Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Low-resolution *ab initio* phasing of *Sarcocystis muris* lectin SML-2

Jürgen J. Müller,[a] Natalia L. Lunina,[b] Alexandre Urzhumtsev,[c] Edgar Weckert,[d] Udo Heinemann[a,e] and Vladimir Y. Lunin[b]*

[a]Macromolecular Structure and Interactions, Max-Delbrück-Centrum für Molekulare Medizin, Robert-Rössle-Strasse 10, D-13092 Berlin, Germany, [b]Institute of Mathematical Problems of Biology, Russian Academy of Sciences, Pushchino, Moscow Region 142290, Russia, [c]Département de Physique, Université H. Poincaré Nancy 1, 54506 Vandoeuvre-lès-Nancy, France, [d]Hasylab at DESY, Notkestrasse 85, D-22603 Hamburg, Germany, and [e]Institut für Chemie und Biochemie, Freie Universität, Takustrasse 6, D-14195 Berlin, Germany

Correspondence e-mail: jjm@mdc-berlin.de

Received 12 December 2005
Accepted 2 March 2006

Structural analysis of the lectin SML-2 faced difficulties when applying standard crystallographic phasing methods. The connectivity-based *ab initio* phasing method allowed the computation of a 16 Å resolution Fourier synthesis and the derivation of primary structural information. It was found that SML-2 crystals have three dimers in the asymmetric part of the unit cell linked by a noncrystallographic symmetry close to translation by (0, 0, 1/3). A clear identification of the noncrystallographic twofold axis explains the space-group transformation from the primitive $P2_12_12_1$ to the *C*-centred $C222_1$ observed during annealing procedures within an N_2 cryostream for cocrystals of SML-2 and galactose. Related packing considerations predict a possible arrangement of SML-2 molecules in a tetragonal unit cell. Multiple noncrystallographic symmetries and crystal forms provide a basis for further image improvements.

1. Introduction

Sarcocystis muris is an intracellular cyst-forming parasite that propagates in mice as the intermediate and cats as the final host (Müller *et al.*, 2001; Klein *et al.*, 2003). The parasite lectin SML-2 belongs to a family of three highly conserved galactose-specific lectins (Müller *et al.*, 2001). Entzeroth *et al.* (1992) demonstrated the secretion of these proteins during an early phase of host-cell invasion and localized the lectins at the moving junction, where they contribute to parasite motility. The biologically active unit is a homodimer (Montag *et al.*, 1997) and each monomer consists of 138 amino acids corresponding to a molecular weight of 15 066 Da.

Traditional methods of structure solution were unsuccessful. Because of a low sequence match to any other protein of known structure, molecular replacement failed. Furthermore, no heavy-atom phasing has yet been successful and selenomethionine derivatives were not available because the protein was prepared directly from the parasites (Müller *et al.*, 2001). At the same time, a complete very low resolution data set collected from the native crystal let us apply *ab initio* phasing methods (Lunin *et al.*, 2002) and obtain the first low-resolution structure images.

2. Experimental

Protein preparation and crystallization has been described previously (Müller *et al.*, 2001). The crystal, with dimensions of $0.2 \times 0.2 \times 0.2$ mm, belongs to the primitive orthorhombic space group $P2_12_12_1$ and diffracts to about 5–6 Å resolution at 100 K at the HASYLAB PETRA I undulator beamline at the PETRA storage ring, DESY, Hamburg. We used a six-circle goniometer (Huber, Germany) in combination with a 165 mm

Table 1
Statistics for native *S. muris* lectin SML-2 crystals and SML-2–galactose cocrystals.

Values in parentheses are for the outer shell.

Crystal	Native SML2	SML-2–galactose cocrystal	SML-2–galactose cocrystal
X-ray source	HASYLAB PETRA I, MAR CCD 165	Home, RU H2B, MAR 300	BESSY II, BL1, MAR CCD 165
Space group	$P2_12_12_1$	$C222_1$	$P2_12_12_1$
Unit-cell parameters			
a (Å)	53.30	74.70	53.49
b (Å)	129.48	81.97	130.71
c (Å)	157.63	131.00	159.86
Mosaicity (°)	0.69	0.37	0.37
Resolution (Å)	110.0–6.0 (20.0–6.0)	20.0–2.45 (2.54–2.45)	50.0–2.56 (2.70–2.56)
Observed/unique reflections	45384/2974	73314/15121	485661/36856
R_{sym} (%)	6.6 (6.9)	4.9 (15.7)	6.9 (18.7)
Completeness (%)	100.0 (98.8)	99.8 (99.0)	99.6 (98.3)
Average redundancy	See Table 2	4.85	13
$\langle I\rangle/\langle\sigma(I)\rangle$	32.7 (30.3)	6.9 (18.7)	28.7 (14.3)

MAR CCD detector at a 425 mm crystal-to-detector distance and with a non-focused parallel beam. A data set was collected with an overall R_{sym} of 6.6% to 6 Å resolution with 100% completeness to 10 Å and 98.9% to 6 Å (Table 1). More detailed statistics as determined by the data-evaluation program *XDS* (Kabsch, 1993) are given in Table 2.

Cocrystallization of the lectin SML-2 with a modified galactose (Müller *et al.*, 2001) resulted in a second crystal form, space group $C222_1$, with unit-cell parameters a = 74.70, b = 81.97, c = 131.00 Å (Table 1). After annealing in an N_2 cryostream, this crystal switched to space group $P2_12_12_1$, yielding a crystal isomorphous to that of the native apoprotein.

Table 2
Very low resolution data statistics calculated with *XDS* (Kabsch, 1993).

