Connectivity properties of high-density regions and *ab initio* phasing at low resolution

Vladimir Y. Lunin,[a,b] Natalia L. Lunina[a] and Alexandre G. Urzhumtsev[b]*

[a]Institute of Mathematical Problems of Biology, Russian Academy of Sciences, Pushchino, Moscow Region, 142292 Russia, and [b]LCM3B, UPRESA 7036 CNRS, Faculté des Sciences, Université Henri Poincaré, Nancy I, 54506 Vandoeuvre-les-Nancy, France. Correspondence e-mail: sacha@lcm3b.u-nancy.fr

It is expected that a correctly phased low-resolution synthesis would show a set of isolated 'blobs' located near to the centres of the macromolecules if the corresponding cut-off level is chosen properly. This is not always the case when using experimentally measured structure-factor magnitudes. Nevertheless, this property can be efficiently used as a constraint in the low-resolution *ab initio* phasing of structure factors. The suggested procedure consists in generating a large number of random phase sets, selecting those that together with the observed magnitudes result in the desired number of blobs in Fourier syntheses, and averaging the selected phase sets. The current paper discusses the formal definitions, analysis of low-resolution syntheses, some phasing algorithms and their application to *ab initio* phasing.

1. Introduction

If a single diffraction experiment data set is used for a structure determination, then additional 'theoretical' information is necessary to supply a set of the observed structure-factor magnitudes with their phases. This information can be specific for the studied object (*e.g.* approximate atomic model) or common for a class of structures. For example, such very general properties as the 'atomicity' of the electron-density distribution in a crystal (Sayre, 1952) or its positivity (Karle & Hauptman, 1950) may be used with different successes to recover structure-factor phases. This paper suggests a way to use another type of general information, namely topological properties of regions of high electron density, for low-resolution *ab initio* phasing in macromolecular crystallography. The simplest example of such a property is the connectivity, which was used for many years to estimate the quality of electron-density maps and was formalized as a quantitative criterion by Baker *et al.* (1993).

When solving the phase problem, it is natural to use different kinds of additional information that complement each other. The goal of the current paper is to study the phasing power of topological information rather than to determine a particular structure. That is why in the tests below the only information used in addition to the observed magnitudes is the simplest topological feature of the molecular region such as the number of connected components forming the region and the volumes of these components.

It must be noted that the true electron-density distribution is very rarely available for macromolecular crystallographers. Usually, one deals with a finite Fourier synthesis approxima-tion of the density distribution rather than with the density distribution itself. Only recently did the analysis of fine details of the density distribution in a protein crystal become possible (Jelsch *et al.*, 1998; Housset, 1999; Podjarny, 1999) using super-high-resolution data in a way similar to small-molecule studies (Lecomte, 1998). The properties of a Fourier synthesis may be similar to those of the true electron-density distributions if a large enough data set was used to calculate the synthesis and they may differ significantly for low- and medium-resolution syntheses. For example, the property of electron-density distributions to be non-negative is conserved to some extent in high-resolution syntheses. On the contrary, low- and medium-resolution syntheses must have negative values when calculated with the exact phases. Similarly, topological properties of regions in the unit cell composed of points with high density values in a Fourier synthesis also depend on the set of structure factors used for the synthesis calculation. Therefore, different topological criteria should be applied to Fourier syntheses when working with data sets of different resolutions. The properties of low-resolution data sets are discussed below.

2. Basic definitions

Let $\rho(\mathbf{r})$ be a function in the unit cell and κ a cut-off level. We define a high-values region corresponding to the chosen cut-off κ as the set of all points $\mathbf{r}$ in the unit cell such that $\rho(\mathbf{r}) > \kappa$:

$$\Omega_\kappa = \{\mathbf{r} : \rho(\mathbf{r}) > \kappa\}. \tag{1}$$

We study below some of the simplest properties of these regions such as the number of components composing the

region Ω_κ and their volumes. (To be precise, the volume is not a topological property of a region, but we include this information in the analysis too.) A set of real-space points is considered as a connected component if every two points in this set may be connected by a continuous curve such that all points of the curve belong to the set.

Some notions must be performed to refine the terms when working with periodic functions. First, when discussing the number of components, we count the number of components in one unit cell. It may occur that some components in Ω_κ can extend continuously through the whole space (Fig. 1, component C). In this case, the component volume is defined as the volume of the part of the component bounded by the unit cell but the component is marked as endless in our analysis (see *e.g.* Table 1, where the endless components are marked by the symbol ∞). Another special case is that when the component is finite but the boundaries of the unit cell cut this component (Fig. 1, component A). In this case, we suppose the component to be continued periodically outside the unit cell.

In practical applications, a Fourier synthesis is calculated at a grid in the unit cell and a region of highest values Ω_κ becomes a discrete set of grid points. It is natural to define now a set of grid points as a connected one if any of its points can be reached from any other by travelling over the points of this region from neighbour to neighbour. The number of points in the set estimates the volume of the corresponding component.

