Small-angle scattering and its interplay with crystallography, contrast variation in SAXS and SANS

Heinrich B. Stuhrmann

GKSS Forschungszentrum, Geesthacht, Germany, and Institut de Biologie Structurale Jean-Pierre Ebel, CEA/CNRS/UJF, F-38027 Grenoble, France. Correspondence e-mail: heinrich.stuhrmann@orange.fr

Methods of contrast variation are tools that are essential in macromolecular structure research. Anomalous dispersion of X-ray diffraction is widely used in protein crystallography. Recent attempts to extend this method to native resonant labels like sulfur and phosphorus are promising. Substitution of hydrogen isotopes is central to biological applications of neutron scattering. Proton spin polarization considerably enhances an existing contrast prepared by isotopic substitution. Concepts and methods of nuclear magnetic resonance (NMR) become an important ingredient in neutron scattering from dynamically polarized targets.

1. Introduction

For a number of reasons, structural biology has been the field where the interplay of small-angle scattering (SAS) with crystallography has been most visible. Very soon after the determination of the structure of myoglobin by Kendrew *et al.* (1960), the question was raised whether this structure from X-ray crystallography would also exist in solution. To decide this question, small-angle X-ray scattering (SAXS) was performed on solutions of myoglobin. First experiments of SAXS in 1964 (Kirste & Stuhrmann, 1967) showed that the radius of gyration of myoglobin in solution appeared to be slightly larger than that calculated from the crystallographic model. Did the structure of myoglobin really differ from that in the crystal, as a referee from *Nature* would have liked to make us say? Or was there simply a shortcoming in the interpretation of the SAXS data? At that time, we hesitated to give a definite answer, and we looked more carefully into the analysis of the data from solution scattering. It was fairly soon realized that myoglobin embedded in a solvent would give rise to a profile of SAXS that is different from that of the 'free' molecule (Kirste & Stuhrmann, 1967). This was the way to contrast variation in SAXS. Its basic idea had been expressed for protein crystallography well before by Bragg & Perutz (1952):

'Suppose a region [= volume] v of constant shape to be occupied by the molecule which for the present purposes is defined as the region into which salt does not penetrate when a salt solution is substituted for water. It is assumed that the density of the liquid outside v is uniform. The effect on the value of F [= structure factor] of increasing density everywhere outside v is equivalent to that of decreasing density by an equal amount everywhere inside v, since a uniform distri-

bution of density has no effect on F. It follows that the values F(water) − F(salt) are the F values for a region v of uniform density equal to the difference between the electron concentration of the salt solution and water.'

This argument was introduced by Bragg & Perutz (1952) to explain reversible changes in the intensities of low-order X-ray reflections from a haemoglobin crystal and to deduce from these the shape and orientation of the haemoglobin molecules occupying the volume v within the unit cell. It is also the description of what later became known as contrast variation in small-angle scattering from macromolecules in solution (Stuhrmann & Kirste, 1965).

Contrast variation by solvent exchange is a low-resolution method. It is for this reason it has been superseded by a method using heavy-metal derivatives, which allowed the first determination of a protein structure at nearly atomic resolution (Kendrew *et al.*, 1960). With the advent of powerful sources of thermal neutrons and of X-rays from storage rings, methods of anomalous X-ray diffraction and of isotopic substitution in neutron scattering became powerful tools in condensed-matter research.

The tools of contrast variation being equal both for crystallography and for SAS, the principal difference between these two methods originates from the order of the molecules in the sample. It is perfect in crystals and non-existent in molecules in solution. Hence, the structural information that can be extracted from the scattered intensity is vastly different: there is structure determination from crystals and there is model building at low resolution in SAS from monodisperse solutions. We will put this situation into a mathematical form. Then we will turn to the methods of contrast variation in both crystallography and SAS.

