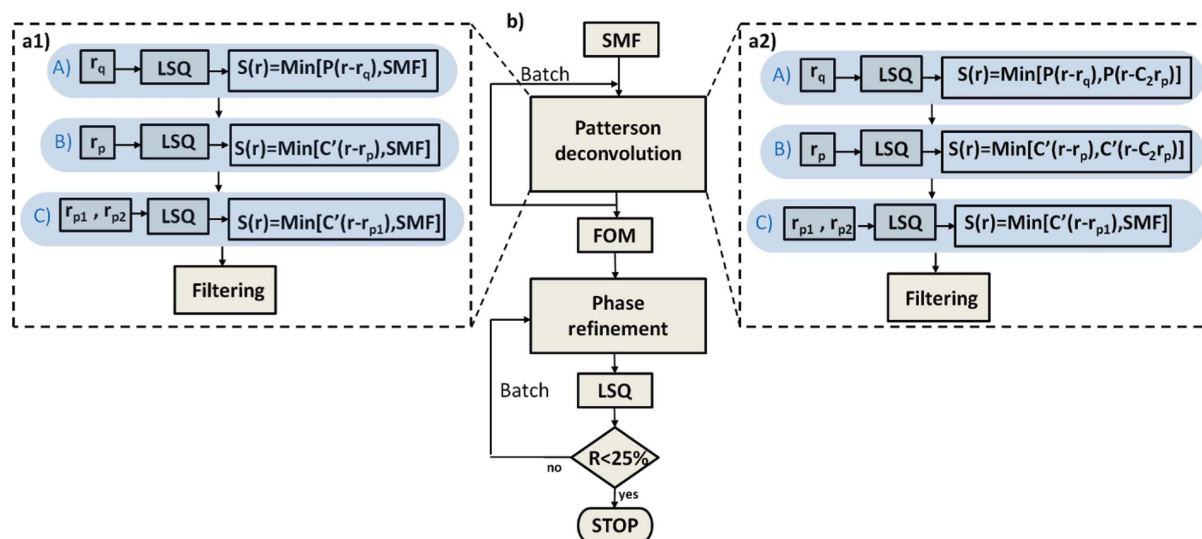


Figure 3 Schematic view of the new Patterson deconvolution algorithm in the first (a1) and second (a2) iteration and of the *ab initio* structure solution strategy of *SIR2011* (b).

obtained are ordered by a figure of merit (FOM) (Burla *et al.*, 2004), and then submitted to phase refinement, which is based on cycles of EDM and DEDM, followed by routines that interpret the electron-density map in terms of a structural model. The *RELAX* procedure (Burla *et al.*, 2005) is also used at this stage for the most promising trials. The structural model is then refined by diagonal least squares (block LSQ in Fig. 3), after which the crystallographic residual is calculated. If it falls below a given threshold (0.25 by default), the program stops.

If the solution is not obtained at the end of the whole procedure, a second iteration is performed using a different version of the deconvolution module: equations (24) and (25) are used in steps A and B instead of equations (12) and (21), respectively.

A last point deserves to be clarified. The reader has certainly noticed that we choose a one- or two-atom symmetry-independent model structure, while the target structure may contain a large number of atoms. According to Appendix A, the *C'* map will show only the interatomic vectors relating model to target atoms, and therefore is a small or a large part of the Patterson map depending on whether N_p is small or large. If N_p is very small, the number of interatomic vectors in the unit cell is much smaller than in the Patterson map: this strongly reduces the vector overlap, improves the peak location and, consequently, minimizes the error in the model structure.

7. Applications

Any new algorithm should be checked by a large number of test structures to prove its general applicability. We used a total of 188 test structures, whose code names, space groups, chemical compositions and references are listed as supplementary information (Tables S1 and S2). To obtain statistical figures summarizing the results of the phasing process, we divided the set of test structures into subsets according to the

Table 1 Statistics for the test structures.