In the discrete case, the neighbouring points can be defined in various ways. In the tests below, the set of neighbours of the grid point (i, j, k) was formed by the six points $(i-1, j, k)$, $(i+1, j, k)$, $(i, j-1, k)$, $(i, j+1, k)$, $(i, j, k-1)$, $(i, j, k+1)$. More complicated definitions of the set of neighbours could be applied too [including 'diagonal' ones $(i+1, j+1, k)$ *etc.*] but the difference in results is not large if a fine enough grid is used. The numbers of grid points used in the connectivity analysis for the test structures considered below are given in Table 1. They correspond approximately to $d_{min}/8$ spacing in the unit cell.

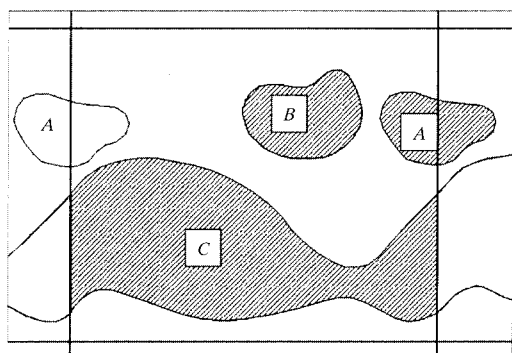The properties of a region Ω_κ selected with the use of an exactly phased Fourier synthesis depend on the synthesis resolution and on the chosen cut-off level. If the resolution is low and the cut-off level is high enough, it is expected that the region Ω_κ consists of a small number of 'blobs' corresponding to individual molecules. When the cut-off level is chosen lower and lower, at some moment these blobs merge into endless continuous regions. For medium resolution, a high-value region may represent the trace of the polypeptide chain, while for very high resolution it may consist of peaks corresponding to individual atoms. The full connectivity analysis of a function $\rho(\mathbf{r})$ consists in the study of connectivity properties for regions Ω_κ corresponding to different resolutions of Fourier syntheses and to different cut-off levels κ. We consider below a simpler case when the resolution and cut-off level are fixed.

The function $\rho(\mathbf{r})$ may be calculated on different scales and with the use of different weights for individual reflections, so it is convenient to define the cut-off level in some invariant manner. For the analysis of the Fourier syntheses at low resolution, we found it convenient to fix a given volume per residue. We say that the cut-off level $\kappa = \kappa(\alpha)$ corresponds to the specific volume α (Å^3 per residue) and denote the corresponding region as $\Omega^\alpha = \Omega_{\kappa(\alpha)}$ if

$$\frac{\text{volume of } \Omega_{\kappa(\alpha)}}{\text{number of residues in the unit cell}} = \alpha. \quad (2)$$

If the value of α is fixed, then the scale of the observed magnitudes affects the $\kappa(\alpha)$ value but does not change the $\Omega_{\kappa(\alpha)}$ region. Obviously, other invariant characteristics may be introduced (*e.g.* volume per atom or volume per electron *etc.*) but the ratio 'volume per residue' seems to be more convenient in practical applications.

3. Connectivity-based selection criterion

The study of low-resolution syntheses for different macromolecules (Table 1) confirms the hypothesis that for a high-enough cut-off level the regions Ω_κ usually (but not always) are formed by a small number of compact finite regions and the number of these regions is equal to the number of molecules in the unit cell. Naturally, the regions linked by crystallographic symmetry have exactly the same volume. If noncrystallographic symmetry is present, then the regions may be slightly different in shape and volume owing to series-truncation effects. The tests with known structures have shown that the specific volume 25 Å^3 per residue is usually suitable to isolate one connected region per molecule if about 15 low-resolution reflections per molecule were used when calculating the Fourier synthesis. It must be noted that these regions do not correspond to what is usually considered as a 'molecular envelope'. They are small and cover central regions of the molecules only. To trap almost all the atomic positions, the region Ω_κ must have about 200 Å^3 per residue specific volume.

As a result, the following property of well-phased low-resolution syntheses may be postulated:

The selection rule. The number of connected components in the high-values region Ω^α is equal to the number of molecules in the unit cell. The components must have the same volume in

Figure 1
Two-dimensional example of the connectivity analysis. The shaded Ω_κ region consists of two finite (A and B) and one infinite (C) connected components.

Table 1
Unit-cell parameters and connectivity analysis for test macromolecular crystals.

The analysed Fourier syntheses were obtained with the observed magnitudes and the phases calculated from the refined atomic model. The presence of endless connected components is marked by (∞).