 181

feature articles

2. Uniqueness

The relation between diffraction from a single crystal and scattering from a dilute solute solution of equal randomly oriented particles is most easily shown in terms of spherical harmonics. The scattering amplitude $F(\mathbf{Q})$ is

$$
\begin{aligned}
F(\mathbf{Q}) &= \sum_{n=1}^{N} f_n \exp[i(\mathbf{Q} \cdot \mathbf{r}_n)] \\
&= \sum_{n=1}^{N} f_n \sum_{l=0}^{\infty} i^l j_l(Qr_n) \sum_{m=-l}^{l} Y_{l,m}^*(\theta_n, \varphi_n) Y_{l,m}(\Omega) \\
&\equiv \sum_{l=0}^{\infty} \sum_{m=-l}^{l} F_{l,m}(Q) Y_{l,m}(\Omega),
\end{aligned}
\tag{1}
$$

where r_n, θ_n, φ_n are the polar coordinates of the nth atom. The first line refers to $\mathbf{r}_n$ in Cartesian coordinates. j_l is the spherical Bessel function of the lth order and $Y_{l,m}$ are the spherical harmonics. Q is the modulus of the scattering vector $\mathbf{Q}$, and Ω is its unit vector in Fourier space. Note that in the crystal one sums over the N atoms of the entire unit cell, which usually has its own space-group symmetries, whereas for isolated molecules one sums over a molecule which has only point-group symmetries at most. The intensity diffracted by a crystal depends on the absolute square of the amplitude,

$$
S(\mathbf{Q}) \propto |F(\mathbf{Q})|^2,
\tag{2}
$$

where $\mathbf{Q}$ meets the Laue condition. In solution, the observable scattering intensity is the sum of the intensities scattered by each molecule according to its random orientation.

$$
I(Q) \propto \sum_{l=0}^{\infty} \sum_{m=-l}^{l} |F_{l,m}(Q)|^2.
\tag{3}
$$

Each multipole in real space, $\rho_{l,m}(r) Y_{l,m}(\theta, \varphi)$, gives rise to its own scattering function $|F_{l,m}(Q)|^2$. An interesting way to create structures leaving $I(Q)$ unchanged starts from the definition of partial structures $\rho_l(\mathbf{r})$.

$$
\rho(\mathbf{r}) = \sum_{l=0}^{\infty} \rho_l(\mathbf{r}) \equiv \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \rho_{l,m}(r) Y_{l,m}(\theta, \varphi).
\tag{4}
$$

The rotation of a partial structure $\rho_l(\mathbf{r})$ by an arbitrary angle with respect to the rest of the structure has no influence on $I(Q)$ but it does change the scalar field $\rho(\mathbf{r})$ (Stuhrmann, 1970a). Any analysis of $I(Q)$ aiming at a possibly unique solution will rely on assumptions that restrict the number of structures, the scattering curves of which agree with measured data.

In some cases, the symmetry of the particles studied in solution might be known from other techniques, e.g. from electron microscopy. A relatively small number of multipoles might suffice to describe isometric structures, like viruses (Finch & Holmes, 1967). Ideally, the structure might be spherical, which would allow a structure analysis which is quite close to that in crystallography. Another assumption is a particle of uniform density as was proposed by Bragg & Perutz (1952), which will be discussed in §§4 and 5.

Table 1
Variation of the intensity of X-ray and neutron scattering with contrast as described by (6).

	B	C	D
Anomalous diffraction from single crystals: $f = f_0 + f' + if''$	f'	f''	$f'^2 + f''^2$
Anomalous scattering from powder/fibre diffraction	f'	0	$f'^2 + f''^2$
Nuclear spin polarization, P	$\pm P$		P^2
Isotopic/chemical substitution	X		X^2

3. Contrast variation

Generalization of the recipe given by Bragg & Perutz (1952) makes the scattering amplitude $F(\mathbf{Q})$ assume the following form:

$$
F(\mathbf{Q}) = U(\mathbf{Q}) + X V(\mathbf{Q}).
\tag{5}
$$

$U(\mathbf{Q})$ is the amplitude of the molecule *in vacuo*. In the case of solvent contrast variation, X would be the negative value of the scattering density (sum of scattering lengths per unit volume) of the solvent and $V(\mathbf{Q})$ would be the amplitude of the macromolecular volume. If the labelled atoms belong to the macromolecule, X will be positive. With resonant (anomalous) scattering, the atomic scattering factor $f = f_0 + f' + if''$ will have to be used. From (5) and (2), one obtains the intensity $S(\mathbf{Q})$ diffracted by a crystal:

$$
\begin{aligned}
S(\mathbf{Q}) = A|U(\mathbf{Q})|^2 &+ B \operatorname{Re}\{U(\mathbf{Q})V^*(\mathbf{Q})\} \\
&+ C \operatorname{Im}\{U(\mathbf{Q})V^*(\mathbf{Q})\} + D|V(\mathbf{Q})|^2.
\end{aligned}
\tag{6}
$$