Values are grouped according to the number of non-H atoms in the asymmetric unit (N_{asym}) and the presence or absence of atoms heavier than Ca (*H/L*). Entries correspond to the number of structures in each class (Num), and the phasing efficiencies of the *C'*-based ($Eff_{C'}$) and of the Patterson-based deconvolution (Eff_P) algorithms

	<i>H</i>			<i>L</i>		
	Num	$Eff_{C'}$	Eff_P	Num	$Eff_{C'}$	Eff_P
$N_{\text{asym}} \leq 20$	23	100	100	14	100	100
$20 < N_{\text{asym}} \leq 80$	21	100	100	45	100	96
$80 < N_{\text{asym}} \leq 150$	14	100	100	46	100	93
$150 < N_{\text{asym}}$	4	100	100	21	90	81

number of non-H atoms in the asymmetric unit (N_{asym}) and their heavier atomic species ($L = \text{light}, H = \text{heavy}$): for practical usefulness an atom is considered heavy if its atomic number is larger than that of Ca. The number of test structures for each subset are reported in Table 1, where we also show the average phasing efficiencies of two deconvolution methods, the first *C'*-based (as described in §5), the second Patterson-based, carried out by using equation (12). For each structure subset $Eff_{C'}$ and Eff_P are the ratio (number of solved structures)/(number of structures) when the *C'*-based and the Patterson-based algorithms are used, respectively. In Fig. 4 we show, for the *C'*-based algorithm and for the *H* and *L* subsets, the average values of: the crystallographic residual (R_{cr}), the sequential number of the pivot peak for which the solution is obtained (n_p), and the cpu time necessary to reach the solution (t in min) versus the average values of N_{asym} ($\langle N_{\text{asym}} \rangle$). The averages have been calculated over the test structures contained in each of the four subsets considered in Table 1.

It can be noted that:

(i) The deconvolution algorithm based on the *C'* map is more efficient than that based on the Patterson map for structures containing only light atoms, and in particular for

difficult cases, *i.e.* for structures with $N_{\text{asym}} > 150$. If heavy atoms are present, both algorithms work with full efficiency.

(ii) On average, the crystallographic residual does not depend on the structure complexity within the range of N_{asym} considered.

(iii) For structures containing heavy atoms, the solution is obtained by the very first (higher) SMF peaks, indicating that the true heavy-atom position is always found in the very early stage of the phasing procedure. As a consequence, the cpu time needed to reach the solution simply scales with the size of the structure. The efficiency of the deconvolution algorithm does not depend on the size of the structure in this case.

(iv) For structures containing only light atoms, a steady increase of the pivot peak order number as a function of the

structure complexity is seen, indicating that locating light atoms in their true positions by superposition techniques is more and more challenging as the size of the structure increases. Therefore the cpu time increases with N_{asym} more rapidly than for the heavy-atom case.

Most of the medium-size test structures are a challenge for any *ab initio* phasing program, especially if they include only light atoms: thus the results obtained for the *L* structures with $N_{\text{asym}} > 150$ and no atom heavier than O are reported in Table 2. The efficiency of the algorithm does not seem to depend on the crystal symmetry. The only structures resistant to the new phasing method are *ceho2z* and *cemc2z*: the failure may be attributed to the large measurement errors for the high-resolution spots ($R_{\text{int}} > 45\%$ for data at atomic resolution), which strongly affect the Patterson or *C'* map quality. A unique structure (*cyclo_bnz*) has been solved by iterating the whole decomposition procedure (part of the standard approach): after having explored all the 23 SMF peaks by using the SMF approach, for a total of 1472 s, the AMS approach took 180 s to reach the solution. It is worth noting that the AMS approach is less efficient (it has 5 failures), but it constitutes an alternative to the SMF approach.

8. Conclusions

The *C* map, the Fourier transform of the product $|F_{\text{h}}F_{\text{ph}}| \exp i(\varphi_{\text{h}} - \varphi_{\text{ph}})$, has been crystallographically characterized, together with its centric modification *C'*, the Fourier transform of the product $m_{\text{h}}|F_{\text{h}}F_{\text{ph}}|$. The *C'* map, easily computable as soon as a model is available, shows, with reduced intensity, peaks corresponding to interatomic vectors between atoms not belonging to the model. This property may be very useful when the *C'* map (instead of the Patterson map) is combined with implication transformation methods and superposition techniques, because it is less subject to noise.

