Low-resolution data analysis for low-density lipoprotein particle

Vladimir Y. Lunin, Natalia L. Lunina, Stephan Ritter, Ingrid Frey, Aloys Berg, Kay Diederichs, Alberto D. Podjarny, Alexandre Urzhumtsev and Manfred W. Baumstark

Low-resolution data analysis for low-density lipoprotein particle

Vladimir Y. Lunin,[a] Natalia L. Lunina,[a] Stephan Ritter,[b] Ingrid Frey,[b] Aloys Berg,[b] Kay Diederichs,[c] Alberto D. Podjarny,[d] Alexandre Urzhumtsev[e]*† and Manfred W. Baumstark[b]†

[a]Institute of Mathematical Problems of Biology, Russian Academy of Sciences, Pushchino, Moscow Region 142292, Russia, [b]Medezinische Universitätsklinik Freiburg, Hugstetterstrasse 55, D-79106 Freiburg, Germany, [c]Fakultät für Biologie, Universität Konstanz, Fach M656, D-78457 Konstanz, Germany, [d]UPR de Biologie Structurale, IGBMC, BP 163, 67404 Illkirch, CU Strasbourg, France, and [e]LCM3B, UPRESA 7036 CNRS, Faculté des Sciences, Université Henry Poincaré, Nancy I, 54506 Vandoeuvre-lés-Nancy, France

† Correspondence concerning methods should be sent to AU; correspondence concerning low-density lipoprotein should be sent to MWB (maba@ukl.uni-freiburg.de).

Correspondence e-mail: sacha@lcm3b.u-nancy.fr

The knowledge of the molecular structure of LDL, a large lipoprotein complex, is of great interest for medical investigations. Currently available LDL crystals do not diffract to high resolution and do not allow the application of standard crystallographic techniques. Additional difficulties arise because of a very dense crystal packing and the presence of several components with quite different mean densities. Several *ab initio* phasing methods previously reported by the authors have been successfully applied to find a crystallographic image of LDL at a resolution of 27 Å. The most promising results have been obtained using direct phasing with a connectivity analysis of the electron-density maps. The current image makes it possible to discern a single particle covered by a layer of relatively high density that is asymmetrically distributed on the particle surface. It shows a partition of high and low densities inside the particle and, in particular, strips of varying density in the lipid core.

1. Introduction

Low-density lipoprotein (LDL) belongs to the family of lipoproteins which are involved in the transport of lipids from the liver to extrahepatic cells (Brown & Goldstein, 1986) and plays a major role in the development of coronary heart disease (Gordon *et al.*, 1981). Knowledge of the three-dimensional structure of this large macromolecular complex and its components is important in order to explain the molecular mechanisms involved in the development of atherosclerosis. LDL is composed of a single polypeptide chain of apolipoprotein B (apoB; about 22% of the total mass), cholesterol (8%), cholesteryl esters (42%), triglycerides (6%) and phospholipids (22%). The total particle mass is about 2000–3000 kDa, depending on the number of associated lipid molecules. Depending on the different lipid:protein ratio, total LDL can be divided into different subfractions based on the concomitant change in size and density. These subfractions are assumed to have different atherogenic potential (Baumstark *et al.*, 1994; Krauss, 1995).

ApoB is one of the largest known proteins. It has a molecular weight of 550 kDa and is composed of 4536 amino-acid residues. The primary structure of apoB is known (Yang *et al.*, 1986; Knott *et al.*, 1986) and reveals a large number of long (>70 residues) internal repeats, suggesting that the apoB sequence was derived largely from internal duplications. The content of secondary-structure elements for apoB is estimated as 43% α-helix, 21% β-sheet structure, 20% random coil structure and 16% β-turns (Chan, 1992). Immunoelectron microscopic analysis of apoB on LDL suggests that the protein

is extended and spans at least one hemisphere of the particle surface (Chatterton *et al.*, 1991, 1995). Evidence for the accessibility of certain epitopes was derived from studies with monoclonal antibodies (Wiklund *et al.*, 1985; Krul *et al.*, 1988) and by investigating the digestion pattern of different proteases (Chen *et al.*, 1994). Based on the trypsin accessibility of different residues, apoB can be divided into five sub-domains. The amino terminus, structurally homologous to the N-terminal domains of lipovitellin and of microsomal tri-glyceride transfer protein (Segrest *et al.*, 1994, 1999), is followed by four strongly amphipathic domains: two β-sheet and two α-helical. The three-dimensional structure of lipo-vitellin recently reported by Anderson *et al.* (1998) showed the N-terminal domains to be a huge α-helical arc surrounded by β-structural elements.

X-ray and neutron-scattering studies (Laggner & Müller, 1978; Müller *et al.*, 1978; Luzzati *et al.*, 1979; Laggner *et al.*, 1981, 1984; Baumstark *et al.*, 1990, 1991) suggested the model of an LDL particle to be a sphere with a diameter varying from 180 to 240 Å and an approximate radial symmetry of the electron-density distribution. The internal structure of the particle is thought to be formed by spherical shells of different electron density, indicating that the neutral lipids, mainly cholesteryl esters, are arranged in the core of the particle surrounded by an electron-dense surface layer consisting of a monolayer of phospholipids, free cholesterol and the protein. It was concluded that apoB covers approximately 50% of the particle surface (Baumstark *et al.*, 1990).

Electron-microscopy studies have been performed by several groups. By freeze-etching electron microscopy Gulik-Krzywicki *et al.* (1979) reported the LDL structure as being a pseudo-spherical particle with several symmetrically arranged bumps at the surface. More details were revealed by electron cryomicroscopy. van Antwerpen *et al.* (1997) showed a series of comparative images of LDL in different states and suggested a discoidal model. In a recent paper using electron cryomicroscopy followed by three-dimensional reconstruction, Orlova *et al.* (1999) gave a quite detailed model nomin-ally at 27 Å. Beyond the biochemical problems, the major difficulties in these studies were caused by the highly symmetrical shape of the particle and resulted in eventual ambiguities in the identification and interpretation of experi-mental electron-microscopy images.

Conditions for three-dimensional crystallization of LDL have been reported previously (Prassl *et al.*, 1996; Ritter *et al.*, 1997). However, no crystallographic analysis has so far been reported in order to obtain the three-dimensional image of LDL. Two major prerequisites had to be met in order to undertake such an analysis. Firstly, complete native low-resolution data had to be available (Ritter *et al.*, 1999); secondly, a suitable method for the solution of the phase problem had to be found. A promising approach was to solve the phase problem by new direct-phasing methods at low resolution (Lunin, Lunina *et al.*, 2000; Lunin, Podjarny *et al.*, 2000). Previous low-resolution studies explicitly or implicitly used a two-component low-resolution model to approximate the compartments occupied by protein and bulk solvent. In

the LDL crystals, however, components with an electron density higher than the bulk solvent (protein and phospholipid head groups) and lower than the bulk solvent (fatty acids) are present. Therefore, for the first time, a three-component low-resolution model had to be introduced to adequately describe the electron-density distribution. The results of applying low-resolution direct-phasing methods, both previously developed ones and those specially updated for the LDL case, are presented in this paper.

2. Crystallization and data collection

Human LDL subfraction LDL-2 (density 1.031–1.034 g ml^{-1}) was crystallized as reported previously (Ritter *et al.*, 1997). Crystals were grown at 290 K, which is well below the temperature of the phase transition of the cholesterol ester core of the LDL particle between an ordered and a disordered state (Deckelbaum *et al.*, 1977). Native data sets from LDL-2 crystals were collected within a resolution range of 300–27 Å under cryogenic conditions using a special beamstop and a helium tunnel to allow collection of the lowest order reflec-tions (Ritter *et al.*, 1999) at beamline ID14 EH3 (ESRF, Grenoble). Indexing and integration of the data sets was performed using *XDS* (Kabsch, 1993). The LDL crystals are not well ordered. A number of reflections can be measured to 15 Å resolution; however, owing to a strong diffraction anisotropy the efficient resolution of the measured diffraction data set is only approximately 27 Å. Data reduction was optimized in order to obtain as high a data coverage as possible, resulting in a data set that has all reflections at the low-resolution end (Tables 1 and 2). The space group is most probably $C2$, with unit-cell parameters $a = 183$, $b = 421$, $c = 385$ Å, $\alpha = \gamma = 90°$, $\beta \simeq 90°$ (one LDL particle per asymmetric unit); however, owing to the absence of higher resolution reflections, the choice of the space group is not unambiguous and the data can possibly be reduced in space groups $C222$ and $C222_1$ (half an LDL particle per asymmetric unit).

3. The analysis of possible particle packing for the orthorhombic space groups

As mentioned above, the intensity statistics are compatible with both the monoclinic group $C2$ and the orthorhombic space groups $C222$ and $C222_1$. This section discusses the possible particle packing for orthorhombic groups. Small-angle X-ray and neutron scattering and electron-microscopy studies suggest that at low resolution the LDL particle may be approximated by a sphere with a diameter of about 200 Å. A comparison of the particle and the unit-cell volumes shows with high probability that there are four particles in the unit cell. In space groups $C222$ or $C222_1$ the only possibility for this is to assume that the particle has an internal twofold symmetry and that its centre occupies a special position on a dyad. This does not agree with the fact that the LDL contains only one molecule of apolipoprotein B, which does not possess exact rotational symmetry. Nevertheless, such an approximation

Table 1
Merging statistics in space group $C222_1$ for a native data set of a LDL-2 measured at ESRF beamline ID14-3.

Values were calculated using the program *XSCALE* of the *XDS* package (Kabsch, 1993). R_{meas} is the redundancy-independent R factor (Diederichs & Karplus, 1997).