Resolution limit (Å)	No. of reflections: Observed	Unique	Possible	Completeness (%)	*R* factor (%)	$\langle I\rangle/\langle\sigma(I)\rangle$	R_{meas} (%)
100	4	1	1	100	5.4	27.9	6.2
70	3	1	1	100	43.6	4.5	53.3
50	27	5	5	100	39.7	15.0	43.7
30	227	26	26	100	3.9	31.8	4.3
20	964	70	70	100	9.2	32.9	9.6
10	10073	589	597	98.7	5.9	44.2	6.0
6	34086	2282	2308	98.9	6.9	30.3	7.1

Only one of multiple efforts to obtain heavy-atom derivatives was successful. A Pt-derivative data set (formal resolution 3.8–19.9 Å; practically no data lower than 16.6 Å resolution; detailed statistics not shown) was collected on a rotating-anode generator at home. The program *OASIS* (Hao *et al.*, 2000), with six heavy atoms per asymmetric part of the unit cell, followed by *DM* (Cowtan, 1994) from the *CCP*4 program suite (Collaborative Computational Project, Number 4, 1994) provided us with phase values for data in the resolution range 5.0–15.0 Å. The phasing power was very low and the corresponding Fourier syntheses were completely noisy and useless.

Further work with the native data set was performed using the program *MOLREP* (Vagin & Teplyakov, 1997), the *ab*

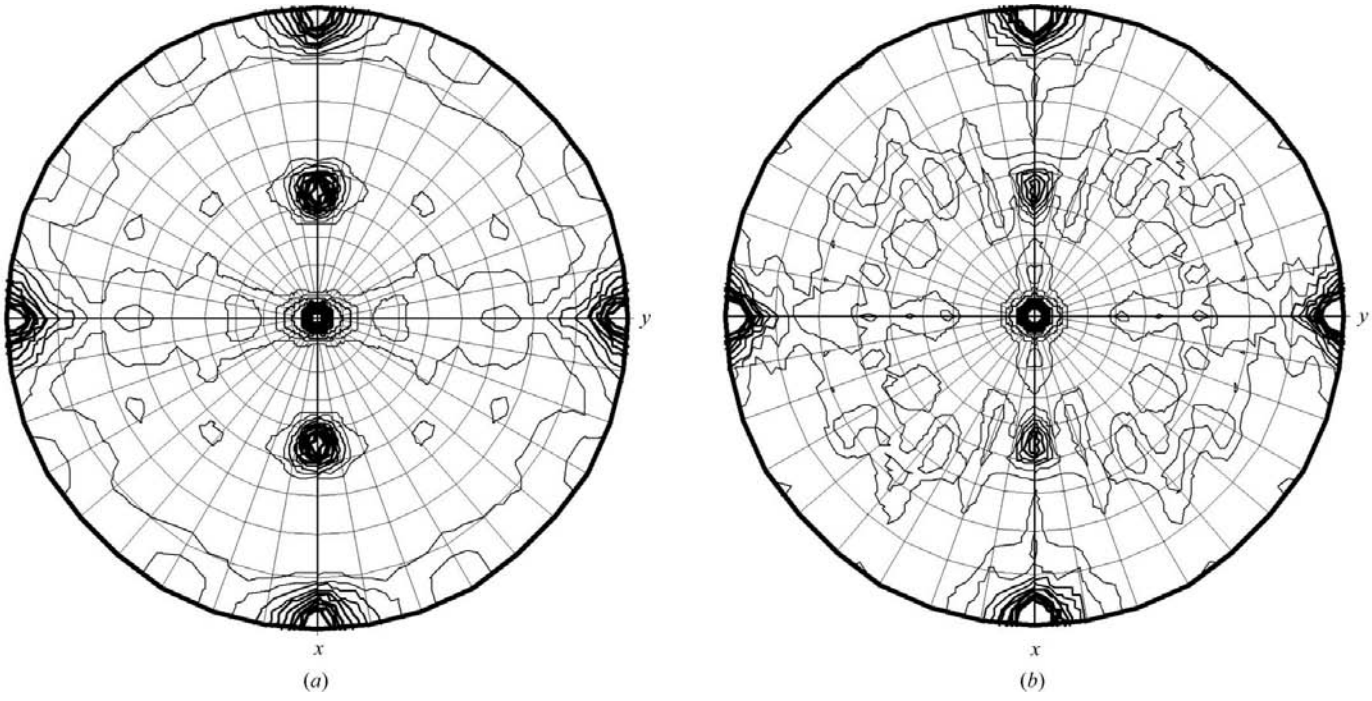

Figure 1
Self-rotation functions (*a*) for native SML-2 ($P2_12_12_1$) and (*b*) for native SML-2 cocrystallized with galactose ($C222_1$) at 6 Å resolution. The sections corresponding to rotation by $\kappa = 180°$ are shown.

Table 3
Mean figure of merit during *ab initio* phasing.

Resolution zone (Å)	∞–28.0	∞–25.0	∞–20.0	∞–16.0
No. of reflections	38	56	103	189
Phasing step				
1	0.26	0.21	0.17	0.13
2	0.31	0.26	0.23	0.18
3	0.34	0.30	0.27	0.22
4	0.36	0.32	0.31	0.25
5	0.39	0.36	0.34	0.28
6	0.41	0.38	0.36	0.30

initio phasing program *GENMEM* (Lunina *et al.*, 2003), the program *IMP* from the *RAVE* program suite (Kleywegt *et al.*, 2001) and the graphic programs *FFT-CAN* (Vernoslova & Lunin, 1993) and *PyMOL* (DeLano, 2002), as described below.

3. Preliminary analysis

For the Matthews coefficient to be in the range 2.3–3.0 Å^3 Da^{-1}, the asymmetric part of the unit cell of the $P2_12_12_1$ space-group crystals could contain 6–8 monomers. The protein content of the unit cell is then 41.2–54.2%. The unit cell in space group $C222_1$, being roughly 1.5 times smaller in volume, provides space for 4–6 monomers in the asymmetric unit.