Our applications show that superposition methods, combined with the *C'* map, can solve small- and medium-size structures (up to about 400 non-H atoms in the asymmetric unit) even when no heavy atom (*i.e.*, no atom heavier than O) is present. This result dramatically changes the common judgement about the versatility of Patterson techniques. In *International Tables for Crystallography* Volume B, Rossmann & Arnold (1993) write

The feasibility of structure solution by the heavy-atom method depends on a number of factors which include the relative size of the heavy atom and the extent and quality of the data. A useful rule of thumb is that the ratio

$$r = \frac{\sum_{\text{heavy}} Z^2}{\sum_{\text{light}} Z^2}$$

should be near unity if the heavy atom is to provide useful starting phase information (Z is the atomic number of an atom). The condition that $r > 1$ normally guarantees interpretability of the Patterson function in terms of the heavy-atom positions.

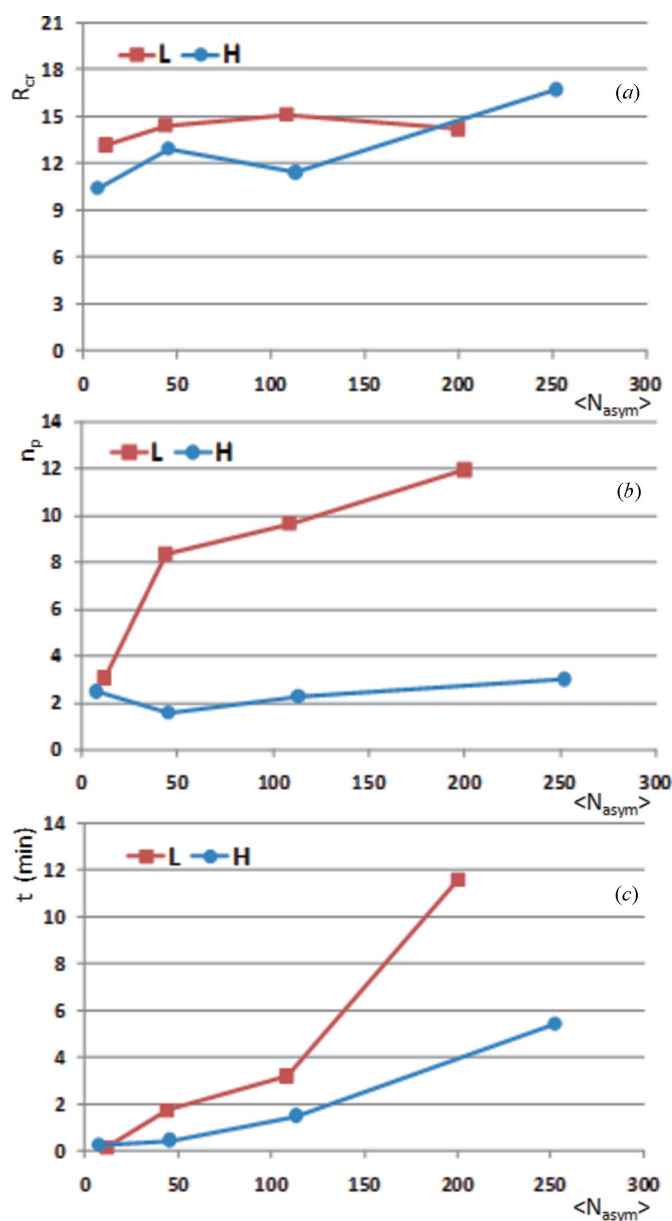


Figure 4
For *H* and *L* structures we show, versus $\langle N_{\text{asym}} \rangle$: (a) the average values of the crystallographic residual R_{cr} ; (b) the pivot peak sequential order n_p at which the solution has been obtained; and (c) the cpu time (t in min) needed to reach the solution.

They also state that the rule is rather conservative and quote as an outstanding example vitamin B₁₂ with formula C₆₂H₈₈CoO₁₄P (Hodgkin *et al.*, 1957), which gave $r = 0.14$ for the Co atom alone.

Our application clearly shows that heavy atoms are no longer strictly necessary for the success of Patterson– C' procedures, which may successfully be applied also to structures with atoms not heavier than O. This result makes Patterson techniques probably the most versatile phasing method, given also their ability to solve protein structures with heavy atoms up 7900 atoms in the asymmetric unit with a data resolution of 1.65 Å (Caliandro *et al.*, 2008).