Protein	γ-Crystallin IIIb	Elongation factor G	Protein G	RNAase *sa*
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Unit-cell dimensions (Å)	58.7	75.9	34.9	64.9
	69.5	105.6	40.3	78.32
	116.9	115.9	42.2	38.79
No. of independent molecules	2	1	1	2
No. of residues per molecule	173	689	61	96
Synthesis resolution (Å)	24.	36.	15.	18.
No. of independent reflections	28	16	15	29
Grid numbers	$30\times32\times60$	$20\times24\times24$	$18\times20\times20$	$32\times40\times20$
Specific volume (Å³ per residue)	No. of grid points in connected regions			
5.	$4\times106+4\times100$	4×45	4×38	$4\times119+4\times5$
10.	$4\times211+4\times207$	4×88	$4\times71+4\times3$	$4\times209+4\times39$
15.	$4\times319+4\times308$	4×129	$4\times98+4\times13$	4×376
20.	$4\times420+4\times410$	4×170	4×129	4×496
25.	$4\times526+4\times521$	4×215	4×187	4×618
30.	2×2502 (∞)	2×526 (∞)	4×217	4×745
35.	1×5860 (∞)	1×1200 (∞)	2×532 (∞)	2×1738 (∞)
40.	1×6672 (∞)	1×1368 (∞)	2×598 (∞)	2×1984 (∞)

the absence of non-crystallographic symmetry otherwise the volume is allowed to be slightly different.

It must be noted that some exceptions from this rule are possible (see RNAase column in Table 1) when the peaks corresponding to neighbouring molecules overlap owing to close intermolecular contacts. Nevertheless (see §4.4), even in such cases the use of this (idealized) selection criterion may result in useful phase information.

4. Connectivity-based low-resolution phasing

4.1. Test objects

To study the phasing power of the connectivity-based selection criterion, several structures with known atomic models were used. We discuss below the results of tests with only two samples, namely γ-crystallin IIIb (Chirgadze *et al.*, 1991) and RNAase *sa* (Ševčic *et al.*, 1991) since the results of tests with other structures are very similar to one of these two cases. Parameters for the test structures are listed in Table 1. Experimental structure-factor magnitudes were used in both cases. To check the results, the phases calculated with the refined models were considered as the exact ones.

In what follows, we call any considered phase set a 'variant' and any set of variants a 'population'. A random population is formed by a large number of variants obtained by means of random and independent generation of all phase values. (Obviously, the symmetry restrictions on possible values of phases for centric reflections must be taken into account.) A selected population consists of all variants from the random population that satisfy the connectivity-based selection criterion formulated above.

4.2. Connectivity-based selection procedure

Table 1 shows that in the case of γ-crystallin the Fourier synthesis calculated with the observed magnitudes and the exact phases satisfies the selection criterion. So, the first obvious question is whether the formulated selection procedure allows one to pick up variants that are close to the true structure-factor phases starting from a random population. There are various ways to define 'close phase sets' and the choice depends on the problem under study. In this paper, we define the measure of closeness of two phase sets as the correlation of two *aligned* Fourier syntheses calculated with these phase sets and the same experimental magnitudes (see Appendix *A*). In what follows, we consider a phase variant as a good one if it has large enough correlation with the exact phases and as a bad variant otherwise. The alignment of two maps before their comparison (Lunin & Lunina, 1996) is crucial since there is no reason for the random phase sets to correspond to the same choice of the permitted unit-cell origin/enantiomorph.

To answer the initial question on the possibility of identifying good variants by the connectivity criterion, 100 000 phase sets (each consisting of all 28 independent phases corresponding to 24 Å resolution) were generated randomly and then the variants satisfying the selection criterion (576 variants) were picked out. The distribution of the variants in the random and selected populations with their correlation with the exact phases (Fig. 2*a*) shows that the selected population contains variants of very different phase quality. In particular, some of them are significantly different from the exact phases. Furthermore, a large number of reasonable variants were lost in the selection as they did not possess the desired connectivity (Table 2). This means that the selection criterion does not allow selection of the good variants

research papers

The following observations can be noted:
  (i) the best phase sets (the ones with the correlation above 0.9 in the case of γ-crystallin) were not selected;
  (ii) the highest percentage of the selected variants was reached for the batches with correlation between 0.8 and 0.9; it was similar for the two objects and was much higher than the mean value (0.6); the success rate for the variants from the previous column, with a correlation between 0.7 and 0.8, is similar for the two cases and is also higher than the mean value.

	Phase correlation										
Phase-set populations	0.0–0.1	0.1–0.2	0.2–0.3	0.3–0.4	0.4–0.5	0.5–0.6	0.6–0.7	0.7–0.8	0.8–0.9	0.9–1.0	Total
γ-Crystallin											
Random	0	0	0	0	1335	28121	48644	19776	2101	23	100000
Selected	0	0	0	0	3	52	251	222	48	0	576
% of selection	0	0	0	0	0.2	0.2	0.5	1.1	2.3	0	0.6
RNAase											
Random	1	1474	12614	21197	20028	21529	17801	5058	298	0	100000
Selected	0	0	9	9	65	209	199	61	7	0	559
% of selection	0	0	0.1	0.0	0.3	1.0	1.1	1.2	2.3	0	0.6

unambiguously. It may select bad variants and lose good ones. A Fourier synthesis corresponding to one of phase sets resulting in the desired connectivity but very different from the true phases is shown in Fig. 3(b).