The self-rotation function calculated by *MOLREP* with the available native data strongly indicated a noncrystallographic twofold axis orthogonal to the *b* axis and forming an angle with the *a* axis of approximately 45° (Fig. 1*a*). The observed height of 6.96σ for the corresponding self-rotation peak is comparable with the height of 7.60σ for the origin peak and exceeds the height of all other peaks by more than fourfold. It is worth noting that, together with the crystallographic 2_1 axis parallel to *a*, this noncrystallographic symmetry generates a fourfold noncrystallographic axis parallel to *b*, producing an equally strong peak in the self-rotation function in the $\kappa = 90°$ section (not shown) of the self-rotation function.

4. Connectivity-based *ab initio* phasing

Application of the connectivity-based *ab initio* phasing method to the lectin data was feasible owing to the exceptional completeness of this data set at low resolution. The details of this method have been described previously (Lunin *et al.*, 2000; Lunina *et al.*, 2003; Urzhumtseva *et al.*, 2004) and here we recall only that the main steps of the procedure are as follows.

(i) Large numbers of random phase sets are generated; in the initial run the phases are generated uniformly, while in the following iterations they are generated taking into account the values obtained previously and their reliability.

(ii) For each phase set the corresponding Fourier map is calculated with experimental magnitudes; in each map, the region of a relatively high density is analyzed, the connected components in this region are identified and their volumes are estimated.

(iii) The maps (and corresponding phase sets) that satisfy prescribed conditions (for example, a given number of connected components, absence of small 'noisy drops' *etc.*) are accepted for further analysis.

(iv) Optimal alignment is performed followed by averaging of selected images; the averaging produces approximate phase values and also their figures of merit; this information is used in the following iterations.

(v) Initially giving phases for a few tens of reflections, the procedure improves and extends this phase set.

5. *Ab initio* phasing of the native SML-2 data

To start the procedure of connectivity-based phasing, we should have some idea of what the high-density region looks like in terms of connectivity analysis. It is reasonable to expect that at low resolution molecules appear as isolated 'blobs' of approximately equal volume if an appropriate cut-off level is used (our experience suggests choosing the level resulting in a molecular mask with a volume of 25–30 Å^3 per residue). Nevertheless, if two molecules form, for example, a dimer with a close interface, then such dimer may be seen as a single 'blob' of larger volume. In the study of SML-2 the precise number of monomers (or dimers) in the unit cell was not known, so we tried phasing using different start hypotheses, namely two or three or four dimers in the asymmetric part of a molecular mask covering 15% of the unit-cell volume or four or six or eight monomers in a molecular mask covering 7% of the unit cell.

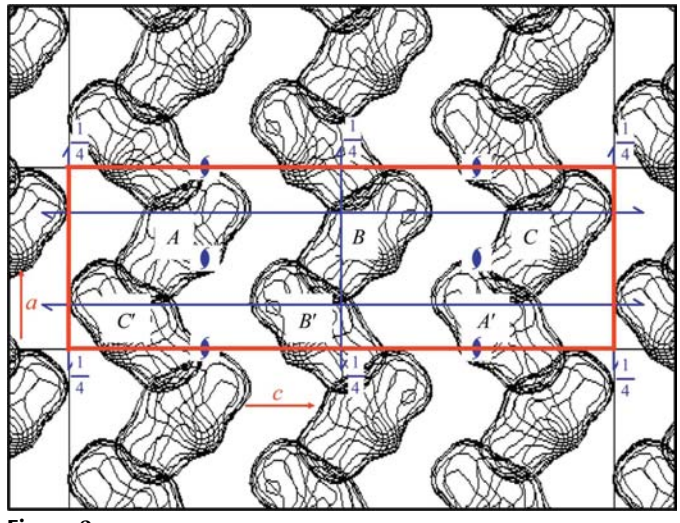

Figure 2
Fourier synthesis at 18 Å resolution calculated with the phases obtained after three cycles of connectivity-based *ab initio* phasing. Each 'blob' corresponds to a lectin dimer. The slice $0 \le y \le \frac{1}{2}$ is shown in projection along the *b* axis. The unit-cell border is shown in red and the $P2_12_12_1$ space-group symmetry in blue. The dimers *A*, *B* and *C* are linked by noncrystallographic symmetry close to the translation by (0, 0, 1/3). The dimers *A* and *A*′, *B* and *B*′, *C* and *C*′ are linked by crystallographic symmetry. The cutoff level of 1.0σ used isolates 15% of the unit-cell volume that corresponds to a specific volume of 50.0 Å^3 per residue.

The search for a phase set that gives a 21 Å resolution Fourier synthesis with three connected isolated components (supposed to correspond to three dimers) per asymmetric unit resulted in a number of images with three blobs of approximately the same volume. The search for two blobs resulted in a number of images with two blobs, one of them being twice as large than the other. This may be interpreted as three components of equal volume, two of them being merged. The search for four dimers resulted in a number of images with three blobs of nearly the same volume and an additional very small marginal 'drop'. This may be interpreted again as three equal blobs plus noise. The attempts to search for isolated monomers resulted in images with drops of very different volumes. These tests led us to accept, according to Urzhumtseva *et al.* (2004), the hypothesis of three dimers in the asymmetric unit seen as three blobs of connected density as a selection rule for initial steps of phasing.

In the first three steps of phasing, the selection rule was simply 'three blobs per asymmetric unit' for Fourier syntheses in several resolution zones. The particular resolutions were (27, 25, 22, 20 Å), (27, 25, 22, 20, 18 Å) and (27, 18, 17 Å), respectively, for these steps. At every step 100 selected phase sets were aligned and averaged, producing approximate phases and their individual figures of merit as the output. The alignment was carried out at resolutions of 21, 19 and 18 Å, respectively.

In the next three steps the additional requirement was introduced into the phasing that the three blobs have roughly the same volume. In addition, the resolution zone was extended to 15 Å (228 unique reflections). The results obtained are discussed in the next section. The variation of the figures of merit from step to step is shown in Table 3.