APPENDIX A

Symmetry properties of the C map

The symmetry properties of the C map may be obtained by a simple analysis of the product $F_{\mathbf{h}}F_{\rho\mathbf{h}}$. In accordance with §2, we will suppose that the unit cells of the target and of the model structure contain N and N_p atoms, respectively, and that t and t_p are the corresponding numbers of atoms in the asymmetric units. Let $\mathbf{C} = (\mathbf{R}, \mathbf{T})$ be the generic symmetry operator of the space group and m its order; then

$$F_{\mathbf{h}}F_{-\rho\mathbf{h}} = \sum_{i=1}^t \sum_{j=1}^{t_p} \sum_{s,k=1}^m f_i f_j \exp[2\pi i \mathbf{h} \cdot (\mathbf{C}_s \mathbf{r}_i - \mathbf{C}_k \mathbf{r}_{pj})]. \quad (26)$$

We subdivide the summation over the target atoms into two parts:

$$F_{\mathbf{h}}F_{-\rho\mathbf{h}} = \sum_{i=1}^{t_p} \sum_{j=1}^{t_p} \sum_{s,k=1}^m f_i f_j \exp[2\pi i \mathbf{h} \cdot (\mathbf{C}_s \mathbf{r}_i - \mathbf{C}_k \mathbf{r}_{pj})] + \sum_{i=t_p+1}^t \sum_{j=1}^{t_p} \sum_{s,k=1}^m f_i f_j \exp[2\pi i \mathbf{h} \cdot (\mathbf{C}_s \mathbf{r}_i - \mathbf{C}_k \mathbf{r}_{pj})]. \quad (27)$$

Table 2

L test structures with $N_{\text{asym}} > 150$ and no atom heavier than O.

The number of the pivot peak at which the solution is obtained by the C' -based algorithm, the final crystallographic residual and the cpu time necessary for achieving the solution are given.

Name	Reference	Space group	N_{asym}	Cell content	Pivot peak	R (%)	Time (min)
babu	(a)	Cc	204	C ₅₂₈ H ₅₇₆ N ₉₆ O ₁₉₂	1	13.1	0.8
ceho2z	(b)	$P2_1$	232	C ₂₆₄ H ₃₀₄ N ₃₂ O ₁₆₈	—	—	—
cemc2z	(c)	$P2_1$	207	C ₂₁₀ H ₄₄₄ N ₄ O ₂₀₀	—	—	—
cyclo_bnz	(d)	$P2_1$	206	C ₂₁₆ H ₂₂₄ N ₈ O ₁₈₉	18	13.8	27.5
cyclo_dba	(d)	$P2_12_12_1$	219	C ₅₀₈ H ₄₄₈ N ₁₆ O ₃₅₂	43	13.8	39.5
dodecak	(e)	$P2_12_12_1$	200	C ₅₆₈ H ₉₁₂ N ₉₆ O ₁₃₆	4	17.5	1.4
helix	(f)	$C2$	164	C ₄₇₂ H ₇₃₆ N ₉₂ O ₉₂	45	19.6	19.2
inclu_rt	(g)	$P1$	194	C ₉₇ H ₁₅₇ O ₉₇	1	13.1	0.6
ohba_p1	(h)	$P1$	188	C ₁₆₆ H ₁₅₈ N ₂ O ₂₀	1	13.5	0.6
tb02rlmk	(i)	$P1$	197	C ₉₆ H ₂₂₄ O ₁₀₁	1	12	0.3
tb	(j)	$P2_12_12_1$	186	C ₅₂₀ H ₇₈₄ O ₂₂₄	22	11.2	52.5
tp	(j)	$P2_1$	161	C ₂₁₀ H ₃₀₈ O ₁₁₂	17	9.1	12.2
tval	(k)	$P1$	156	C ₁₀₈ H ₁₈₀ N ₁₂ O ₃₆	1	11.9	0.3

References: (a) Narendra Babu *et al.* (2009); (b) Onagi *et al.* (2003); (c) Edwards (2003); (d) Giastas *et al.* (2003); (e) Saviano *et al.* (2004); (f) Rudresh *et al.* (2004); (g) Makedonopoulou & Mavridis (2000); (h) Ohba & Miyamoto (2004); (i) H. Gornitzka, personal communication; (j) K. Gessle, personal communication; (k) Karle (1975).