4.3. Enrichment and averaging

While the performed analysis removes the hope of finding a reasonable solution of the phase problem immediately, Fig. 2 shows nevertheless an encouraging feature of the selected population. There is a significantly higher concentration of good variants in selected populations than in random ones. So, the connectivity-based selection procedure produces a kind of 'enrichment' of a random population by increasing the share of good variants in the population. It is likely that the average of the maps corresponding to the selected variants can give a reasonable approximation to the true map. This effect is schematically illustrated in Fig. 4.

The averaging of 576 variants selected in the test with γ-crystallin IIIb has resulted in the map possessing a correlation coefficient of 0.87 in comparison with the exactly phased map at 24 Å resolution. Some sections of the averaged map overlapped with the atomic positions are shown in Figs. 5 and 6. The use of a relatively high cut-off level allows one to see the peaks corresponding to centres of all eight dimeric molecules in the unit cell (Fig. 5). A lower cut-off level (Fig. 6) results in a joint region of all molecules rather than separated regions corresponding to individual molecules. In the space group $P2_12_12_1$, the phase values for four strong reflections can be fixed in advance to specify the origin/enantiomorph so that these phases may be considered as known. These reflections may dominate when calculating the map correlation coefficient producing an artificially high correlation value. To eliminate the influence of these reflections, the map correlation coefficient was calculated for the set of 24 newly phased reflections and was found to be as great as 0.77.

4.4. Test ab initio phasing: RNAase sa

The main difference between the γ-crystallin and RNAase cases is that the Fourier synthesis calculated with the observed magnitudes and the exact phases satisfies the selection criterion for the case of γ-crystallin and does not satisfy it for the case of RNAase. For RNAase, the synthesis does not reveal the same number of connected regions as the number of
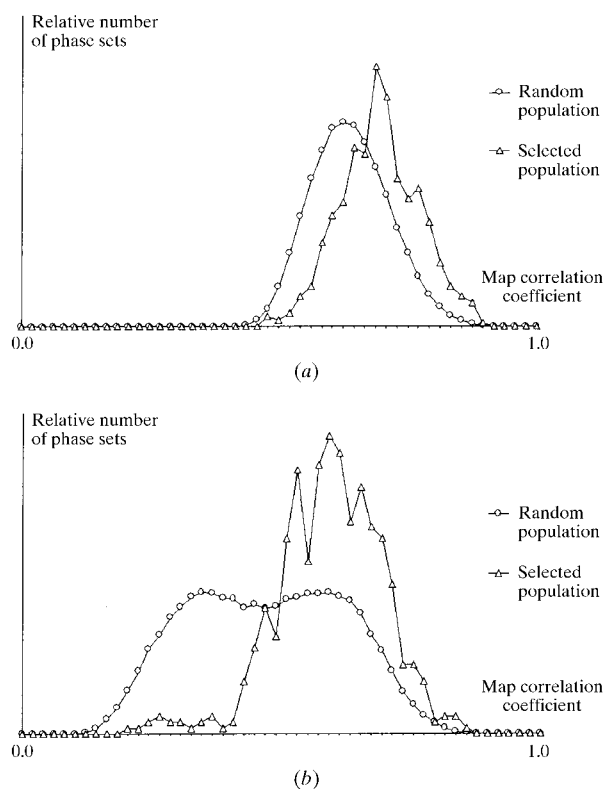