Resolu-tion (Å)	Observed	Unique	Unique $(I > 3\sigma)$	Complete-ness (%)	I/σ	R_{meas} (%)	R_{meas} $(I > 3\sigma)$ (%)
100	83	13	13	100.0	48.1	8.9	8.9
60	469	38	38	100.0	52.9	6.4	6.4
50	523	34	34	100.0	42.1	8.0	8.0
40	1124	73	70	100.0	33.3	9.3	9.1
30	3163	193	160	99.5	14.7	25.3	22.6
27	853	126	64	100.0	3.9	49.6	36.5
15	13890	1973	116	94.3	1.4	114.5	47.2
Total	20105	2447	346	95.2	5.1	9.3	8.5

Table 2
Merging statistics in space group $C2$ for a native data set of a LDL-2 measured at ESRF beamline ID14-3.

Resolu-tion (Å)	Observed	Unique	Unique $(I > 3\sigma)$	Complete-ness (%)	I/σ	R_{meas} (%)	R_{meas} $(I > 3\sigma)$ (%)
100	89	22	22	95.6	44.7	4.4	4.4
60	459	65	64	100.0	44.9	5.0	5.0
50	525	63	61	100.0	31.8	8.2	8.2
40	1131	132	121	100.0	25.6	9.5	9.2
30	3165	358	269	99.7	10.9	26.3	21.5
27	853	238	62	100.0	2.4	49.7	26.8
15	13895	3784	63	94.4	0.8	114.5	27.5
Total	20094	4655	522	95.2	3.6	8.2	6.9

Table 3
Observed structure-factors magnitude for $00l$ reflections.

Even l						Odd l					
h	k	l	F^{obs}	σ^{obs}	d	h	k	l	F^{obs}	σ^{obs}	d
0	0	2	950	68	190	0	0	3	13	3.8	126
0	0	4	356	5.6	95	0	0	5	3.5	1.2	76
0	0	6	176	2.6	63	0	0	7	6.8	2.1	54
0	0	8	73	1.2	47	0	0	9	3	1.1	42
0	0	10	27	0.81	38	0	0	11	2	1.2	34
0	0	12	8	0.99	3	0	0	13	2	1.2	29
0	0	14	3	1.6	27						

may be reasonable at low resolution, so that the possibility of a $C222$ or $C222_1$ space group was not rejected by this argument.

3.1. *C222* group

The analysis of structure-factor magnitudes for $00l$ reflections (Table 3) suggests the presence of a screw axis. However, the small number of the measured reflections does not allow a definite conclusion to be drawn and the analysis of both space groups $C222$ and $C222_1$ is necessary.

In the $C222$ space group, a simple geometrical analysis of the rotation axes reveals the only possible way to place four spheres of 200 Å diameter in the LDL unit cell without significant overlapping of symmetry-linked particles. The sphere centres should be placed at the point $\mathbf{r}_1 = (0, 0, 1/4)$ and the symmetry-related points $\{\mathbf{r}_j\}$. Naturally, the points linked to them by an origin/enantiomorph transformation permitted for this space group are also possible.

If the density distribution in the particle has an approximate internal symmetry (*e.g.* spherical symmetry; see *Appendix A* for a more general case), it is possible to express structure factors as

$$\mathbf{F}(\mathbf{h}) \simeq \mathbf{F}_{ind}(\mathbf{h})\mathbf{T}(\mathbf{h}). \tag{1}$$

Here, $\mathbf{F}_{ind}(\mathbf{h})$ are the structure factors corresponding to a single particle placed at the origin and $\mathbf{T}(\mathbf{h})$ are trigonometric structure factors defined by the particle packing

$$\mathbf{T}(\mathbf{h}) = \sum_{j=1}^{n} \exp[2\pi i(\mathbf{h}, \mathbf{r}_j)]. \tag{2}$$

The summation in (2) is over all the particles present in the unit cell. For the case considered above with $\mathbf{r}_1 = (0, 0, 1/4)$, $\mathbf{r}_2 = (0, 0, -1/4)$, $\mathbf{r}_3 = (1/2, 1/2, 1/4)$ and $\mathbf{r}_4 = (1/2, 1/2, -1/4)$, the trigonometric structure factors become

$$\mathbf{T}(\mathbf{h}) = 2\{1 + \exp[\pi i(h + k)]\}[1 + \exp(-\pi i l)]\exp(\pi i l/2). \tag{3}$$

It follows that $\mathbf{T}(\mathbf{h}) = 0$ for all reflections with an odd l index and the corresponding structure-factor magnitudes must be close to zero,

$$F(\mathbf{h}) \simeq 0 \text{ if } l \text{ is odd.} \tag{4}$$

[We say that the magnitude is close to zero and not equal to it because (1) is only an approximation.] The diffraction data do not agree with the condition (4) (Table 4) and the hypothesis about the $C222$ space group must be rejected.

3.2. *C222₁* group

In the $C222_1$ space group two possibilities exist to place a 200 Å diameter sphere into the LDL unit cell without a significant overlapping of symmetry-linked particles. The sphere centre can be placed either on the $[x, 0, 0]$ axis or on the $[0, y, 1/4]$ axis. In the former case, the trigonometric structure factors (2) for $0kl$ reflections are

$$\mathbf{T}(0, k, l) = 2[1 + \exp(\pi i l)][1 + \exp(\pi i k)]) \tag{5}$$

and

$$F(0, k, l) \simeq 0 \text{ if } l \text{ is odd.} \tag{6}$$

For the latter case, for $h0l$ reflections we obtain

$$\mathbf{T}(h, 0, l) = 2\exp(\pi i l/2)[1 + \exp(\pi i h)][1 + \exp(\pi i l)] \tag{7}$$

and

$$F(h, 0, l) \simeq 0 \text{ if } l \text{ is odd.} \tag{8}$$

The diffraction data (Tables 5 and 6) do not fit the condition (5), while the condition (8) is satisfied with reasonable accuracy.

The space group $C222_1$ was chosen as the result of the analysis performed and the position of the particle centre was restricted to the $[0, y, 1/4]$ axis. The determination of the co-ordinate y is discussed below.

Table 4
Observed structure-factor magnitudes.

Even l						Odd l					
h	k	l	F^{obs}	σ^{obs}	d	h	k	l	F^{obs}	σ^{obs}	d
0	2	0	437	11	208	0	2	1	784	30.1	182
0	0	2	950	68	190	1	1	1	576	8.4	151
1	1	0	332	4.6	165	0	2	3	228	2.8	108
0	2	2	168	1.9	140	1	3	1	16	0.30	106
1	1	2	117	1.1	125	1	1	3	195	1.5	100
1	3	0	70	1.0	110	0	4	1	290	3.9	100

Table 5
Observed structure-factor magnitudes for low-resolution $0kl$ reflections.

Even l						Odd l					
h	k	l	F^{obs}	σ^{obs}	d	h	k	l	F^{obs}	σ^{obs}	d
0	2	2	168	1.9	140	0	2	3	228	2.8	108
0	4	0	159	3.3	104	0	4	1	290	3.9	100
0	4	2	139	1.5	91	0	4	3	139	1.7	80
0	2	4	136	1.6	86	0	2	5	54	0.60	71
0	4	4	51	0.50	70	0	4	5	117	1.3	61
0	2	6	68	7.9	60	0	2	7	16	0.53	52

It may be noted that the given analyses of structure-factor magnitudes was based on the approximate equation (1) valid for pseudo-spherically symmetrical particles. Similar conclusions may be obtained for isolated particles of a more general shape (*Appendix A*).

3.3. The search for the position of the centre of the particle: one-shell model

If the position of the centre of the particle is restricted to the line $[0, y, 1/4]$, then the trigonometric structure factors (2) are

$$\mathbf{T}(\mathbf{h}) = T(h, k, l) = \begin{cases} 4\cos(2\pi ky) & \text{for } l = 4n \\ -4\sin(2\pi ky) & \text{for } l = 4n+1 \\ -4\cos(2\pi ky) & \text{for } l = 4n+2 \\ 4\sin(2\pi ky) & \text{for } l = 4n+3, \end{cases} \quad (9)$$

where n is an integer. Now suppose that the distribution of the points of high electron density in the particle has a radial symmetry and that these points are concentrated mostly at the distance R from the particle centre (Fig. 1).

To be more precise, let the density distribution be approximately

$$\rho_{ind}(r) \simeq C\left(\frac{4\pi}{B}\right)^{3/2}\left\{\exp\left[-\frac{4\pi^2}{B}(r-R)^2\right] + \exp\left[-\frac{4\pi^2}{B}(r+R)^2\right]\right\}. \quad (10)$$

(The second term in this sum is small enough for the positive r values, but allows calculations to be performed in the close form.) In this case, the scattering by the electron density (10) is

Table 6
Observed structure-factor magnitudes for low-resolution $h0l$ reflections.