Both terms on the right-hand side of equation (27) involve acentric distributions of interatomic distances. However, if the model and target structures are highly correlated, the first term is pseudo-centric: it becomes completely centric when $D = 1$ (that is, when $\mathbf{r}_{pj} \equiv \mathbf{r}_j$, $j = 1, \dots, N_p$). In this case

$$F_{\mathbf{h}}F_{-\rho\mathbf{h}} = \sum_{i=1}^{t_p} \sum_{j=1}^{t_p} \sum_{s,k=1}^m f_i f_j \exp[2\pi i \mathbf{h} \cdot (\mathbf{C}_s - \mathbf{C}_k) \mathbf{r}_{pj}] + \sum_{i=t_p+1}^t \sum_{j=1}^{t_p} \sum_{s,k=1}^m f_i f_j \exp[2\pi i \mathbf{h} \cdot (\mathbf{C}_s \mathbf{r}_i - \mathbf{C}_k \mathbf{r}_{pj})]. \quad (28)$$

This feature was not noticed in the symmetry analysis described by Carrozzini *et al.* (2010). Let us now derive the symmetry of the product $F_{\mathbf{h}}F_{-\rho\mathbf{h}}$, and therefore of the C map. Defining

$$\mathbf{C}_s = \mathbf{C}_k \mathbf{C}_\mu = (\mathbf{R}_k \mathbf{R}_\mu, \mathbf{R}_k \mathbf{T}_\mu + \mathbf{T}_k)$$

gives

$$F_{\mathbf{h}}F_{-\rho\mathbf{h}} = \sum_{i=1}^t \sum_{j=1}^{t_p} \sum_{s,k=1}^m f_i f_j \exp[2\pi i \mathbf{h} \mathbf{R}_k \cdot (\mathbf{C}_\mu \mathbf{r}_i - \mathbf{r}_{pj})]. \quad (29)$$

If the model is highly correlated with the target structure, equation (29) may be approximated by

$$F_{\mathbf{h}}F_{-\rho\mathbf{h}} = \sum_{i=1}^{t_p} \sum_{j=1}^{t_p} \sum_{s,k=1}^m f_i f_j \exp[2\pi i \mathbf{h} \mathbf{R}_k \cdot (\mathbf{C}_\mu \mathbf{r}_{pi} - \mathbf{r}_{pj})] + \sum_{i=t_p+1}^t \sum_{j=1}^{t_p} \sum_{s,k=1}^m f_i f_j \exp[2\pi i \mathbf{h} \cdot (\mathbf{C}_s \mathbf{r}_i - \mathbf{C}_k \mathbf{r}_{pj})]. \quad (30)$$

Each interatomic vector in equations (29) or (30) depends on the space-group symmetry of the target structure, but the symmetry of the interatomic vectors obeys that of the corresponding symmorphic space group. The reader, however, should notice that, according to equation (28), the interatomic vectors among model atoms will show the symmetry of the Laue group, because their distribution is centric.

References

- Abrahams, J. P. (1997). *Acta Cryst.* **D53**, 371–376.
 Abrahams, J. P. & Leslie, A. G. W. (1996). *Acta Cryst.* **D52**, 30–42.
 Buerger, M. J. (1959). *Vector Space*, ch. 11. New York: Wiley.
 Burla, M. C., Caliandro, R., Camalli, M., Carrozzini, B., Cascarano, G. L., De Caro, L., Giacovazzo, C., Polidori, G. & Spagna, R. (2005). *J. Appl. Cryst.* **38**, 381–388.
 Burla, M. C., Caliandro, R., Camalli, M., Carrozzini, B., Cascarano, G. L., Giacovazzo, C., Mallamo, M., Mazzzone, A., Polidori, G. & Spagna, R. (2012). *J. Appl. Cryst.* **45**, 357–361.
 Burla, M. C., Caliandro, R., Carrozzini, B., Cascarano, G. L., De Caro, L., Giacovazzo, C. & Polidori, G. (2004). *J. Appl. Cryst.* **37**, 791–801.
 Burla, M. C., Caliandro, R., Carrozzini, B., Cascarano, G. L., De Caro, L., Giacovazzo, C., Polidori, G. & Siliqi, D. (2007). *J. Appl. Cryst.* **40**, 834–840.
 Caliandro, R., Carrozzini, B., Cascarano, G. L., De Caro, L., Giacovazzo, C., Mazzzone, A. & Siliqi, D. (2008). *J. Appl. Cryst.* **41**, 548–553.