Figure 2
Distribution of phase variants in accordance with the map correlation coefficient in random (circles) and selected (triangles) populations: (a) γ-crystallin IIIb; 24 Å resolution data; (b) RNAase sa; 18 Å resolution data.
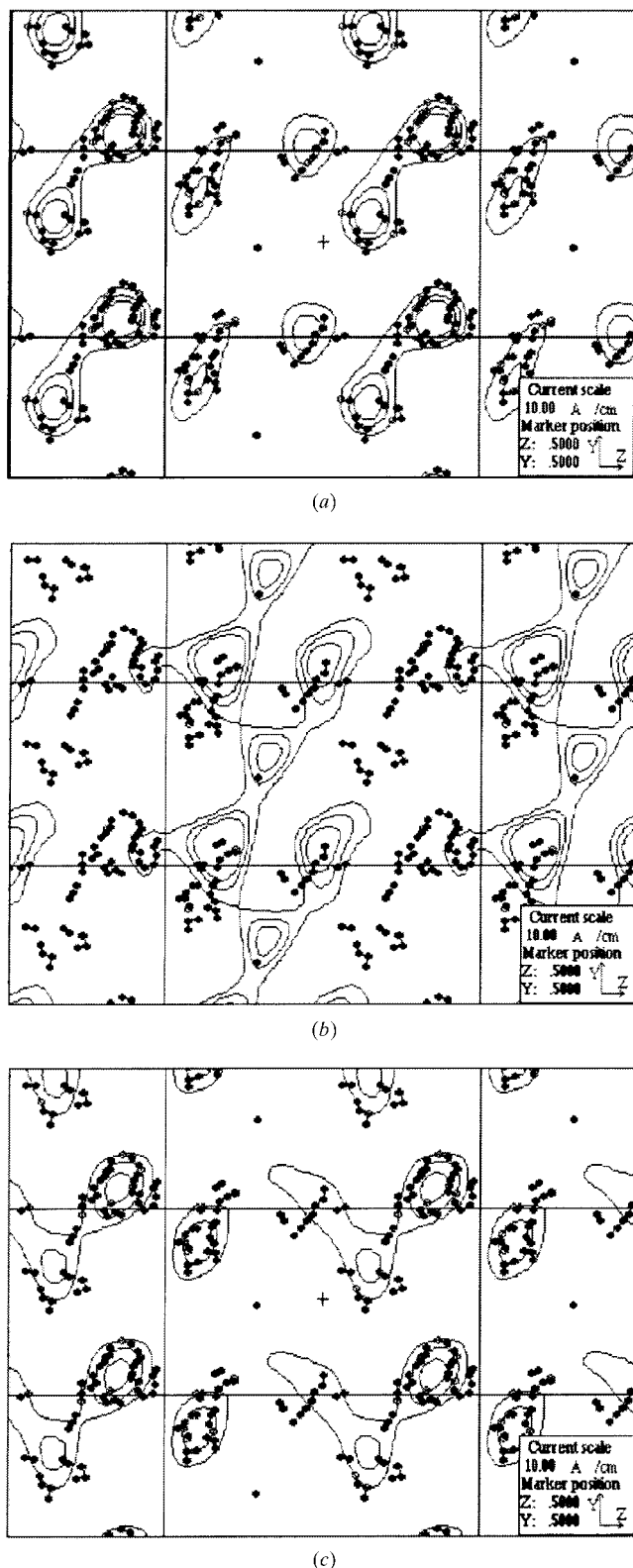
Figure 3
C_α-atom positions in γ-crystallin IIIb overlapped with the sections $x = 2/32$ of the 24 Å resolution Fourier syntheses: (*a*) the exactly phased synthesis; (*b*) an example of a randomly phased synthesis that satisfies the selection criterion (eight finite connected components at 25 Å³ per residue cut-off) but nevertheless has a low correlation (0.49) with the exact synthesis; (*c*) the result of averaging the connectivity-selected syntheses (576 syntheses from 100 000 generated). The lowest shown cut-off level corresponds to 100 Å³ per residue.

molecules (Table 1). So, when the connectivity-based selection procedure is applied to the random variants, the ones that are very close (or equal) to the exact phases will be lost. Nevertheless, the share of the lost bad variants can be essentially higher, which could enrich the selected population. It is worth remembering that in the case of γ-crystallin the essential part of the good variants was also lost in the selection (Table 2) but the result was the enrichment of the population.

Fig. 2(*b*) shows the distribution of variants in a random and in the corresponding selected populations with their correlation with the exact phases at 18 Å resolution. Similar to the case of γ-crystallin, the selection procedure did not allow selection of the good variants unambiguously and found also a number of wrong solutions. At the same time, some good variants (Table 2) were lost in the selection. Nevertheless, the tendency to enrichment of the population was very strong again.

The averaging of 558 variants selected in the test produced the synthesis possessing a correlation coefficient of 0.75 in comparison with the exact 18 Å synthesis. The map correlation coefficient was equal to 0.44 for the 25 newly phased reflections in the 18 Å resolution zone and was equal to 0.73 for the nine newly phased reflections in the 24 Å resolution zone. Some sections of the averaged synthesis are shown in Fig. 7. Again, the highest peaks in the *ab initio* phased synthesis correspond to eight molecule positions, while a lower cut-off level isolates a joint region composed by all molecules.

5. Discussion

The exploited property of well phased Fourier syntheses to reveal compact 'blobs' corresponding to molecules is the simplest example of what may be called topological properties.
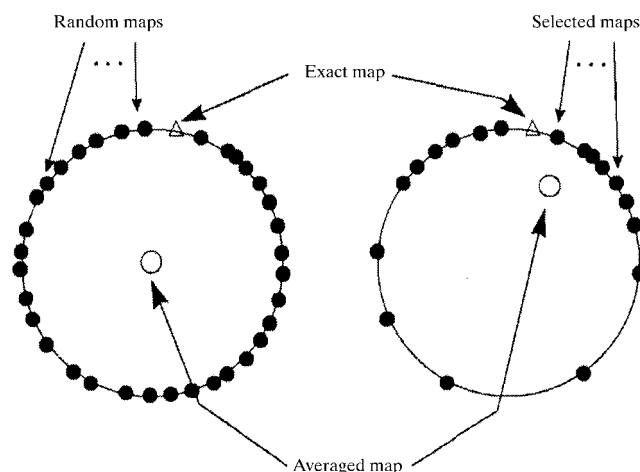

Figure 4
Schematic representation of the results of an enrichment procedure. Every point corresponds to a Fourier-synthesis map. The set of selected points contains those that are quite different from the exact solution. Nevertheless, the concentration of 'good' points is higher in the selected set than in the initial one and the averaged variant is close to the true solution.

research papers

Nevertheless, even this simplest property has some phasing power. At the same time, it has undesirable features that are common with those of some other low-resolution selection criteria, such as Fourier synthesis histograms (Lunin *et al.*, 1990), correlation of a few atom model magnitudes with the observed ones (Lunin *et al.*, 1995), or statistical likelihood (Petrova *et al.*, 1999):

(i) there exist 'bad' phase sets that satisfy the considered selection criterion but are too far from the correct solution;

(ii) there exist phase sets that are close to the exact phases but do not satisfy the selection criterion.