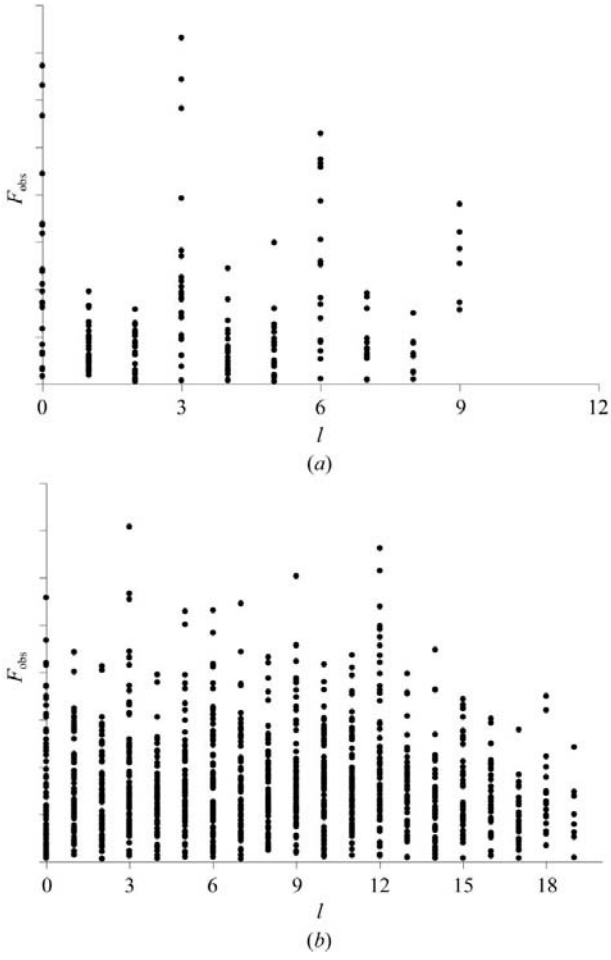

Figure 3
Observed structure-factor amplitudes *versus* *l*-index value. (*a*) $d \geq 16.0$ Å resolution zone; (*b*) $16.0 > d \geq 8.0$ Å resolution zone. Low-resolution data demonstrate relatively large values for h, k, $3n$ reflections. This is not the case for medium-resolution data.

6. Results

6.1. Molecular packing

The Fourier maps obtained after the first three steps of *ab initio* phasing showed the molecular packing very clearly (Fig. 2). These maps reveal three high-density regions in the asymmetric part of the unit cell that are similar in their shapes and volumes. It is worth noting that neither equivalent volumes nor similar shapes were demanded during these steps of phasing. The internal symmetry of the obtained high-density regions (this symmetry is discussed in more detail below) made us conclude that each of these regions corresponds to a lectin dimer. This was confirmed by the size of these regions when an appropriate density-cutoff level was chosen so that the corresponding envelopes touch each other.

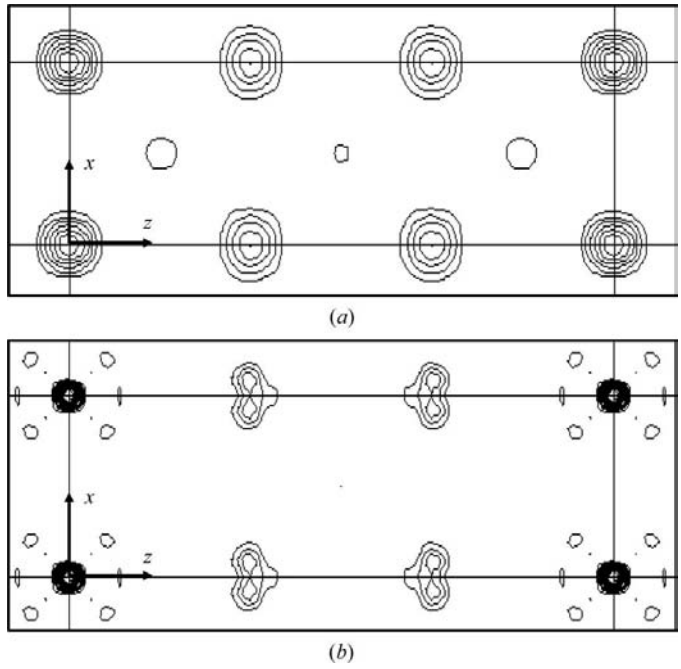

Figure 4
Section $y = 0$ of Patterson maps calculated with SML-2 data. (*a*) 16 Å resolution, contours starting from 2σ with step 2σ are shown; (*b*) 6 Å resolution, contours starting from 2σ with step 4σ are shown. Strong peaks in the 16 Å map, comparable in height with the peak at the origin, indicate pseudo-translation symmetry. These peaks are split in the 6 Å resolution map, indicating the noncrystallographic nature of the translation close to (0, 0, 1/3).

Table 4
Internal correlation of the Fourier syntheses calculated with *ab initio* phases at a resolution of 16 Å.

The table shows correlation of Fourier syntheses values in three regions presumably corresponding to the dimers *A*, *B* and *C* and also the internal twofold symmetry for each of these regions.

(*a*) Unweighted.

Unweighted	*A*	*B*	*C*
A	1.00	0.65	0.76
B	0.65	1.00	0.74
C	0.76	0.74	1.00
Internal	0.78	0.31	0.64

(*b*) Weighted.

Weighted	*A*	*B*	*C*
A	1.00	0.95	0.97
B	0.95	1.00	0.95
C	0.97	0.95	1.00
Internal	0.84	0.83	0.84

Therefore, the agreement of the number, shape and size of the obtained regions with independent experimental information can be considered as experimental evidence of the results in the situation when an atomic model is still unknown.