- Caliandro, R., Carrozzini, B., Cascarano, G. L., De Caro, L., Giacobozzo, C. & Siliqi, D. (2007). *J. Appl. Cryst.* **40**, 883–890.
- Carrozzini, B., Cascarano, G. L. & Giacobozzo, C. (2010). *J. Appl. Cryst.* **43**, 221–226.
- Cowtan, K. (1999). *Acta Cryst.* **D55**, 1555–1567.
- Edwards, A. (2003). Personal communication.
- Giacobozzo, C. & Siliqi, D. (1997). *Acta Cryst.* **A53**, 789–798.
- Giastas, P., Yannakopoulou, K. & Mavridis, I. M. (2003). *Acta Cryst.* **B59**, 287–299.
- Harker, D. (1936). *J. Chem. Phys.* **4**, 381–390.
- Hodgkin, D. C., Kamper, J., Lindsey, J., MacKay, M., Pickworth, J., Robertson, J. H., Shoemaker, C. B., White, J. G., Prosen, R. J. & Trueblood, K. N. (1957). *Proc. R. Soc. London Ser. A*, **242**, 228–263.
- Karle, I. L. (1975). *J. Am. Chem. Soc.* **97**, 4379–4386.
- Makedonopoulou, S. & Mavridis, I. M. (2000). *Acta Cryst.* **B56**, 322–331.
- Narendra Babu, S. N., Abdul Rahim, A. S., Osman, H., Jebas, S. R. & Fun, H.-K. (2009). *Acta Cryst.* **E65**, o1560–o1561.
- Ohba, S. & Miyamoto, H. (2004). *Acta Cryst.* **E60**, o216–o218.
- Onagi, H., Carrozzini, B., Cascarano, G. L., Easton, C. J., Edwards, A. J., Lincoln, S. F. & Rae, A. D. (2003). *Chem Eur. J.* **9**, 5971–5977.
- Patterson, A. L. (1934a). *Phys. Rev.* **45**, 763.
- Patterson, A. L. (1934b). *Phys. Rev.* **46**, 372–376.
- Pavelčík, F. (1986). *J. Appl. Cryst.* **19**, 488–491.
- Pavelčík, F. (1988). *Acta Cryst.* **A44**, 724–729.
- Pavelčík, F., Kuchta, L. & Sivý, J. (1992). *Acta Cryst.* **A48**, 791–796.
- Pavelčík, F. & Pivovarcikova, O. (2002). *J. Appl. Cryst.* **35**, 526–532.
- Read, R. J. (1986). *Acta Cryst.* **A42**, 140–149.
- Refaat, L. S. & Woolfson, M. M. (1993). *Acta Cryst.* **D49**, 367–371.
- Richardson, J. W. & Jacobson, R. A. (1987). *Patterson and Pattersons*, edited by J. P. Glusker, B. K. Patterson & M. Rossi, pp. 310–317. Oxford University Press.
- Rossmann, M. G. & Arnold, E. (1993). *International Tables for Crystallography*, Vol. B, *Reciprocal Space*, edited by U. Shmueli, ch. 2.3. Dordrecht: Kluwer.
- Rudresh, R. S., Ramagopal, U. A., Inai, Y., Goel, S., Sahal, D. & Chauhan, V. S. (2004). *Structure*, **12**, 389–396.
- Saviano, M., Improta, R., Benedetti, E., Carrozzini, B., Cascarano, G. L., Didierjean, C., Toniolo, C. & Crisma, M. (2004). *ChemBioChem*, **5**, 541–544.
- Sheldrick, G. M. (1992). *Crystallographic Computing 5*, edited by D. Moras, A. D. Podjarny & J. C. Thierry, pp. 145–157. Oxford University Press.
- Sim, G. A. (1959). *Acta Cryst.* **12**, 813–815.
- Simpson, P. G., Dobrott, R. D. & Lipscomb, W. N. (1965). *Acta Cryst.* **18**, 169–179.
- Srinivasan, R. & Ramachandran, G. N. (1965). *Acta Cryst.* **19**, 1008–1014.