The coefficients B, C and D are given in Table 1. The coefficient A takes into account the non-resonant atomic scattering factors.

The internal structure appears most clearly when the scattering-length density of the dissolved particle, ρ_p, equals that of the solvent ρ_s, i.e. when the contrast $\rho = \rho_p - \rho_s$ vanishes. The variation of the intensity of SAS from solutions is often written in the following form:

$$
I(Q) = \rho^2 I_C(Q) + \rho^2 I_{C0}(Q) + I_0(Q).
\tag{7}
$$

With the amplitude $F_0(\mathbf{Q})$ at the matching point and $F_C(\mathbf{Q})$ as the amplitude associated with the contrast ρ, the basic scattering functions of (7) are $I_0(Q) = \langle |F_0(\mathbf{Q})|^2 \rangle$, $I_C(Q) = \langle |F_C(\mathbf{Q})|^2 \rangle$ and $I_{C0}(Q) = \langle 2\operatorname{Re}\{F_0(\mathbf{Q})F_C(\mathbf{Q})\} \rangle$. The brackets $\langle \dots \rangle$ denote the average of intensity from randomly oriented particles. One way to obtain the averaged scattering intensity is to replace the amplitudes by the radial functions of the multipole expansion and to perform the double sum as in (3).

At small Q, the radius of gyration R_g is obtained from a Guinier plot. The variation of R_g with the solvent scattering density, ρ_s, provides an unmistakable feature of the low-resolution architecture of complex particles in solution (Ibel & Stuhrmann, 1975).

$$
R_g^2 = R_\infty^2 + \frac{\alpha}{\rho_p - \rho_s} - \frac{\beta}{(\rho_p - \rho_s)^2}.
\tag{8}
$$

Particles with a high density core (*e.g.* ribosomes, ferritin) will give rise to a negative α, whereas particles with a low-density core (*e.g.* nucleosome core particle, low-density lipoprotein) give rise to a positive α, in a spherical approximation. A non-vanishing β would be due to a dipolar structure, *i.e.* the centres of mass of the different components of the particle would not coincide.

Anomalous SAS from randomly oriented molecules is less powerful than anomalous diffraction from single crystals because at low resolution the effects of f' and f'' are relatively much weaker than at high resolution. The relative change of the intensity, $\Delta I/I$, of anomalous diffraction has been estimated by Crick & Magdoff (1956):

$$\frac{\Delta I}{I} = \left(\frac{2N_r}{N_l}\right)^{1/2}\frac{f_r}{f_l}. \tag{9}$$

N_r = number of resonant atoms per unit cell, N_l = number of non-resonant atoms C, N and O (taken to be molecular weight/14.7), f_r = resonant scattering factor, f_l = taken to be 7 electrons, an average of C, N and O).