On the other hand, there is a tendency to have a higher percentage of good phase sets in the selected population than in a random one, therefore the following phasing procedure may be applied to reasonably small low-resolution data sets:

(i) generation of a large number of random independent phase sets;

(ii) selection of the phase sets that, when coupled with the observed magnitudes, result in Fourier syntheses with the desired properties;

(iii) averaging the selected variants.

Some modifications may be introduced into the process. First, the random generation of a phase set is not the only way to explore possible phase combinations. Alternative ways may be:

(i) to perform a complete study of all phase combinations (Lunin *et al.*, 1999);

(ii) to explore some 'regular grids' in the space of all phase sets, *e.g.* by means of the use of error correcting codes (Woolfson, 1954; Gilmore *et al.*, 1999), which allow a significant reduction in the number of checked phase combinations.
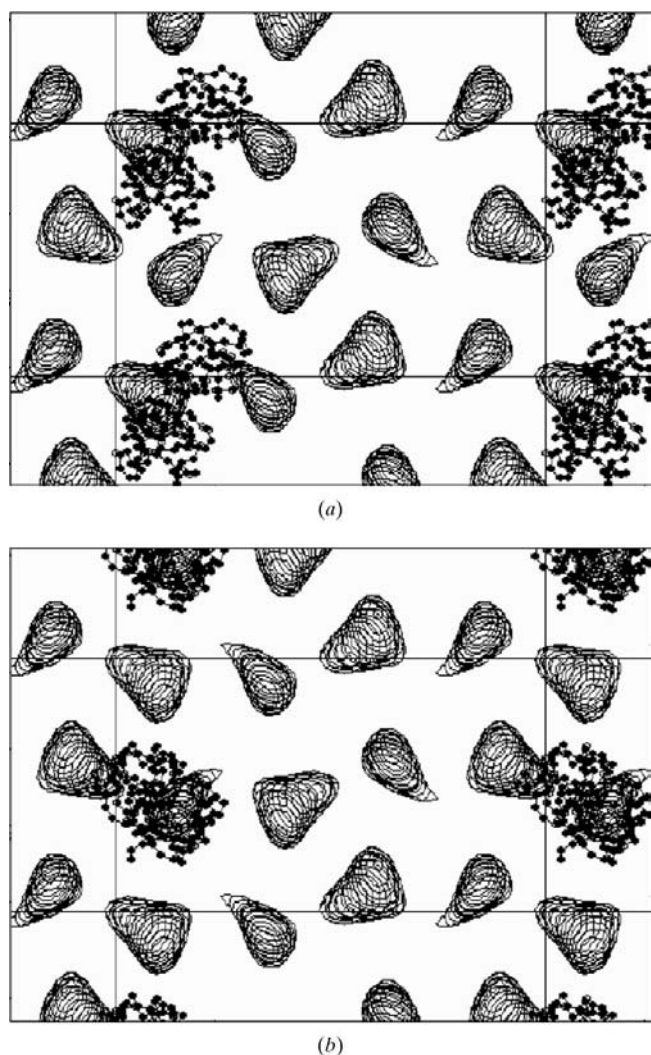

Figure 5
The *ab initio* phased synthesis for γ-crystallin IIIb at 24 Å resolution overlapped with the positions of C$_\alpha$ atoms in the projection 0yz: (*a*) one copy of the heterodimer *A*; (*b*) one copy of the heterodimer *B*. The synthesis was obtained by averaging 576 syntheses selected from 100 000 randomly generated ones. The symmetry-linked molecules are not present in the figure. 15 Å^3 per residue cut-off-level contour is chosen to indicate the molecular positions.
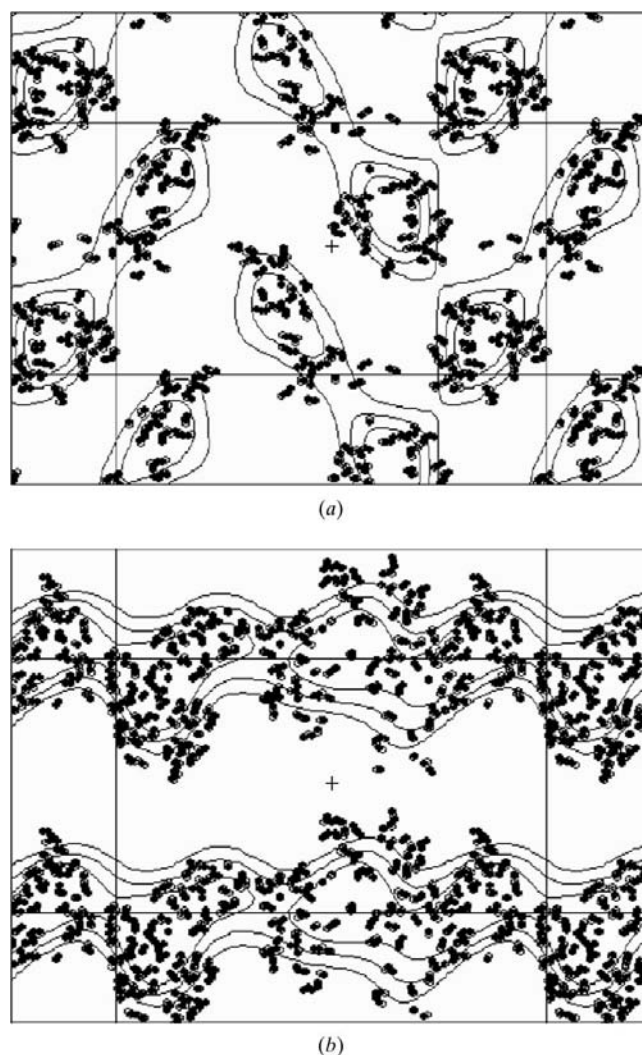