The three dimer regions demonstrate quasi-regular arrangement in the unit cell close to a translation by (0, 0, 1/3). This is taken to indicate that at low resolution this molecular packing may be described by a crystal with identical parameters *a* and *b* and a three times smaller parameter *c*. This 'hidden' additional periodicity in the *c* direction implies that the reflections with indices $h, k, 3n + 1$ and $h, k, 3n + 2$ should be negligible. Fig. 3 indicates a weak form of such a phenomenon in the 16 Å resolution zone and its absence at higher resolution. As a result, the two noncrystallographic symmetries close to (0, 0, 1/3) and (0, 0, 2/3) are indeed approximate and noncrystallographic (with a possible rotation component) rather than additional exact periodicity. The presence of quasi-periodicity in the crystal is confirmed by Patterson syntheses (Fig. 4) that demonstrate prominent peaks at (0, 0, 1/3) and (0, 0, 2/3). In a 16 Å resolution map these peaks have a height of 10σ (compared with 14σ for the origin peak) and are about five times stronger than the other peaks. In a 6 Å resolution map these peaks are of 14σ height and still dominate all other peaks, with the exception of a 83σ origin peak. However, now they are split, demonstrating a component along the *a* axis: a small shift additional to (0,0, 1/3) and (0, 0, 2/3). We used the *ab initio* phased syntheses and the program *IMP* (Kleywegt *et al.*, 2001) to refine noncrystallographic symmetry transformations. This analysis indeed revealed the presence of a rotation component of about 5°.

The same program *IMP* estimated the similarity of these regions of Fourier syntheses numerically (Table 4). The corresponding density correlation was high, especially when the weighted synthesis was analysed. A lower correlation in the non-weighted synthesis and lower values of figures of merit for higher resolution reflections indicate the need for further refinement of corresponding phase values.

6.2. Noncrystallographic dyad

Visual analysis of the shape of three dimer regions linked by noncrystallographic symmetry shows that each of them has an approximate twofold symmetry (Figs. 5 and 6). It is present both at low- and high-density cutoff levels, while some breaks in the symmetry caused by phase errors are observable at high levels. The directions of the dyads for the three crystallographically independent dimers are close to each other and coincide with the direction of the rotation axes indicated by the self-rotation function for $\kappa = 180°$ (Fig. 1). It is worth

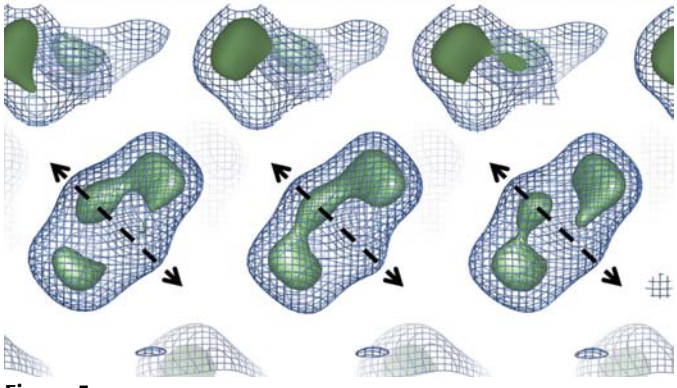

Figure 5
Three-dimensional image of three dimers. Fourier synthesis (18 Å resolution) calculated with phases obtained after three cycles of connectivity-based *ab initio* phasing viewed following the *b* axis. Three dimers linked by noncrystallographic symmetry are shown. The blue contour corresponds to a 1.0σ cutoff level and the green contour corresponds to 2.0σ. The image demonstrates the approximate symmetry of the dimers. The high-level region is broken in two parts for two of three dimers owing to the phase errors present. Broken arrows indicate the direction of the noncrystallographic dyad as defined by the self-rotation function.

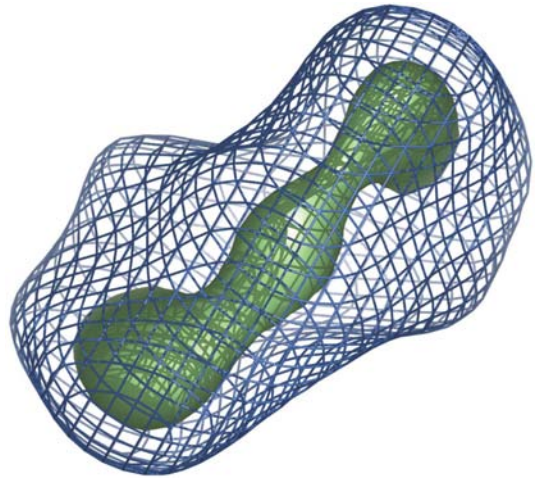

Figure 6
Three-dimensional image showing a lectin dimer. Fourier synthesis (18 Å resolution) calculated with phases obtained after three cycles of connectivity-based *ab initio* phasing looking down the direction of the twofold noncrystallographic symmetry axis as defined from the self-rotation function. The blue contour corresponds to a 1.0σ cutoff level; the green contour corresponds to 2.0σ.

noting that no condition relevant to this symmetry was included in the phasing procedure. We concluded that the appearance of this symmetry in the images obtained may serve as a further indicator of the correctness of the phase values obtained by our procedure.

This internal symmetry detected by a visual analysis of maps has been confirmed numerically (Table 4), especially when the proper weighting is used.

6.3. Switch between crystal forms

To simplify the presentation, in this section we use the notation $^{[a]}2_1$, $^{[b]}2_1$ and $^{[c]}2_1$ to distinguish the 2_1 axes parallel to the *a*, *b* and *c* coordinate axes, respectively. *a*′, *b*′ and *c*′ represent the coordinate axes in the new coordinate system and *x*′, *y*′, *z*′ the corresponding coordinates.