Even l						Odd l					
h	k	l	F^{obs}	σ^{obs}	d	h	k	l	F^{obs}	σ^{obs}	d
2	0	0	429	7.4	90	2	0	1	19	0.29	88
2	0	2	335	5.3	81	2	0	3	5.4	1.2	73
2	0	4	97	1.2	65	2	0	5	5.1	0.43	58
2	0	6	23	0.47	52	2	0	7	8.2	0.50	46
4	0	0	196	6.9	45	4	0	1	19	0.97	45
4	0	2	190	8.8	44	4	0	3	21	0.24	42
4	0	4	55	1.1	41	4	0	5	2.7	0.46	39

$$\mathbf{F}_{ind}(\mathbf{h}) = \frac{2}{|\mathbf{h}|}\int_0^\infty r\rho_{ind}(r)\sin(2\pi|\mathbf{h}|r)\,\mathrm{d}r \quad (11)$$

$$= 2C\left[\cos(2\pi|\mathbf{h}|R) + \frac{4\pi R}{B|\mathbf{h}|}\sin(2\pi|\mathbf{h}|R)\right]\exp\left(-\frac{B|\mathbf{h}|^2}{4}\right)$$

so that

$$F^{calc}(h, k, l) \simeq |\mathbf{F}_{ind}(\mathbf{h})\mathbf{T}(\mathbf{h})|$$

$$\simeq |\mathbf{F}_{ind}(\mathbf{h})| \begin{cases} |\cos(2\pi ky)| & \text{for even } l \\ |\sin(2\pi ky)| & \text{for odd } l. \end{cases} \quad (12)$$

The approximate value of the particle radius $R \simeq 100$ Å is known experimentally, while the B value may be estimated from the low-resolution Wilson plot as $B = 15\,000$ Å^2. It is reasonable to expect that the coordinate y of the particle centre maximizes the correlation coefficient

$$C = \frac{\sum_{\mathbf{h}}[F^{calc}(\mathbf{h}) - \langle F^{calc}\rangle][F^{obs}(\mathbf{h}) - \langle F^{obs}\rangle]}{\left\{\sum_{\mathbf{h}}[F^{calc}(\mathbf{h}) - \langle F^{calc}\rangle]^2\right\}^{1/2}\left\{\sum_{\mathbf{h}}[F^{obs}(\mathbf{h}) - \langle F^{obs}\rangle]^2\right\}^{1/2}} \quad (13)$$

for the observed structure-factor magnitudes and those calculated by (11). Fig. 2 shows the dependence of the correlation coefficient value on the y coordinate of the particle centre. The four peaks correspond to the particle-centre positions linked by the space-group symmetries or by permitted origin shifts. As the values of R and B chosen were approximate values, the calculations were repeated varying the value for R within the range 100–115 Å and the value for B within the range 6000–20 000 Å^2, but this did not change the
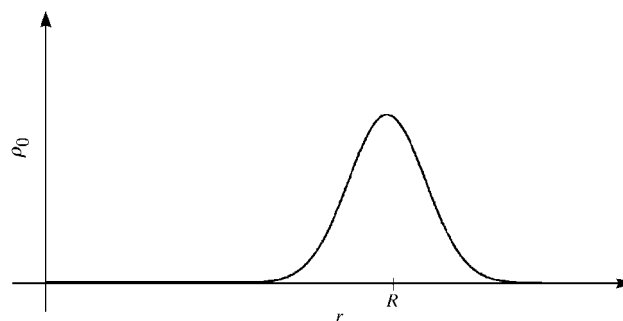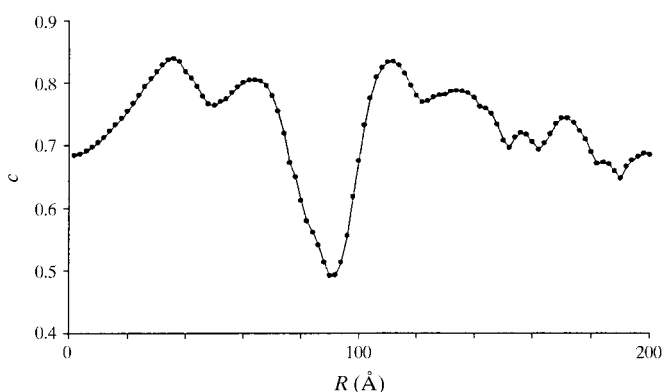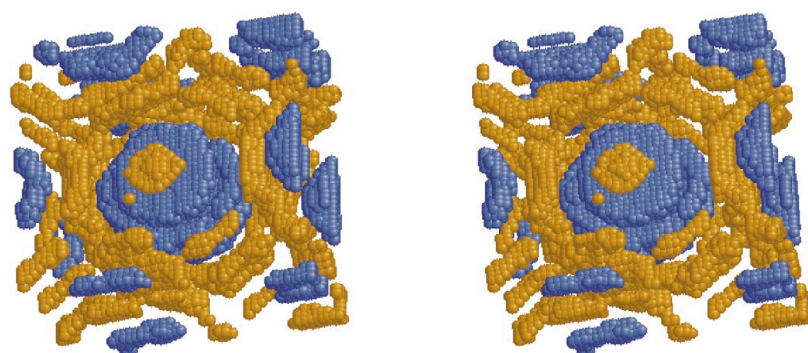

Figure 1
One-shell density distribution.

optimal value $y = 0.16$. Maximal correlation was obtained for $R = 110$ Å and $B = 15\,000$ Å^2. The approximate equation $B \simeq 8\pi^2 a^2$ allows the width of the high-density layer to be estimated to be $2a = 28$ Å, which corresponds to the expected value.
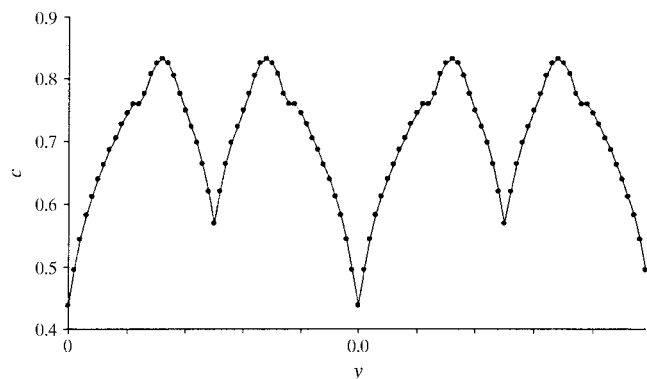
Figure 2
The correlation (13) of the calculated (for $R = 105$ Å and $B = 15\,000$ Å^2) and observed magnitudes *versus* the y coordinate of the particle centre. 456 reflections in the 27 Å resolution zone were used. The position of four maxima are linked by space-group symmetries and permitted origin shifts.

Figure 3
The correlation (13) of the calculated (with $y = 0.16$ and $B = 15\,000$ Å^2) and observed magnitudes *versus* radius R of the shell. 456 reflections in the 27 Å resolution zone were used. The two main maxima at $R = 30$ Å and $R = 110$ Å correspond to the shells of the lowest and highest density in the particle.

Figure 4
Three-shell phased Fourier synthesis for LDL particle (stereoview). High density is shown in yellow and low density in blue.

3.4. Three-shell model

The plot of the dependence of the magnitude correlation coefficient on radius of the shell R (Fig. 3) shows the presence of additional maxima corresponding to $R = 30$ Å and $R = 60$ Å. These maxima are separated from the $R = 110$ Å maximum by a deep cleft. The presence of the maxima on the left agrees with the hypothesis that the lowest density in the particle is located in a shell; however, this shell is not an ideal one so the peak splits into two maxima. The deep cleft might be interpreted as the existence of a spherical shell inside the particle with either an almost equal ratio of high and low density or filled uniformly with the density close to the mean density value in the crystal. The further studies confirmed the latter hypothesis (see Figs. 14 and 15 below).

The parameters of the shells of the lowest and highest density may be refined if the calculated magnitudes are based on a three-shell model composed of three Gaussian-type shells (10). In this case,

$$F^{\mathrm{calc}}(h, k, l) \simeq 2C\sum_{j=1}^{3}\alpha_j R_j \left| \cos(2\pi|\mathbf{h}|R_j) + \frac{4\pi R_j}{B|\mathbf{h}|}\sin(2\pi|\mathbf{h}|R_j) \right|$$
$$\times \exp\left(-\frac{B|\mathbf{h}|^2}{4}\right) \begin{cases} |\cos(2\pi ky)| & \text{for even } l, \\ |\sin(2\pi ky)| & \text{for odd } l, \end{cases}$$
$$(14)$$

where the weights α_j may be positive or negative and their magnitudes reflect the deviations of the density in the low-density shell from the mean density for the crystal. If the weight α_1 for $R_1 = 110$ Å is fixed as 1, the maximization of the correlation coefficient (13) resulted in $\alpha_2 = -1.2$ ($R_2 = 60$ Å) and $\alpha_3 = -0.4$ ($R_3 = 30$ Å). The negative values of α_2 and α_3 show that the inner-shell mean densities deviate in the opposite direction to the mean density in the crystal compared with the mean density for the outer shell.

The three-shell model allowed the magnitude correlation coefficient to increase from 0.83 to 0.87 in the 27 Å resolution zone in comparison with a single-shell model. At the same time, this did not practically influence the Fourier syntheses calculated with the observed magnitudes and model phases.

The map correlation coefficient (Lunin & Woolfson, 1993) for one-shell and three-shell phased syntheses is 0.97 in the entire 27 Å resolution zone and 0.76 in the resolution shell 30.0–27.0 Å. Fig. 4 represents the three-shell phased synthesis.

It is worthy of note that while the phases were calculated from the density distribution possessing radial symmetry, the Fourier synthesis does not reveal proper radial symmetry. Rather, it shows the particle as composed of layers separated along the x axis. This effect of modulation along the x axis follows from the anisotropy of the diffraction field as discussed in §4 and *Appendix B*.

4. Intensity analysis: low-resolution projection of the crystal structure

One consequence of the spherical symmetry of the LDL particle would be the spherical symmetry of the corresponding structure factors $F_{\mathrm{ind}}(\mathbf{h})$. These structure factors might be real (as opposed to complex) numbers, with the magnitudes depending on resolution $d = s^{-1}$ only. As a consequence, the plot of structure-factor magnitudes $F_{\mathrm{ind}}(\mathbf{h})$ *versus* s^2 value might be represented by a set of points distributed along a smooth line. Fig. 5 shows a plot of $\ln F$ *versus* s^2 for the experimental data in the 30 Å resolution zone and does not reveal a proper spherical symmetry for the observed magnitudes: magnitudes corresponding to close s^2 values differ significantly. Nevertheless, some kind of waves may even be seen in this picture when looking carefully. These waves may be obtained more clearly for modified magnitude values corresponding to an isolated LDL particle rather than the whole unit-cell content.