At low resolution, and in particular with SAS from solutions, $\Delta I/I$ is

$$\frac{\Delta I}{I} = \frac{2N_r f_r}{N_l f_l}. \tag{10}$$
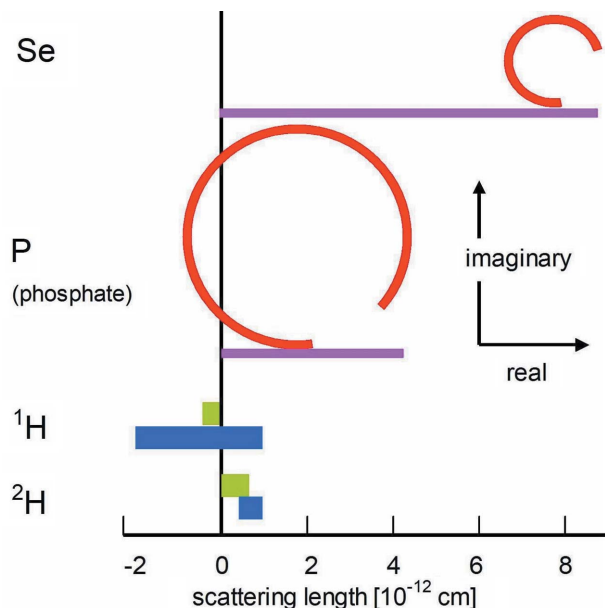
Figure 1
Comparison of complex scattering lengths of resonant X-ray scattering with isotope/spin-dependent neutron scattering. The plot of the imaginary part f'' *versus* the real part of the scattering length, $f_0 + f'$, is understood by the inspection of Fig. 2 which shows the resonant components f' and f'' of phosphorus. The lengths b of neutron scattering from H atoms are real: -0.374×10^{-12} cm for ^{1}H and $+0.667 \times 10^{-12}$ cm for ^{2}H (green). The scattering lengths of both isotopes vary with nuclear polarization P. Using a completely polarized neutron beam, $b({}^1$H$) = (-0.374 \pm 1.456P) \times 10^{-12}$ cm and $b({}^2$H$) = (0.667 \pm 0.28P) \times 10^{-12}$ cm (blue). The sign $\pm$ refers to the polarization direction of the neutron beam. With dynamic nuclear polarization (DNP), P may vary within the limits of -1 and $+1$.

For one resonant atom among 100 non-resonant atoms, the relative change of the intensity due to anomalous diffraction at high resolution would exceed that of anomalous scattering at low Q by a factor of ten. It is for this reason that experiments of anomalous SAXS have remained rare in biological structure research (Kühnholz, 1991). Anomalous SAXS is more often used for the study of materials (Goerigk *et al.*, 2004; Simon, 2007). Examples of the variation of the scattering length of resonant X-ray scattering and of nuclear polarization-dependent neutron scattering are shown in Fig. 1.

4. The solvent phase

Many applications of SAS in solution aim at the determination of macromolecular shapes from $I_C(Q)$ in (7). In protein crystallography, such attempts have been rare and they will be presented as proofs of principle.