The observed molecular packing suggests an explanation of the possible switches between the crystal forms for lectin SML-2. Firstly, pseudo-translation symmetry suggests the existence of crystals with a threefold-shortened *c* axis, *i.e.* unit-cell parameters of around 53, 130 and 53 Å. These crystals possess $P2_12_12_1$ space-group symmetry. The asymmetric part of the unit cell contains two monomers that form a dimer through a noncrystallographic dyad.

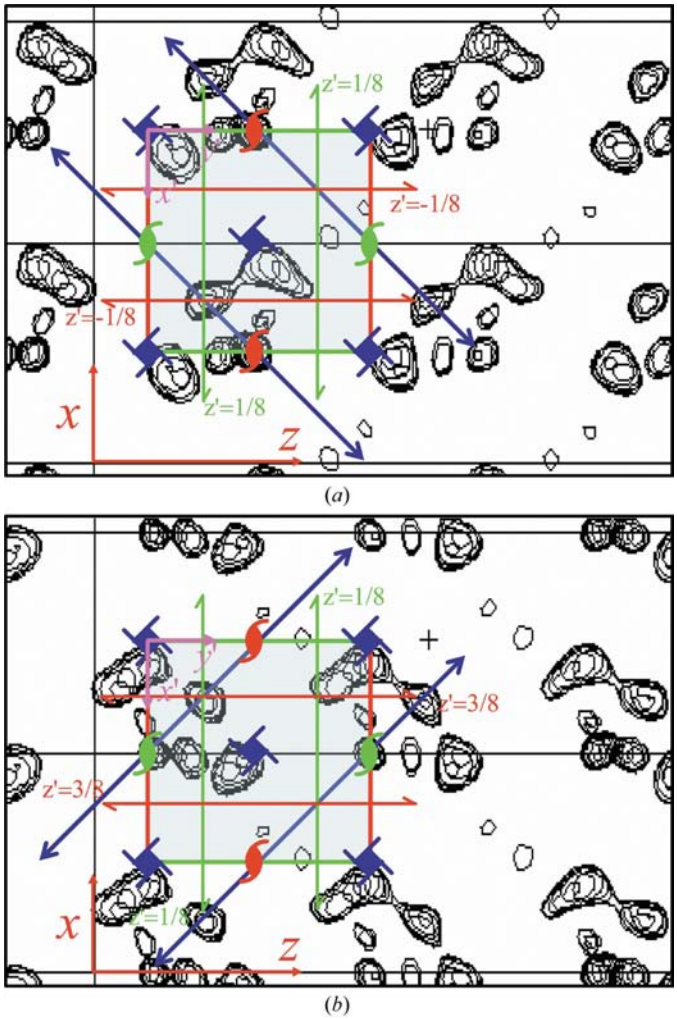

Figure 7
Fourier synthesis map demonstrating an approximate $P4_32_12$ space-group symmetry. One of the possible choices for the unit cell is indicated. Two subsequent slices (*a*) $-1/8 \le z' \le 1/8$ and (*b*) $1/8 \le z' \le 3/8$ are shown. Symmetry elements of the initial $P2_12_12_1$ space group are in red, symmetry elements induced by noncrystallographic translations are in green and symmetries induced by a noncrystallographic dyad are in blue; (*x*′, *y*′, *z*′) are coordinates in the $P4_32_12$ coordinate system. The map contours are shown for 16 Å resolution. The Fourier synthesis was obtained after six steps of *ab initio* phasing. The cutoff level corresponds to 2.0σ (13 Å^3 per residue).