It follows from (1) that for particles possessing additional internal symmetry (for example, spherical symmetry) the structure factor corresponding to the individual particle placed into the origin may be estimated as
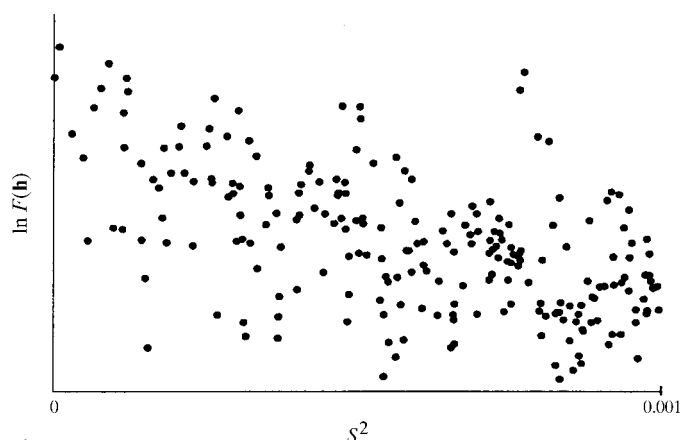
Figure 5
Observed structure-factor magnitudes *versus* s^2. Every point represents an individual structure factor.
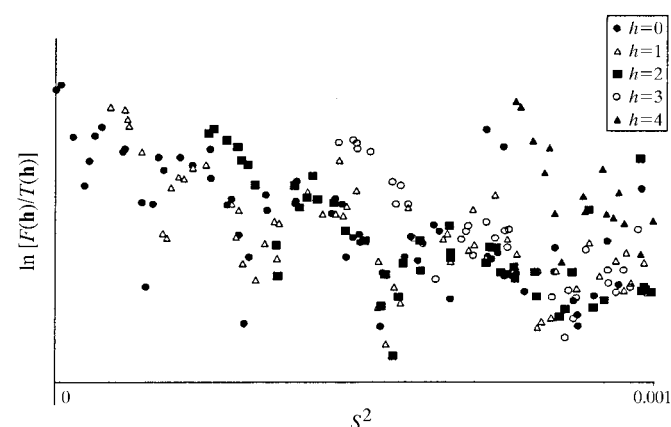
Figure 6
Observed structure-factor magnitudes modified in accordance with (15) plotted *versus* s^2 value. Every point represents a structure factor.

$$\mathbf{F}_{\mathrm{ind}}(\mathbf{h}) \simeq \mathbf{F}(\mathbf{h})/\mathbf{T}(\mathbf{h}). \tag{15}$$

Fig. 6 presents the diagram corresponding to the structure-factor magnitudes modified in accordance with (15), assuming the coordinates of the centre of an independent particle are equal to (0.0, 0.16, 0.25). This diagram reveals more spherical symmetry for the modified structure-factor magnitudes $F_{\mathrm{ind}}(\mathbf{h})$ than Fig. 5 does for the native $F(\mathbf{h})$. Nevertheless, Fig. 6 suggests that the diagram consists of several curves shifted along the s^2 axis rather than of a single smooth curve, as would be the case in the presence of the strong spherical symmetry in $F_{\mathrm{ind}}(\mathbf{h})$ values. A possible explanation is the modulation of the density distribution along the x axis (see *Appendix B* for details).

Fig. 7 shows the diffraction pattern in the plane $l = 0$ of reciprocal space. The essential feature of this image is a strong anisotropy in the h direction. Such anisotropy might be a result of a modulation of the density distribution along the x axis

$$\rho(x, y, z) = \mu(x)\rho_0(x, y, z), \tag{16}$$

where ρ_0 is a smooth spherically symmetric distribution and μ is a function oscillating along the x axis (see *Appendix B*). As another consequence of the modulation, each one of the subsets of points in Fig. 6 corresponding to the planes defined by a fixed h index is close to a curve, but these curves are shifted by different amounts for different values of h. These diagrams calculated for several reciprocal-space planes (Fig. 8) agree with the modulation hypothesis. One more confirmation of this hypothesis is the presence of parallel strips of high density which are seen in electron-microscopic pictures (Orlova *et al.*, 1999).

Equation (15) and Fig. 8 suggest how to obtain the projection of the density distribution along the x axis. Such a projection
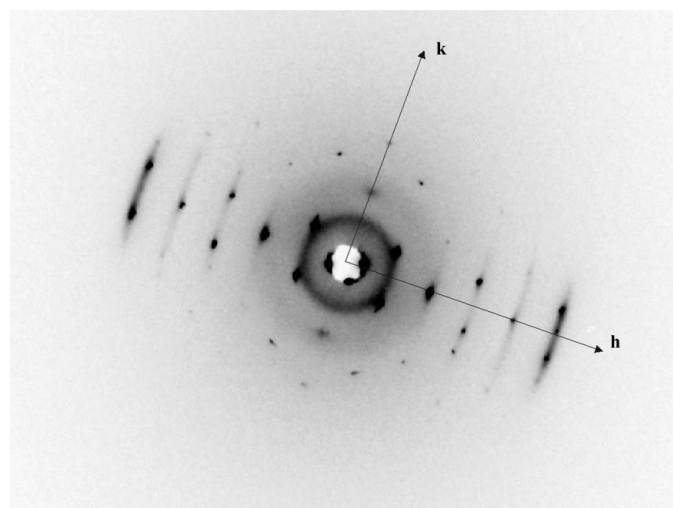
Figure 7
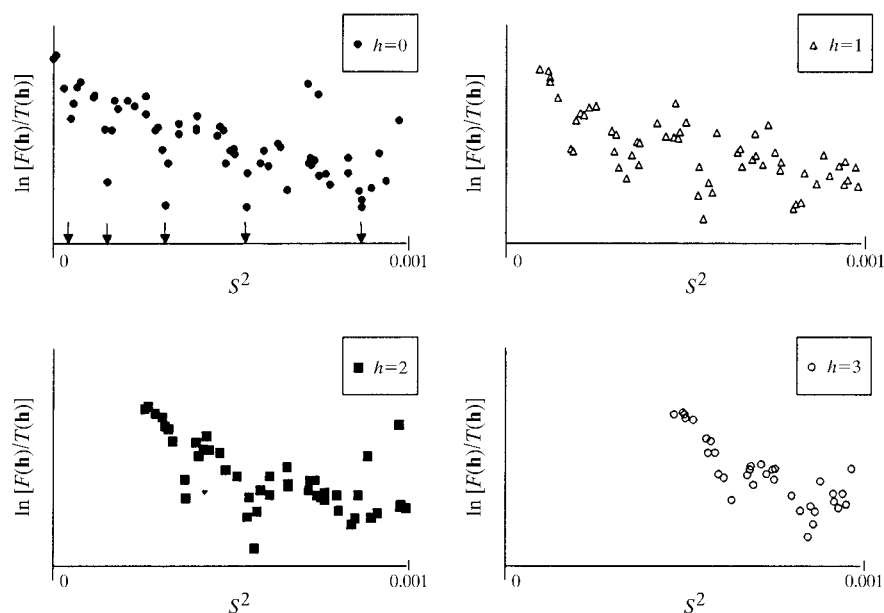Diffraction pattern in the plane $l = 0$ for the LDL crystal.

$$\rho_x(y, z) = \int_0^1 \rho(x, y, z) \, \mathrm{d}x \qquad (17)$$

may be calculated as the Fourier series restricted to $\mathbf{F}(0, k, l)$ structure factors. If the approximation (1) is valid and the position of a particle is known, then the problem of phasing $\mathbf{F}(0, k, l)$ structure factors is reduced to the phasing of values corresponding to one particle. On the other hand, Fig. 8 supports the idea that the corresponding structure factors $\mathbf{F}_{\mathrm{ind}}(0, k, l)$ are derived from a spherically symmetric particle and thus have only a real part. Therefore, the problem of their phasing is the problem of the choice of their sign. The points where structure factors are expected to change their sign may be determined from the diagram as those for which $\mathbf{F}_{\mathrm{ind}}(0, k, l) \simeq 0$ (marked by arrows in Fig. 8). Fig. 9 shows the projection $\rho_x(y, z)$ of the Fourier synthesis calculated with the observed magnitudes and phases defined from (1).

5. FAM phasing

The few atoms model (FAM) approach was developed recently (Lunin *et al.*, 1995, 1998) for very low resolution *ab initio* phasing. It consists of several steps:

(i) generation of a large number of models composed of a few Gaussian spheres; random choice of sphere coordinates in the whole unit cell in the first step and in a current molecule region in the next steps;

(ii) calculation of sets of structure-factor magnitudes and phases for every generated model;

(iii) selection of the phase sets corresponding to FAMs whose structure-factor magnitudes reveal good correlation with the observed magnitudes;

(iv) separation of the selected phase sets into a few clusters;

Table 7
Protocol of FAM-based phasing.

Step	Resolution (Å)	No. of independent FAM atoms	Fixed part of FAM	No. of FAMs generated	No. of selected phase sets
1	80	2	No	478	100
2	60	4	High	955	100
3	60	4	Low	1297	100
4	40	8	High	670	100
5	40	8	Low	1351	100
6	30	16	High	1290	100
7	30	16	Low	2028	100

(v) averaging phases inside every cluster.

The sets of cluster-average phases produce alternative solutions for the phase problem. A simplified approach which could be applied at the first stages of analysis is just to average all selected variants. In the latter case, we are not faced with the problem of chosing the best cluster, but the quality of the resulting image would be diminished.