Figure 6
Two sections of the *ab initio* phased synthesis for γ-crystallin IIIb at 24 Å resolution overlapped with all atom positions. The synthesis was obtained by averaging 576 syntheses selected from 100 000 randomly generated ones. The cut-off levels correspond to the specific volumes 150, 100 and 50 Å^3 per residue. (*a*) Section $x = 0$; (*b*) section $x = 6/32$.

Second, a more detailed analysis of the selected population may be worthwhile. The selected population may have additional features of distribution of variants in the configurational space. Sometimes, it may form several clusters of variants and it is reasonable to perform an averaging for every cluster separately rather than average all selected variants together (Lunin *et al.*, 1990, 1995).

APPENDIX *A*
The map correlation coefficient

We consider an interpretable Fourier synthesis as the main goal of low-resolution phasing. So the closeness of two phase sets $\{\varphi_1(\mathbf{h})\}_\mathbf{h}$ and $\{\varphi_2(\mathbf{h})\}_\mathbf{h}$ is estimated through the closeness of the two Fourier syntheses $\rho_1(\mathbf{r})$ and $\rho_2(\mathbf{r})$ calculated with these phase sets and the same set $\{F^{obs}(\mathbf{h})\}_\mathbf{h}$ of observed magnitudes.

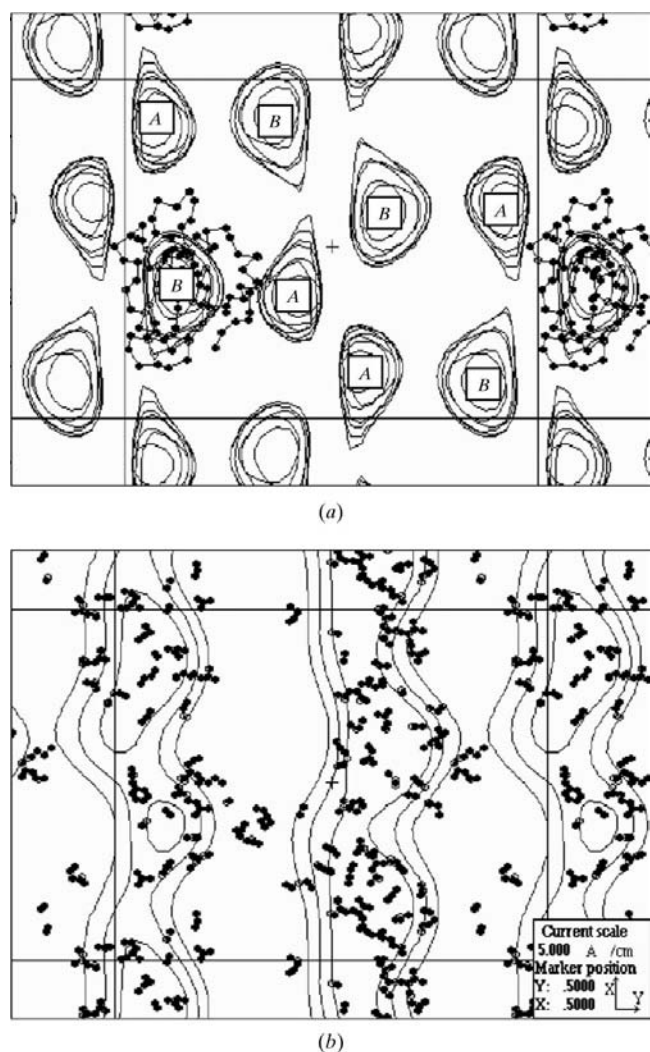A formal map correlation coefficient (Lunin & Woolfson, 1993) is defined as

$$
\begin{aligned}
CP\{\rho_1, \rho_2\} &= CP(\{\varphi_1(\mathbf{h})\}, \{\varphi_2(\mathbf{h})\}) \\
&= \frac{\int \rho_1(\mathbf{r})\rho_2(\mathbf{r})\,dV_\mathbf{r}}{[\int \rho_1(\mathbf{r})^2\,dV_\mathbf{r} \int \rho_2(\mathbf{r})^2\,dV_\mathbf{r}]^{1/2}} \\
&= \frac{\sum_\mathbf{h} F^{obs}(\mathbf{h})^2 \cos[\varphi_1(\mathbf{h}) - \varphi_2(\mathbf{h})]}{\sum_\mathbf{h} F^{obs}(\mathbf{h})^2}.
\end{aligned} \tag{3}
$$

Here and below, we suppose that all the maps and all structure-factor sums are calculated without the F_{000} term.