4.1. Lipid phase of purple membrane

Purple membrane lends itself to a comparison of different methods of contrast variation using resonant X-ray scattering and isotopic substitution in neutron scattering. While the structure of its membrane protein is well known (Henderson & Unwin, 1975), this is much less the case with the lipid phase. The latter is rich in phosphates and sulfates. Anomalous X-ray scattering from both phosphorus (Fig. 2) and sulfur may therefore elucidate the lipid–protein interaction. Neutron
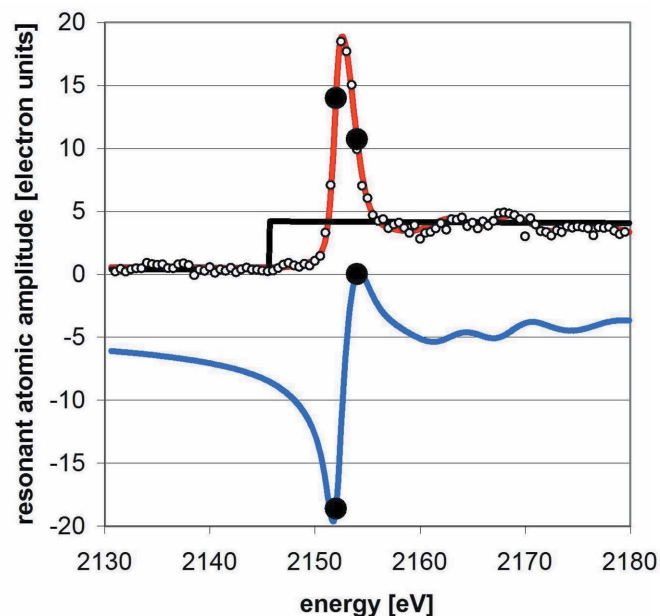
Figure 2
Anomalous dispersion of phosphorus from phospholipids of purple membrane at energies near the K-edge. The red curve denotes the imaginary part f'' of the atomic scattering factor (in anomalous electron units). The real resonant part f' (in blue) is negative, except at 2154 eV (5.755 Å), where it is close to zero. The large circles denote the energies used in anomalous diffraction from purple membrane. The absorption edge of elementary phosphorus is at 2144 eV (5.782 Å) (black line) (Biou *et al.*, 2005).
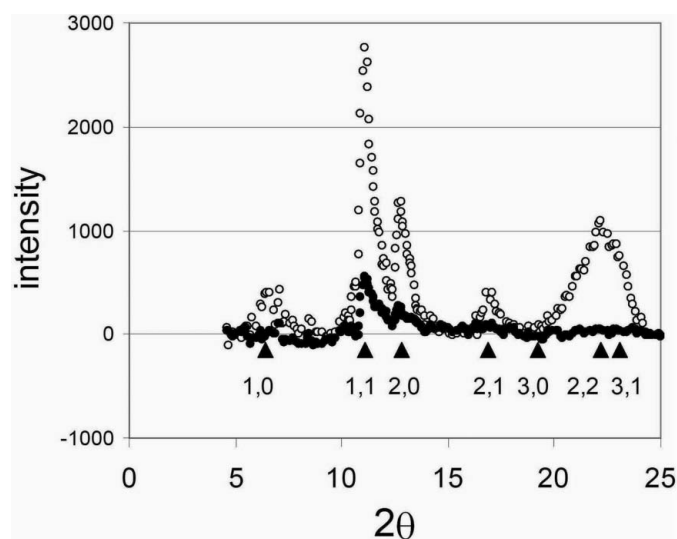
feature articles

scattering in mixtures of H_2O/D_2O reveals the repartition of the aqueous phase in purple membrane.

Anomalous dispersion near the K-edge makes phosphorus an excellent native label in structural studies (Fig. 1). For a purple-membrane stack, strong anomalous diffraction is expected at $E_1 = 2152$ and $E_2 = 2154$ eV, where the difference in f' is largest and the change in absorption (proportional to f'') is smallest. Significant changes in the diffracted intensity are observed at low values of (h, k) (Fig. 3), notably with the $(1, 1)$ and $(2, 0)$ reflections. This result agrees remarkably well with corresponding neutron diffraction data from purple membrane in H_2O/D_2O mixtures (Worcester, 1975; Zaccai & Gilmore, 1979). Phosphorus, as part of the polar head groups of the phospholipids, occupies the space between the proteins as does water (Fig. 4). Both provide a negative print of the hexagonal protein lattice.

At this point, a technical remark seems appropriate. The penetration depth of soft X-rays in nearly any kind of matter is low, e.g. it amounts to 0.02 mm water for 5.7 Å photons. While thin flat samples of purple membrane of that thickness and also solutions in flat sample holders are rather easily prepared and studied by soft X-rays, this is less obvious for protein crystals (Stuhrmann et al., 1997; Biou et al., 2005). The reduced penetration depth imposes some constraints on the design of a set up for soft X-ray diffraction which are not unsurmountable (Djinović Carugo et al., 2005; Biou et al., 2005).

4.2. The solvent phase in proteins, MASC

As pointed out by Bricogne (1993), anomalous scattering at several wavelengths by the solvent of a macromolecular crystal can be used to phase Bragg reflections and thus pave the way to structure determination. The similarity of this approach with that of Bra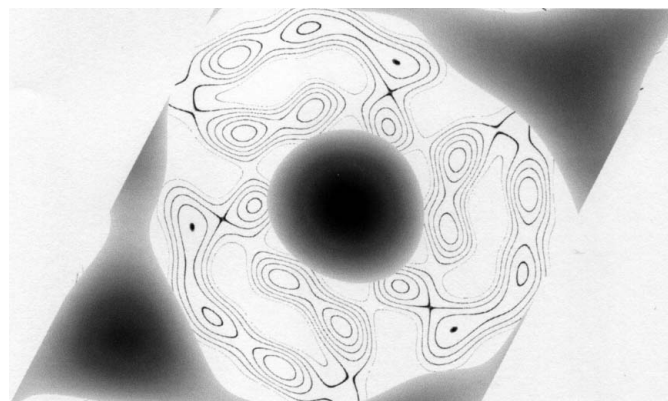gg and Perutz is evident. The new element of the method is that the amplitude of the scattering density of the solvent is a complex number, whence its phasing power as implicitly stated in (6).