Application of this approach to such objects as the LDL particle has a special feature. Owing to the presence of compact regions both for high and low density, high correlation values may be obtained for the blobs placed in any of these regions. Thus, the set of permitted transformations allowed in the map alignment process (Lunin & Lunina, 1996) was extended by an additional degree of freedom, namely density flipping $\rho(\mathbf{r}) \rightarrow -\rho(\mathbf{r})$. Fig. 10 shows the result of averaging 100 maps at 80 Å resolution (24 independent reflections) selected from 478 maps phased with randomly generated two-sphere models. (Because of the eight symmetry transformations in $C222_1$ group and the four LDL particles present in the unit cell, this means four pseudo-atoms per LDL particle.) This synthesis reveals four compact regions of the lowest density and highly merged regions of highest density. The centres of the lowest density regions coincide with those obtained from packing considerations (see §3). The further steps were performed with an increasing number of pseudo-atoms in FAM and an extension of the resolution zone. At every step the phases obtained previously were used (together with observed magnitudes) to define the regions of highest and lowest densities. Alternating from step to step, one of these regions was filled by pseudo-atoms placed at the grid point, while another was used as a mask for random atom generation. Table 7 presents the phasing protocol.
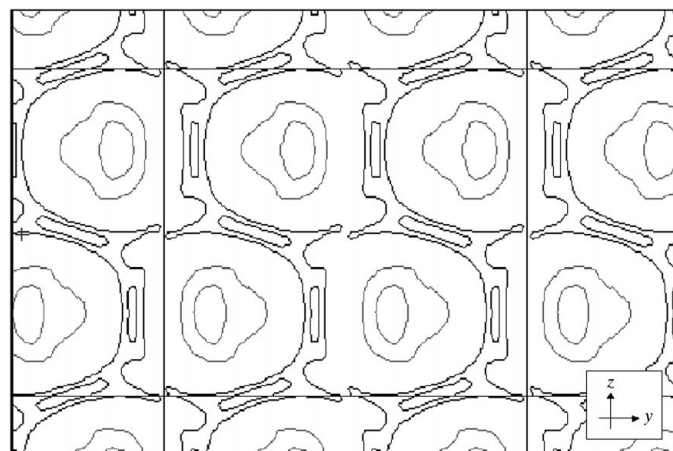
The FAM phasing procedure has resulted in phases reasonably close to those obtained with the use of the three-shell model (§4). The map correlation coefficients for these syntheses calculated for different resolution zones are

Figure 8
Observed structure-factor magnitudes modified in accordance with (15) plotted *versus* s^2 for different planes h = constant in reciprocal space. Every point represents an individual structure factor. Arrows at the left top figure ($h = 0$) indicate the points where $\mathbf{F}_{\mathrm{ind}}(0, k, l)$ changes its sign.

summarized in Table 8 and the corresponding syntheses are shown in Fig. 11. The resulting FAM-phased synthesis is shown in Fig. 12 and reveals a density pattern close to that obtained by the methods given in §3 (see Fig. 9 for comparison). It must be pointed out that the information about the number of particles in the unit cell and about their positions was not used in the FAM phasing procedure. Therefore, the results of the FAM phasing confirm the hypothesis about the special position of this centre and the values of its coordinates.

6. Connectivity-based *ab initio* phasing for the LDL particles

The property of connectivity of the regions of high electron density in protein crystals was used for many years as a qualitative criterion of the goodness of calculated electron-density maps and was formalized by Baker *et al.* (1993) as a quantitative criterion for high-resolution maps. Recently, this

Table 8
Map correlation coefficients for FAM phased and three-shell phased maps in different resolution zones.

Resolution	∞–100	∞–80	∞–60	∞–40	∞–30
C_φ	0.81	0.71	0.70	0.65	0.64
No. of reflections	12	24	49	153	338

idea was adapted to low-resolution *ab initio* phasing (Lunin *et al.*, 1999; Urzhumtsev *et al.*, 2000) and the connectivity criterion was included into the general 'enriching' phasing procedure as one of the selection criteria (Lunin, Lunina *et al.*, 2000).

The basic idea of the approach consists of the observation that topological properties of regions of highest density values in the Fourier synthesis map are different for properly phased syntheses and for those calculated with random phases. Let $\rho(\mathbf{r})$ be a Fourier synthesis calculated with the observed magnitudes and some phases. With this synthesis we associate a high-density region, which is defined formally as a set of

Figure 9
Density projection along x axes calculated at 27 Å resolution. Black contours correspond to high-density values and grey contours to the lowest electron-density values.
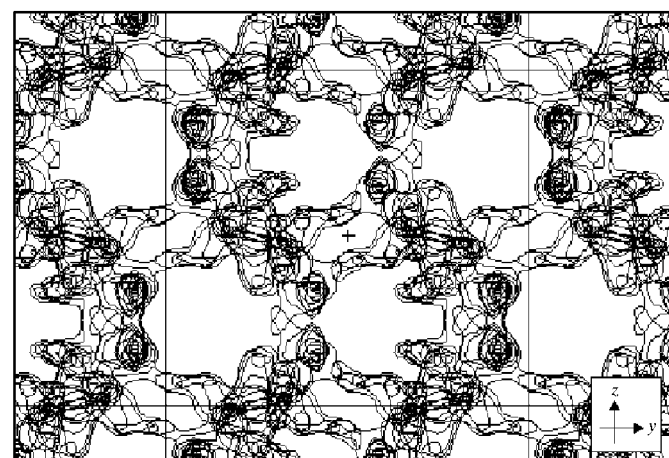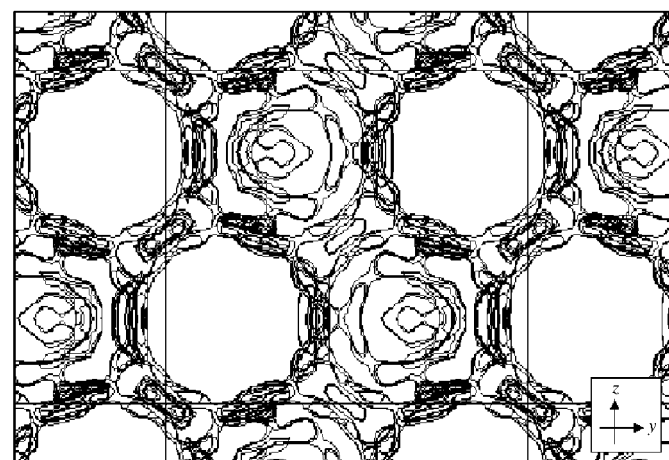

(a)

(b)
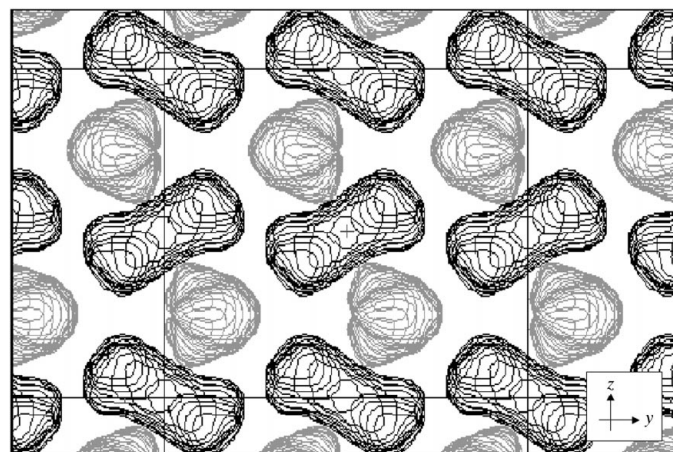
Figure 10
Fourier synthesis obtained after the first step of the FAM procedure. Black contours correspond to high-density region and grey contours to low-density region.

Figure 11
Density projection along the x axis for a part of the unit cell for the FAM phased (a) and the three-shell phased (b) unweighted Fourier syntheses at 27 Å resolution.

Table 9
Connectivity-based *ab initio* phasing for the LDL particles (space group *C*2).

A low-density region was defined as one containing those 10% points of the unit cell with lowest Fourier synthesis values. A high-density region was defined as one containing those 10% points of the unit cell with highest Fourier synthesis values. At all steps except the 12th, the blobs were required to have finite dimensions.

Step	Selection conditions		Alignment resolution (Å)
	Lowest density region [four finite blobs for every specified resolution (Å)]	High-density region (restrictions on 27 Å resolution Fourier synthesis)	
1	150		150
2	150 100 80		100
3	150 100 80 60 40		80
4	150 100 80 60 40		60
5	150 100 80 60 40 30 27		40
6	150 100 80 60 40 30 27		35
7	150 100 80 60 40 35 30 27		30
8	150 100 80 60 40 35 30 27		29
9	100 60 40 35 30 29 27		27
10	100 60 40 35 30 29 27	<20 blobs in whole	27
11	100 60 40 35 30 29 27	<16 blobs	27
12	100 20	4 equal largest domains; possibly infinite	27
13	100 20	4 equal largest domains	27
14	100 20	4 equal largest domains; <20 blobs	27
15	100 20	4 equal largest domains; <16 blobs	27
16	100 20	4 equal largest domains; <12 blobs	27
17	100 20	4 equal largest domains; <8 blobs	27

points in the unit cell possessing the highest values of the function $\rho(\mathbf{r})$,

$$\Omega_\kappa = \{\mathbf{r} : \rho(\mathbf{r}) \geq \kappa\}. \qquad (18)$$

When a low-resolution synthesis is calculated with the exact phases, this region is expected to be composed of a small number of compact 'blobs' corresponding to independent molecules. The number of these blobs is equal to the number of molecules in the unit cell. The blob volumes are equal if all molecules are linked by crystallographic symmetries or are approximately equal if non-crystallographic symmetry is

Figure 12
FAM phased synthesis in the projection along the *x* axis. Black contours correspond to the high-density region and grey contours to the low-density region.

present. On the other hand, the high-density regions built with the use of randomly phased syntheses are likely to show infinite merged regions or a large number of 'drops'.