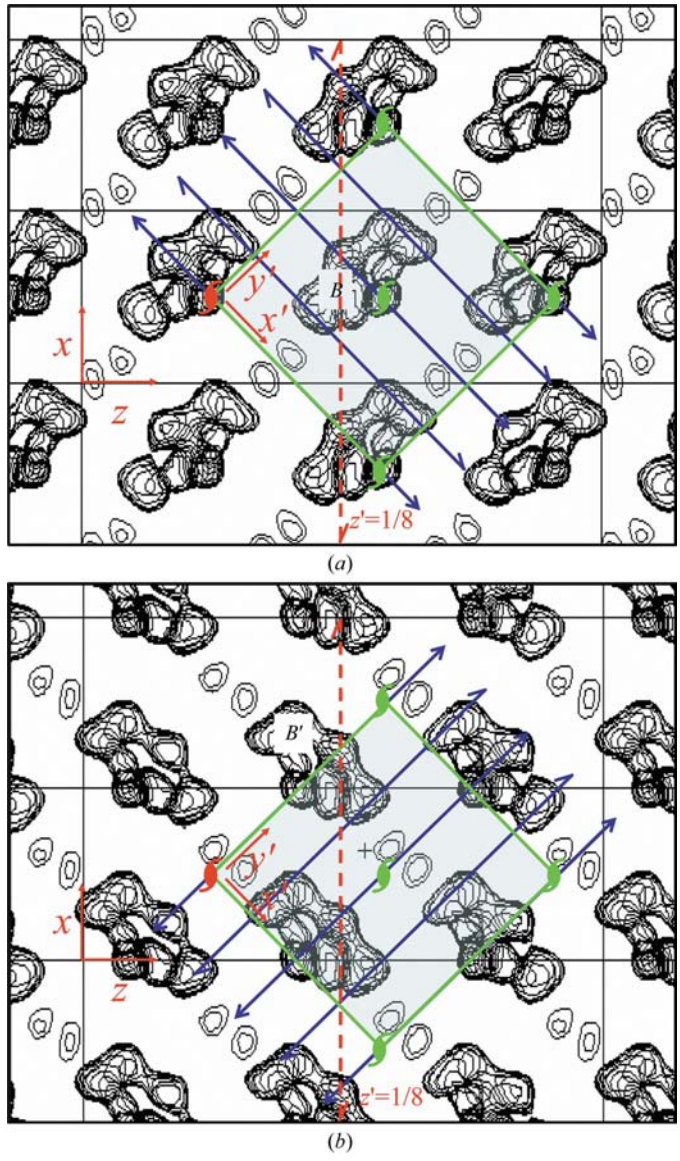

Figure 8
Fourier synthesis map demonstrating an approximate $C222_1$ space group symmetry. Two subsequent slices (*a*) $-1/8 \le z' \le 1/8$ and (*b*) $1/8 \le z' \le 3/8$ are shown. Symmetry elements of the initial $P2_12_12_1$ space group are shown in red, symmetry elements induced by noncrystallographic translations are in green and symmetries induced by noncrystallographic dyad are in blue; (*x*′, *y*′, *z*′) are coordinates in the $C222_1$ coordinate system. Dimers *B* and *B*′ are now linked by noncrystallographic symmetry close to the former $^{[a]}2_1$ crystallographic axis (shown by a dashed red line). The map contours are shown for 18 Å resolution Fourier synthesis as obtained after three steps of *ab initio* phasing. The cutoff level corresponds to 1.5σ (36 Å^3 per residue).

The self-rotation function and analysis of *ab initio* phased Fourier maps define the dimer's dyad as being in the plane parallel to the *ac* plane and making an angle of about 45° with the *a* axis. Furthermore, the map analysis places this dyad into the plane $y = 1/8$. Together with the $P2_12_12_1$ symmetry, this results in $P4_32_12$ space-group symmetry (Fig. 7) with the same unit-cell parameters of about 53, 53 and 130 Å, but now with a single monomer per asymmetric unit. This means that modification of crystallization conditions may lead to a small shift and rotation of dimers, resulting in crystals with $P4_32_12$ symmetry. Taking into account that at this stage of phasing the enantiomer choice is free, space group $P4_12_12$ must be considered as a possible alternative to $P4_32_12$.

Small rotations and shifts of dimers may result in one further different crystal form. Let us suppose that as a result of a small modification of molecule packing, the noncrystallographic translations and dyad become true crystallographic symmetry. This dyad, which initially passes very close to the crystallographic $^{[b]}2_1$ axis, may intersect this axis exactly after such a minor rearrangement. Let us suppose too that this rearrangement slightly distorts two other twofold crystallographic axes $^{[a]}2_1$ and $^{[c]}2_1$, making them noncrystallographic. In this case, the structure adopts the $C222_1$ space-group symmetry (Fig. 8), with unit-cell parameters of about 75, 75 and 130 Å. Now, the asymmetric part of this unit cell contains four monomers linked by noncrystallographic symmetry operators that are close to the former crystallographic axes $^{[a]}2_1$ and $^{[c]}2_1$. These noncrystallographic axes must be nearly parallel to the new $a'b'$ plane and make angles of close to 45° with the new a' and b' axes. It is interesting that crystals of the lectin cocrystallized with galactose belong to this space group and have quite similar unit-cell parameters (see Table 1). A low-resolution (*e.g.* at 6 Å) self-rotation function (Fig. 1*b*) reveals peaks corresponding to twofold rotation axes in accordance with our analysis. When compared with Fig. 1(*a*), these peaks are weaker (4.3σ in comparison with 10.1σ for the peak at the origin and 1.7σ for the next peaks). When the self-rotation function is calculated at a higher resolution (*e.g.* at 3 Å) the peaks decrease to the level of noise and may be missed.

The position of noncrystallographic axes in all these crystal forms is in agreement with the observation by Wang & Janin (1993) on the closeness of noncrystallographic symmetry axes to the crystallographic symmetry axes.

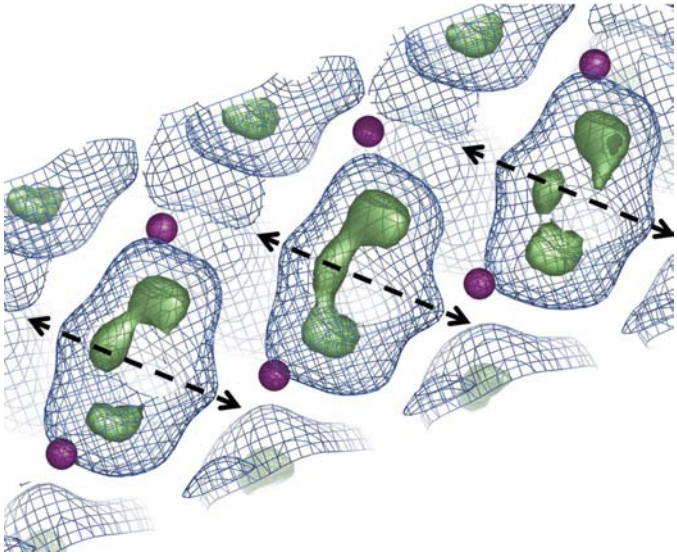

Figure 9
The positions of heavy atoms in a Pt derivative superposed with *ab initio* defined envelopes. Three dimers linked by noncrystallographic pseudo-translation symmetry are shown. The purple spheres indicate the positions of Pt atoms (see §6.4 for details). The arrows indicate the direction of the noncrystallographic dyad, as defined by the self-rotation function. The envelope shown occupies 40% of the unit-cell volume (see §3 for volume estimation). The green surface corresponds to 2.0σ cutoff.

6.4. Heavy-atom analysis

A further independent source of information was the Pt derivative. A direct comparison of the phases obtained by the two approaches was impossible because both phase values were not available simultaneously for any reflection. The same reason prevented us from using difference Fourier synthesis with *ab initio* phases to check the heavy-atom positions. At this level of study, the only way to compare the results of these approaches was to analyze the positions of the Pt atoms.