It must be taken into account when studying *ab initio* phasing procedures that some phase sets while being formally different (and possessing low CP values) may result in similar maps that differ only by a permitted origin shift (and/or by the enantiomer choice). Such maps must be considered as equivalent ones, so the map alignment must be performed before the formal measure of their closeness is calculated (Lunin *et al.*, 1990; Hašek & Schenk, 1992; Lunin & Lunina, 1996). We define the crystallographic map correlation coefficient as the maximal value of the formal map correlation coefficient that can be attained when comparing $\rho_1(\mathbf{r})$ and $\rho_2(s\mathbf{r} + \mathbf{t})$ syntheses, where $s = \pm 1$ if both enantiomorphs are permitted ($s = 1$ otherwise) and $\mathbf{t}$ runs over all permitted origins.

$$
\begin{aligned}
CCP\{\rho_1, \rho_2\} &= CCP(\{\varphi_1(\mathbf{h})\}, \{\varphi_2(\mathbf{h})\}) \\
&= \max_{s=\pm1,\mathbf{t}} CP\{\rho_1(\mathbf{r}), \rho_2(s\mathbf{r} + \mathbf{t})\}.
\end{aligned} \tag{4}
$$

It may be surprising that, as follows from Fig. 2, it is practically impossible to generate random phases that have negative correlation with the true ones in the space group $P2_12_12_1$. The fact is that in this space group there exist 16 possibilities for the choice of the origin and enantiomorph, so it is always possible to find some alignment that results in a reasonable overlapping of high-value regions in the low-resolution Fourier synthesis maps.

Figure 7
The *ab initio* phased synthesis for RNAase at 18 Å resolution overlapped with: (*a*) C$_\alpha$ atoms of one of the symmetry-linked copies of molecule *B* in the projection $0xy$; 20 Å^3 per residue cut-off level; the peaks marked *A* correspond to the position of the molecule *A* (atoms are not shown); (*b*) all atoms in section $z = 4/20$; 150, 100 and 50 Å^3 per residue cut-off levels.

References

Baker, D., Krukowski, A. E. & Agard, D. A. (1993). *Acta Cryst.* D**49**, 186–192.
Chirgadze, Yu. N., Nevskaya, N. A., Vernoslova, E. A., Nikonov, S. V., Sergeev, Yu. V., Brazhnikov, E. V., Fomenkova, N. P., Lunin, V. Yu. & Urzhumtsev, A. G. (1991). *Exp. Eye Res.* **53**, 295–304.
Gilmore, C., Dong, W. & Bricogne, G. (1999). *Acta Cryst.* A**55**, 70–83.
Hašek, J. & Schenk, H. (1992). *Acta Cryst.* A**48**, 693–695.
Housset, D. (1999). *Ec. Phys. Chim. Vivant*, **1**, 9–18.
Jelsch, C., Pichon-Pesme, V., Lecomte, C. & Aubry, A. (1998). *Acta Cryst.* D**54**, 1306–1318.

Karle, J. & Hauptman, H. (1950). *Acta Cryst.* **3**, 181–187.

Lecomte, C. (1998). *NATO ASI and Euroconference*, NATO ASI. Dordrecht: Kluwer Academic Publishers.

Lunin, V. Y. & Lunina, N. L. (1996). *Acta Cryst.* A**52**, 365–368.

Lunin, V. Y., Lunina, N. L., Petrova, T. E., Vernoslova, E. A., Urzhumtsev, A. G. & Podjarny, A. D. (1995). *Acta Cryst.* D**51**, 896–903.

Lunin, V. Y., Lunina, N. L. & Urzhumtsev, A. G. (1999). *Acta Cryst.* A**55**, 916–925.

Lunin, V. Yu., Urzhumtsev, A. G. & Skovoroda, T. P. (1990). *Acta Cryst.* A**46**, 540–544.

Lunin, V. Yu. & Woolfson, M. M. (1993). *Acta Cryst.* D**49**, 530–533.

Petrova, T. E., Lunin, V. Y. & Podjarny, A. D. (1999). XVIIIth IUCr Congress, Glasgow, Scotland, Abstract P12.BB.005.

Podjarny, A. D. (1999). Personal communication.

Sayre, D. (1952). *Acta Cryst.* **5**, 60–65.

Ševčik, J., Dodson, E. & Dodson, G. (1991). *Acta Cryst.* B**47**, 240–253.

Woolfson, M. M. (1954). *Acta Cryst.* **7**, 65–67.