The phasing procedure consists of several steps. Firstly, a large number of random phase sets are generated. These phases are generated as uniformly distributed phases (with standard restrictions for centric phases) at the beginning of the procedure or in accordance with the known phase distributions if such information is already available. Every phase set generated is used together with the observed magnitudes to calculate the Fourier synthesis and to build the corresponding high-density region. This region is subjected to connectivity analysis in order to determine the number of separated connected components in the unit cell and to calculate their volumes. If the region consists of the desired number of components, then the phase set used is considered to be admissible and is stored for further analysis. After a reasonable number of admissible phase sets have been selected (about 100 variants in our study), they are analysed by the cluster-analysis methods to separate them into clusters of close phase sets. In every isolated cluster, the phase sets are averaged for every reflection in order to produce the corresponding 'best' phases $\varphi^{best}(\mathbf{h})$ and figures of merit $m(\mathbf{h})$ which reflect the spread of the corresponding phase in the cluster sets,

$$m(\mathbf{h})\exp[i\varphi^{best}(\mathbf{h})] = \frac{1}{M}\sum_{j=1}^{M}\exp[i\varphi^j(\mathbf{h})]. \qquad (19)$$

Here, M is the number of phase sets in the cluster and $\varphi^j(\mathbf{h})$ is the phase value of the structure factor with the index $\mathbf{h}$ in the *j*th phase set selected. If several clusters were isolated, the question may arise which cluster must be chosen for a further analysis. We do not discuss how to make this choice, as this problem did not occur in the LDL analysis.

Obviously the properties of Ω_κ regions depend on the chosen cutoff level κ. Test calculations (Lunin, Lunina & Urzhumtsev, 2000) have shown that a suitable choice is to use the κ value which results in a volume of Ω_κ of about 25 Å³ per residue. The computer algorithm of the connectivity analysis will be discussed elsewhere.

For LDL, there are some features which make it different from a typical protein. First of all, the protein shell of the LDL particle occupies a relatively small part of the unit cell. At the same time, close contacts between protein molecules belonging to different particles are expected because the solvent content is low. This means that it is necessary to

include a large number of reflections in the synthesis calculation in order to obtain separated protein regions for different particles. This complicates starting the phasing, when quite a small number of reflections could be involved in the search. Another feature is the presence of a region with a level of electron density lower than that of the solvent. Therefore, one can expect that in the correct low-resolution synthesis the regions of lowest values also consist of blobs and that the number of such 'negative' blobs is equal to the number of particles in the unit cell. Furthermore, the inner cores are separated in space much better than the protein shells, giving a chance to see corresponding regions as being separated even when a small number of very low resolution reflections is phased successfully.

The phasing strategy is summarized in Table 9. We started with the generation of uniformly distributed phases for seven reflections in the 150 Å resolution zone. Those variants resulting in four compact blobs of lowest density in a 150 Å resolution Fourier synthesis were selected and averaged, producing the start-phase values $\{\varphi_1^{\text{best}}(\mathbf{h})\}$ for the second step. In the second and subsequent steps, the phases for acentric reflections were generated as values distributed in accordance with the Von Mizes distribution,

$$P(\varphi) \propto \exp[t \cos(\varphi - \varphi_{j-1}^{\text{best}})], \qquad (20)$$

where $\{\varphi_{j-1}^{\text{best}}(\mathbf{h})\}$ are the average phases defined in the previous cycle and the parameter $t = t(\mathbf{h})$ is defined individually for every reflection from the condition that the expected value of cosine of phase deviation from the start phase is equal to the corresponding figure of merit

$$\langle \cos[\varphi(\mathbf{h}) - \varphi_{j-1}^{\text{best}}(\mathbf{h})] \rangle = m_{j-1}(\mathbf{h}). \qquad (21)$$

In a more explicit form this equation may be written as

$$\frac{I_1[2t(\mathbf{h})]}{I_0[2t(\mathbf{h})]} = m_{j-1}(\mathbf{h}), \qquad (22)$$

where I_0 and I_1 are the modified Bessel functions. Centric phases were generated randomly as $\varphi_{j-1}^{\text{best}}(\mathbf{h})$, with probability $0.5[1 + m_{j-1}(\mathbf{h})]$, or alternatively as $\varphi_{j-1}^{\text{best}}(\mathbf{h}) + \pi$. In the first cycle, the only selection criterion was the condition that low-density regions obtained in a 150 Å resolution synthesis consisted of four blobs of equal finite volumes. In further cycles, additional restrictions were involved which required the same property for regions obtained in the syntheses calculated with the same phase sets but at different resolutions. At every step, about 100 phase sets possessing the desired topological properties were selected and averaged to obtain the start phases for the next step. Optimal alignments in accordance with the permitted origin shifts (Lunin & Lunina, 1996) were performed before averaging. The set of reflections defining the alignment varied with the progress of the phasing (Table 9). The cluster analysis usually revealed a single cluster of phase sets with a few extra sets. Thus, the problem of the choice of the best cluster did not appear in this investigation, as has already been mentioned above.

Starting from cycle 10, additional requirements were included which restricted the high-density regions obtained in the 27 Å resolution syntheses. When starting from the set $\varphi_9^{\text{best}}(\mathbf{h})$ we failed to generate even one phase set resulting in the high-density region composed of four blobs only. Therefore, the restrictions were formulated in a weaker form as the condition that the number of blobs in the high density region does not exceed a given limit. After two steps another requirement was introduced: the high-density region must contain four separated (and equal) largest blobs and any number of smaller 'drops'. The number of these additional drops was also restricted during the last stages. The process was completed with the search for the phases which result in low- and high-density regions consisting of four blobs each. Fig. 13 shows the result of the phasing, which is similar in general to those obtained by other approaches (see §3 and §4 above and Fig. 11), but provides a more detailed picture.

The first steps required relatively small CPU time and were carried out on a PC with a Pentium II processor. When the restrictions on the high-density region were involved, the CPU time requirements grew significantly. In the last stages of the analysis, a specially developed computer program was used which was run on a computer with a parallel architecture and message-passing interface support.

The phasing procedure was performed both for the $C222_1$ and the $C2$ space groups. A non-standard setting ($C211$, unique axis a) was chosen for the $C2$ space group in order to have it as a proper subgroup in $C222_1$. Figs. 14 and 15 present the results of phasing in the $C2$ space group. When compared with Fig. 4, these figures show the loss of radial symmetry inherent in the phases obtained with the use of three-shell model (§3.4).

7. Discussion

The knowledge of the molecular structure of the LDL particle is of great interest for medical investigations. Unfortunately, currently available LDL crystals do not diffract at high resolution and do not allow the application of standard crystallo-
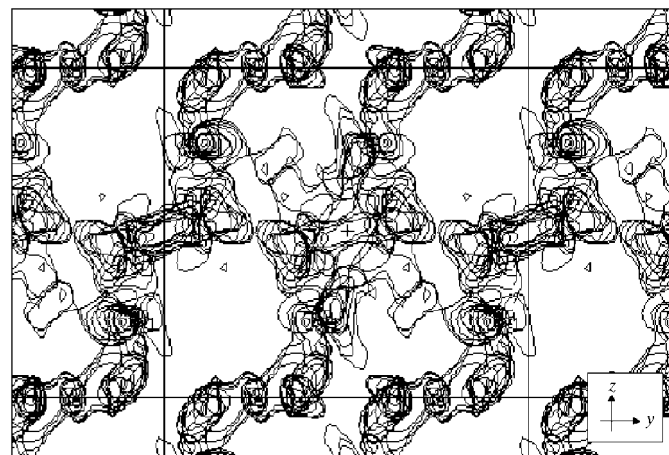
Figure 13
Density projection along the x axis for a part of the unit cell for connectivity-phased Fourier synthesis at 27 Å resolution in the $C2$ space group. The high-density region is shown.

graphic techniques. Additional difficulties arise owing to the very dense crystal packing and to the presence of several components with quite different mean densities. Thus, data sets from LDL crystals constitute a challenge for *ab initio* phasing methods. Several approaches reported in this article have resulted in consistent electron-density maps. Analysis of the low-resolution data from the LDL crystals demonstrates the kind of information which can be obtained, even in quite unfavourable conditions, of an extremely dense form of macromolecular packing in the crystal.

These studies show the robustness of several general *ab initio* phasing procedures. Such techniques as trigonometric structure-factor analysis and modelling by spherical shells were also efficient in this analysis; however, the results obtained with the connectivity analysis were of superior quality. These methods can similarly be applied in other structural investigations.

The crystallographic analysis at its current stage essentially provides the position of the particle in the crystal and some very general features. The resulting overall shell structure of the LDL particle is consistent with models obtained by small-angle scattering. However, the electron density does not show a radial symmetric particle but supports a particle with a pseudo-cylindrical shape, in good agreement with the

electron-microscopy pictures obtained by van Antwerpen *et al.* (1997) and Orlova *et al.* (1999). Both three-dimensional structures correspond to a LDL particle with an ordered lipid core, *i.e.* below the phase-transition temperature of the cholesterol ester core. The distribution of electron-density levels on the particle surface is asymmetric and regions of highest electron density are interpreted as correlating with the distribution of apoB. Areas of reduced (but still high) electron density on the surface could correspond to the phospholipid monolayer. The current resolution of the LDL complex does not yet allow any structural detail at the level of the secondary-structure elements to be seen. It must be noted too that the obtained 27 Å resolution Fourier synthesis corresponds to the electron-density distribution in the whole unit cell. Not all closely contacted surfaces of symmetry-linked particles may be separated unambiguously at this stage of the study. Nevertheless, the electron density inside the particle core shows layer-like structures of low density (Fig. 14) most probably corresponding to the acyl chains of layers of cholesterolester. We assume that these layers are formed during the phase transition of the LDL core. To obtain additional information on the structure of the LDL particle at physiological temperature, *i.e.* above the transition temperature, c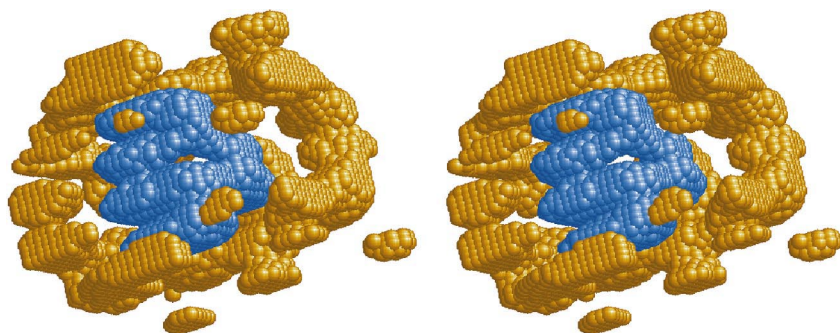rystallization conditions at 310 K are currently being established. Future studies will focus on further analysis of the LDL data sets, the mapping of certain epitopes on the LDL particle by specific labels and on improving the resolution of the crystals.