When comparing two phase sets, the possibility of a different choice of the unit-cell origin and enantiomer must be taken into account (for a general case analysis, see Lunin & Lunina, 1996). The same alignment procedure should be applied when comparing heavy-atom positions with *ab initio* phased images. In a $P2_12_12_1$ crystal there are eight possibilities of permitted origin choice and both enantiomers are compatible with the space group. All 16 possible transformations were applied to the set of heavy atoms. For all but two transformations heavy atoms were obtained either between the envelopes or inside them and the corresponding transformations were eliminated. The transformation $(x, y, z) \rightarrow (-x+\frac{1}{2}, -y, -z+\frac{1}{2})$ positioned the heavy atoms more or less at the surface of the molecular envelope when the envelope volume is equal to 20% of the total unit-cell volume (the envelope volume is defined by an applied density-cutoff level). This volume is unrealistically small and therefore the heavy atoms seem to be too close to the centres of dimers. The transformation $(x, y, z) \rightarrow (-x+\frac{1}{2}, -y+\frac{1}{2}, -z+\frac{1}{2})$ positioned the heavy atoms nicely at the surface of the molecular envelope when the corresponding density-cutoff level is significantly lower and the envelope volume is equal to 50% of the total unit-cell volume (Fig. 9). It is worth noting that the heavy-atom positions are in agreement with the noncrystallographic dyads of dimers as defined from *ab initio* phasing.

This result shows that the *ab initio* phases found do not contradict the preliminary information obtained using the Pt derivative. Unfortunately, at present the low quality of SIR phases together with the lost central zone reflections make a more detailed comparison impossible.

7. Conclusions

This work illustrates the possibilities of *ab initio* phasing in solving problems in protein crystallography that cannot be approached otherwise. Low-resolution images gave an explanation of the existence of several crystal forms and predicted the presence of new crystals, in particular a tetragonal form. In crystals containing multiple copies (six independent monomers in $P2_12_12_1$ crystals), these phases and envelopes are a good starting point for further phase extension through noncrystallographic symmetry averaging. Based on the *ab initio* obtained images, a link between $P2_12_12_1$ and $C222_1$ crystals has been established. This information permits a direct transfer, without any molecular-replacement searches, of the (improved) lectin envelopes into the $C222_1$ crystals, for which higher resolution data are available. These will be the obvious next steps in the crystal structure analysis of the lectin SML-2.

The currently available low-resolution set of experimental magnitudes and *ab initio* determined phases has been deposited as supplementary material[1].

We are grateful to H. Klein, Th. Montag, N. Zyto and B. Löschner from the Paul-Ehrlich-Institut, Bundesanstalt für Seren und Impfstoffe, FG Parasitologie/Diagnostika, 63225 Langen, Germany, who provided the original lectin SML-2. The modified galactose was a gift from U. Pfüller, University of Witten-Herdecke, Germany. This work was supported by the Fonds der Chemischen Industrie. NLL and VYL were supported by RFBR grant 03-04-48155 and joint RFBR–CNRS grant 05-01-22002-CNRS_a. AU was supported by the CNRS-RAS collaboration, GdR 2417 CNRS and by Pole 'Intelligence Logicielle', Nancy.

[1] Supplementary material has been deposited in the IUCr electronic archive (Reference: WD5056). Services for accessing this material are described at the back of the journal.

References

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Cowtan, K. (1994). *Jnt CCP4/ESF–EACBM Newsl. Protein Crystallogr.* **31**, 34–38.
DeLano, W. L. (2002). *The PyMOL Molecular Graphics System.* DeLano Scientific, San Carlos, CA, USA. http://www.pymol.org.
Entzeroth, R., Kerkhoff, H. & Koenig, A. (1992). *Eur. J. Cell Biol.* **59**, 405–413.
Hao, O., Gu, Y.-X., Zheng, C.-D. & Fan, H.-F. (2000). *J. Appl. Cryst.* **33**, 980–981.
Kabsch, W. (1993). *J. Appl. Cryst.* **26**, 795–800.
Klein, H., Mueller, S., Loeschner, B., Toenjes, R. R., Braun, G., Müller, E.-C., Otto, A. & Montag, T. (2003). *Parasitol. Res.* **90**, 84–86.
Kleywegt, G. J., Zou, J. Y., Kjeldgaard, M. & Jones, T. A. (2001). *International Tables for Crystallography*, Vol. *F*, edited by M. G. Rossmann & E. Arnold, pp. 353–356. Dordrecht: Kluwer Academic Publishers.
Lunin, V. Y. & Lunina, N. L. (1996). *Acta Cryst.* A**52**, 365–368.
Lunin, V. Y., Lunina, N. L., Podjarny, A., Bockmayr, A. & Urzhumtsev, A. (2002). *Z. Kristallogr.* **217**, 668–685.
Lunin, V. Y., Lunina, N. L. & Urzhumtsev, A. (2000). *Acta Cryst.* A**56**, 375–382.
Lunina, N. L., Lunin, V. Y. & Urzhumtsev, A. (2003). *Acta Cryst.* D**59**, 1702–1715.
Montag, T., Bornhak, M., Loeschner, B., Klein, H., Otto, A., Zyto, N. & Entzeroth, R. (1997). *Eur. J. Cell Biol.* **74**, *Suppl.* 46, 21.
Müller, J. J., Müller, E.-C., Montag, T., Zyto, N., Löschner, B., Klein, H., Heinemann, U. & Otto, A. (2001). *Acta Cryst.* D**57**, 1042–1045.
Urzhumtseva, L., Lunina, N., Fokine, A., Samama, J.-P., Lunin, V. Y. & Urzhumtsev, A. (2004). *Acta Cryst.* D**60**, 1519–1526.
Vagin, A. & Teplyakov, A. (1997). *J. Appl. Cryst.* **30**, 1022–1025.
Vernoslova, E. A. & Lunin, V. Y. (1993). *J. Appl. Cryst.* **26**, 291–294.
Wang, X. & Janin, J. (1993). *Acta Cryst.* D**49**, 505–512.