Figure 14
27 Å resolution crystallographic image of the LDL particle (stereoview) found *ab initio* using the connectivity-based approach. Low density in the core is shown in blue and high density is shown in yellow.

Figure 15
27 Å resolution crystallographic images of the LDL particle found *ab initio* using the connectivity-based approach. Low density in the core is shown in blue and high density is shown in yellow.

APPENDIX *A*
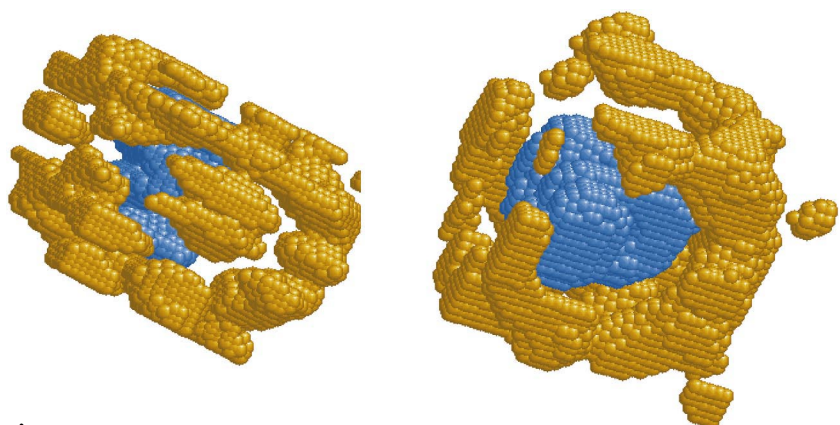## Structure factors for the particles possessing internal symmetry

The structure factors corresponding to a molecular crystal are defined both by the distribution of the electron density in an isolated molecule and by the arrangement of molecules in the unit cell. In some cases, the equations derived below permit an estimate of the magnitudes of structure factors corresponding to an isolated molecule starting from the magnitudes observed for the whole crystal.

Let $F_{ind}(\mathbf{h})$ be complex structure factors corresponding to an isolated molecule with the centre of gravity at the origin of the unit cell, $\mathbf{t}_0$ be the position of the centre of gravity of this molecule in the unit cell and $\Gamma = \{(\mathbf{R}_\nu, \mathbf{t}_\nu)\}_{\nu=1}^n$ be the space group of the crystal, where $(\mathbf{R}_1, \mathbf{t}_1)$ stands for the identity transformation. The structure factors corresponding to the crystal structure are now

$$\mathbf{F}(\mathbf{h}) = \sum_{\nu=1}^n \mathbf{F}_{ind}(\mathbf{R}_\nu^T \mathbf{h}) \exp[2\pi i(\mathbf{h}, \mathbf{R}_\nu \mathbf{t}_0 + \mathbf{t}_\nu)].$$

$$(23)$$

To transform this equation we denote $a_\nu = \mathbf{F}_{\mathrm{ind}}(\mathbf{R}_\nu^T \mathbf{h})$, $b_\nu = \exp[2\pi i(\mathbf{h}, \mathbf{R}_\nu \mathbf{t}_0 + \mathbf{t}_\nu)]$,

$$\mathbf{F}_{\mathrm{ave}}(\mathbf{h}) = \frac{1}{n}\sum_{\nu=1}^{n} \mathbf{F}_{\mathrm{ind}}(\mathbf{R}_\nu^T \mathbf{h}), \tag{24}$$
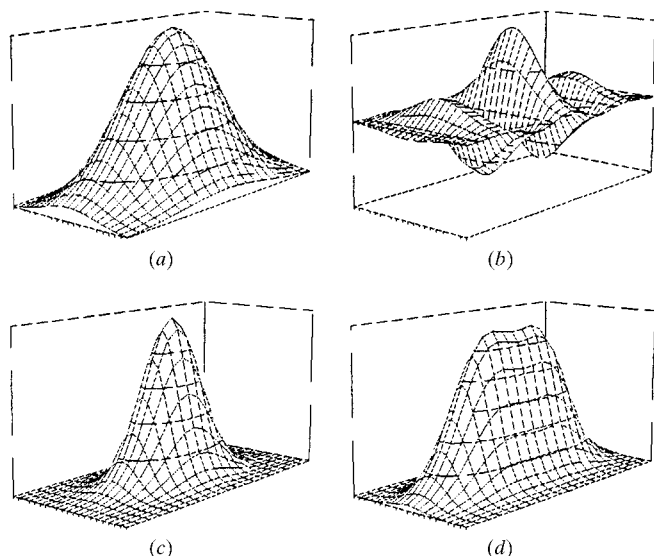
$$\mathbf{T}(\mathbf{h}) = \sum_{\nu=1}^{n} \exp[2\pi i(\mathbf{R}_\nu \mathbf{t}_0 + \mathbf{t}_\nu)] \tag{25}$$
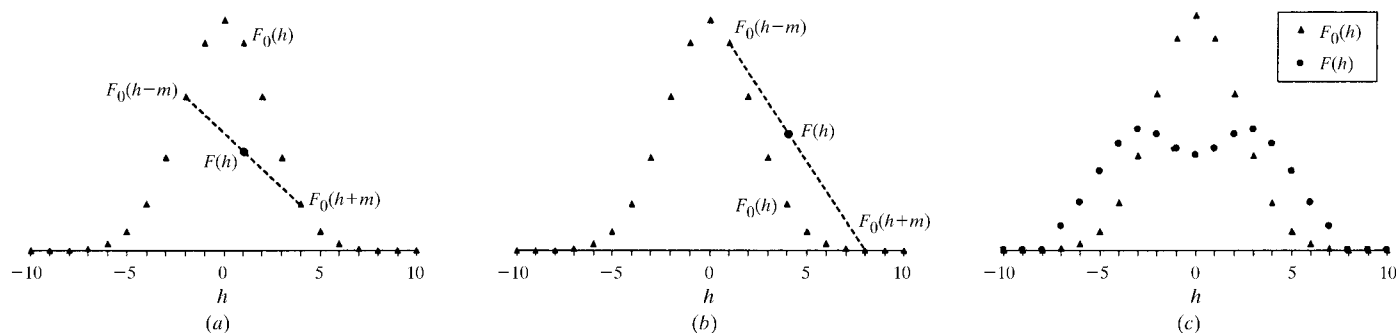
and apply the algebraic identity

$$\sum_{\nu=1}^{n} a_\nu b_\nu = \left(\frac{1}{n}\sum_{\nu=1}^{n} a_\nu\right)\left(\sum_{\mu=1}^{n} b_\mu\right) + \frac{1}{n}\sum_{\nu=1}^{n}\sum_{\mu=\nu+1}^{n}(a_\nu - a_\mu)(b_\nu - b_\mu), \tag{26}$$

which can be verified by inspection.

The result obtained is

Figure 16
The effect of periodical modulation (equations 29 and 30) on two-dimensional Gaussian distribution: (*a*) the Gaussian distribution $\rho_0(x, y)$; (*b*) the modulated distribution $\rho(x, y)$; (*c*) the Fourier transformation of $\rho_0(x, y)$ distribution; (*d*) the Fourier transformation of the modulated distribution.

$$\mathbf{F}(\mathbf{h}) = \mathbf{F}_{\mathrm{ave}}(\mathbf{h})\mathbf{T}(\mathbf{h}) + \frac{1}{n}\sum_{\nu=1}^{n}\sum_{\mu=\nu+1}^{n}\left([\mathbf{F}_{\mathrm{ind}}(\mathbf{R}_\nu^T \mathbf{h}) - \mathbf{F}_{\mathrm{ind}}(\mathbf{R}_\mu^T \mathbf{h})] \right. \tag{27}$$
$$\left. \times \{\exp[2\pi i(\mathbf{h}, \mathbf{R}_\nu \mathbf{t}_0 + \mathbf{t}_\nu)] - \exp[2\pi i(\mathbf{h}, \mathbf{R}_\mu \mathbf{t}_0 + \mathbf{t}_\mu)]\}\right).$$

If the electron-density distribution in an isolated molecule possesses approximately the symmetry of the point group $\{\mathbf{R}_\nu\}_{\nu=1}^{n}$, then $\mathbf{F}_{\mathrm{ave}}(\mathbf{h}) \simeq \mathbf{F}_{\mathrm{ind}}(\mathbf{h})$ and the first multipliers in the double sum are small and the approximate equation may be written as

$$\mathbf{F}_{\mathrm{ind}}(\mathbf{h}) = \mathbf{F}(\mathbf{h})/\mathbf{T}(\mathbf{h}). \tag{28}$$

The validity of this equation depends on the accuracy of the internal symmetry of the molecule.

Obviously, not all reflections may be corrected by means of (28), only those where $\mathbf{T}(\mathbf{h}) \neq 0$.

APPENDIX *B*
The anisotropy of the diffraction pattern for periodically modulated density distributions

A very simple model of a density modulation enables an explanation of some features of the diffraction pattern for the LDL particle.

Let us suppose that the distribution of the electron density in the particle under study may be approximately described as

$$\rho(x, y, z) \simeq \mu(x)\rho_0(x, y, z), \tag{29}$$

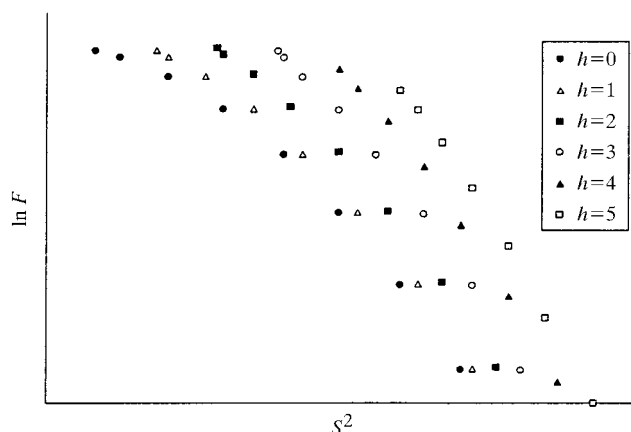where the function $\rho_0(x, y, z)$ is a smooth function possessing spherical symmetry and

$$\mu(x) = \cos(2\pi mx) = \tfrac{1}{2}[\exp(2\pi imx) + \exp(-2\pi imx)]. \tag{30}$$

Fig. 16 shows the effect of the modulation schematically.

The structure factors may be represented in this case as

$$\begin{aligned}\mathbf{F}(h, k, l) &= \Im(\rho) = \Im(\mu) * \Im(\rho_0) \\ &= \tfrac{1}{2}[\delta(h - m, k, l) + \delta(h + m, k, l)] * \Im(\rho_0) \\ &= \tfrac{1}{2}[\mathbf{F}_0(h - m, k, l) + \mathbf{F}_0(h + m, k, l)].\end{aligned} \tag{31}$$

It follows from the last equation that for fixed values of k and l the structure factor $\mathbf{F}(h, k, l)$ will be enhanced in comparison with the $\mathbf{F}(h, k, l)$ for relatively large h indexes and will be weakened when h is small (Fig. 17). At the same time, the structure-factor magnitudes fall at the usual rate with

Figure 17
Modulation (equations 29 and 30) of the electron density decreases structure-factor magnitudes for relatively small indexes h (*a*) and increases them for large h (*b*). Structure factors corresponding to a one-dimensional Gaussian distribution are marked by triangles. The resulting magnitude values are shown (*c*) by dots.

Figure 18
Two-dimensional example of an anisotropy of a diffraction pattern. Structure-factor magnitudes for different set of reflections are shown *versus* the s^2 value. The anisotropy results in the situation that for every resolution shell the strongest reflections are those possessing the maximum h value possible.

increasing k and l values. This results in an anisotropy of the distribution of magnitudes $\mathbf{F}(h, k, l)$ in reciprocal space: in every resolution shell, the strongest reflections will be those possessing the maximum h value possible for the particular resolution shell. Fig. 16 shows the distribution of magnitudes of structure factors for a two-dimensional example and Fig. 18 shows the effect of 'leading' of reflections with the maximum possible value for h indices over other reflections in corresponding resolution shells.

While being extremely simple, these examples enable qualitative explanation to be given of a possible source of features of anisotropy of the diffraction pattern obtained for the LDL particles. The comparison of the diffraction patterns for the LDL (Fig. 7) with Fig. 15 and Fig. 6 with Fig. 18 shows that the diffraction anomalies for the LDL may to some extent be explained by possible modulation of the density distribution similar to (29 and 30) modulation.

References

Anderson, T. A., Levitt, D. G. & Banaszak, L. J. (1998). *Structure*, **6**, 895–909.

Antwerpen, R. van, Chen, G. C., Pullinger, C. R., Kane, J. P., LaBelle, M., Krauss, R. M., Luna-Chavez, C., Forte, T. M. & Gilkey, J. C. (1997). *J. Lipid Res.* **38**, 659–669.

Baker, D., Krukowski, A. E. & Agard, D. A. (1993). *Acta Cryst.* D**49**, 186–192.

Baumstark, M. W., Berg, A., Halle, M., Rensing, U. F. E., Roskamm, H. & Keul, J. (1994). *Atherosclerosis*, **109**, 29–30.

Baumstark, M. W., Kreutz, W., Berg, A., Frey, I. & Keul, J. (1990). *Biochim. Biophys. Acta*, **1037**, 48–57.

Baumstark, M. W., Kreuz, W., Berg, A. & Keul, J. (1991). *Adv. Exp. Med. Biol.* **285**, 123–130.

Brown, M. S. & Goldstein, J. L. (1986). *Science*, **232**, 34–47.

Chan, L. (1992). *J. Biol. Chem.* **267**, 25621–25624.

Chatterton, J. E., Phillips, M. L., Curtiss, L. K., Milne, R., Fruchart, J. C. & Schumaker, V. N. (1995). *J. Lipid Res.* **36**, 2027–2037.

Chatterton, J. E., Phillips, M. L., Curtiss, L. K., Milne, R. W., Marcel Y. L. & Schumaker, V. N. (1991). *J. Biol. Chem.* **266**, 5955–5962.

Chen, G. C., Liu, W., Duchateau, P., Allaart, J., Hamilton, R. L., Mendel, C. M., Lau, K, Hardman, D. A., Frost, P. H. & Malloy, M. J. (1994). *J. Biol. Chem.* **269**, 29121–29128.

Deckelbaum, R. J., Shipley, G. G. & Small, D. M. (1977). *J. Biol. Chem.* **252**, 744–754.

Diederichs, K. & Karplus, P. A. (1997). *Nature Struct. Biol.* **4**, 269–275.

Gordon, T., Kannel, W. B., Castelli, W. P. & Dawber, T. R. (1981). *Arch. Intern. Med.* **141**, 1128–1131.

Gulik-Krzywicki, T., Yates, M. & Aggerbeck, L. P. (1979). *J. Mol. Biol.* **131**, 475–484.

Kabsch, W. (1993). *J. Appl. Cryst.* **26**, 795–800.

Knott, T. J., Pease, R. J., Powell, L. M., Wallis, S. C., Rall, S. C. Jr, Innerarity, T. L., Blackhart, B., Taylor, W. H., Marcel, Y., Milne, R., Johnson, D., Fuller, M., Lusis, A. J., McCarthy, B. J., Mahley, R. W., Levy-Wilson, B. & Scott, J. (1986). *Nature (London)*, **323**, 734–738.

Krauss, R. M. (1995). *Am. J. Cardiol.* **75**, 53B–57B.

Krul, E. S., Kleinman, Y., Kinoshita, M., Pfleger, B., Oida, K., Law, A., Scott, J., Pease, R. & Schonfeld., G. (1988). *J. Lipid Res.* **29**, 937–947.

Laggner, P., Kostner, G. M., Degovics, G. & Worcester, D. L. (1984). *Proc. Natl Acad. Sci. USA*, **81**, 4389–4393.

Laggner, P., Kostner, G. M., Rakusch, U. & Worcester, D. (1981). *J. Biol. Chem.* **256**, 11832–11839.

Laggner, P. & Müller, K. W. (1978). *Quart. Rev. Biophys.* **11**, 371–425.

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

Lunin, V. Yu., Lunina, N. L., Petrova, T. E., Skovoroda, T. P., Urzhumtsev, A. G. & Podjarny, A. D. (2000). *Acta Cryst.* D**56**, 1223–1232.

Lunin, V. Yu., Lunina, N. L., Petrova, T. E., Urzhumtsev, A. G. & Podjarny, A. D. (1998). *Acta Cryst.* D**54**, 726–734.

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

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

Lunin, V. Y., Lunina, N. L. & Urzhumtsev, A. G. (2000). *Acta Cryst.* A**56**, 375–382.

Lunin, V. Yu., Podjarny, A. D. & Urzhumtsev, A. (2000). *Advances in Structure Analysis. XVIIIth European Crystallography Meeting, 16–20 August 1998, Prague, Czech Republic*, edited by R. Kuzel. In the press.

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

Luzzati, V., Tardieu, A. & Aggerbeck, L. P. (1979). *J. Mol. Biol.* **131**, 435–473.

Müller, K., Laggner, P., Glatter, O. & Kostner, G. (1978). *Eur. J. Biochem.* **82**, 73–90.

Orlova, E. V., Sherman, M. B., Chiu, W., Mowri, H., Smith, L. C. & Gotto, A. M. Jr (1999). *Proc. Natl Acad. Sci. USA*, **96**, 8420–8425.

Prassl, R., Chapman, J. M., Nigon, F., Sara, M., Eschenburg, S., Betzel, C., Saxena, A. & Laggner, P. (1996). *J. Biol. Chem.* **271**, 28731–28733.

Ritter, S., Diederichs, K., Frey, I., Berg, A., Keul, J. & Baumstark, M. W. (1999). *J. Cryst. Growth*, **196**, 344–349.

Ritter, S., Frey, I., Diederichs, K., Grathwohl, D., Keul, J. & Baumstark, M. W. (1997). *Proteins*, **28**, 293–297.

Segrest, J. P., Jones, M. K. & Dashti, N. (1999). *J. Lipid Res.* **40**, 1401–1416.

Segrest, J. P., Jones, M. K., Mishra, V. K., Anantharamaiah, G. M. & Garber, D. W. (1994). *Arterioscler. Thromb.* **14**, 1674–1685.

Urzhumtsev, A. G., Lunina, N., Skovororda, T., Podjarny, A. & Lunin, V. Y. (2000). *Acta Cryst.* D**56**, 1233–1244.

Vernoslova, E. A. & Lunin, V. Yu. (1993). *J. Appl. Cryst.* **26**, 291–294.

Wiklund, O., Dyer, C. A., Tsao, B. P. & Curtiss, L. K. (1985). *J. Biol. Chem.* **260**, 10956–10960.

Yang, C.-Y., Chen, S.-H., Gianturco, S. H., Bradley, W. A., Sparrow, J. T., Tanimura, M., Li, W.-H., Sparrow, D. A., DeLoof, H., Rosseneu, M., Lee, F.-S., Gu, Z.-W., Gotto, A. M. Jr & Chan, L. (1986). *Nature (London)*, **323**, 738–742.