The theory of MASC (multiwavelength anomalous solvent contrast) assuming a binary density model of the protein and the solvent volume in the crystal and the first results were reported by Fourme et al. (1995). MASC experiments have been performed on three protein structures of different molecular weight (Ramin et al., 1999). In all cases, ordered anomalous scatterers were found in addition to disordered atoms. When only MASC effects are taken into account, the agreement between experimental and model values is satisfactory at very low resolution ($d > 20$ Å).

While the method of MASC has hardly been used since then, the method of multiwavelength anomalous diffraction (MAD) using selenium (Fig. 1) as a substitute for sulfur in methionine is widely used in protein crystallography (Hendrickson, 1991). Fig. 5 shows the Argand diagram of the $1, 0, -1$ reflection of trypsin crystallized from a solution of ammonium sulfate as mother liquor at wavelengths near the K-edge of sulfur. As the absorption edge of the sulfate ion is shifted by 10 eV to higher energies with respect to that of sulfur in methionine and cystine, the anomalous dispersions of sulfates (MASC) and sulfur of methionine and cystin (MAD) are well separated (Stuhrmann et al., 1997).

4.3. Shapes of proteins from SAS

For a long time, simple shapes, like spheres, hollow spheres, ellipsoids, to mention some, were considered to be sufficient for the interpretation of SAS from compact particles. With a shape scattering function, $I_C(\mathbf{Q})$, obtained cleanly by solvent contrast variation, a more general approach has been introduced. The surface of a particle is described by a single-valued function of the angles θ and φ (Stuhrmann, 1970b,c):

Figure 3
The intensity diffracted by a stack of purple membranes at the energy $E_1 = 2152$ eV (5.761 Å) (open circles) and $E_2 = 2154$ eV (5.755λ) (not shown) and its difference $I(E_1) - I(E_2)$ (black circles) as a function of the scattering angle 2θ (°). Indices (h, k) are marked. The two-dimensional hexagonal plane group is $p3$ with $a = 61$ Å (Biou et al., 2005).
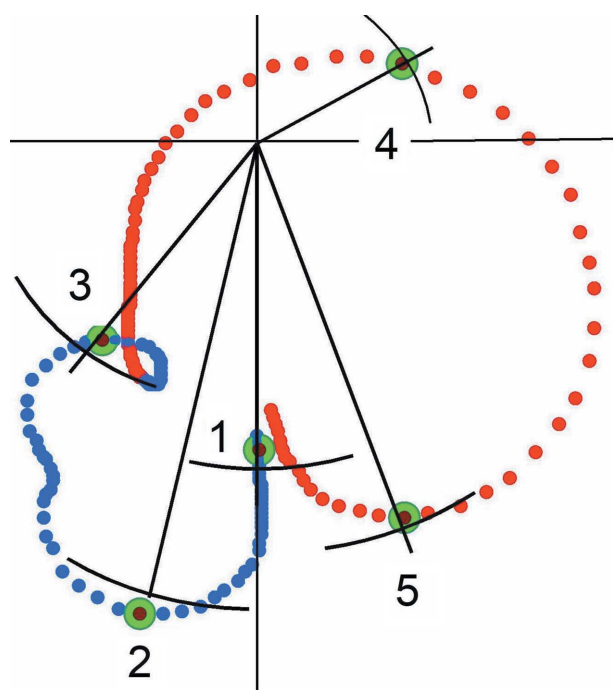

Figure 4
Phospholipid distribution (grey) in the plane diffraction. The difference Fourier map is based on the known structure of bacteriorhodopsin (Seiff et al., 1985) and the intensity differences from anomalous diffraction of X-rays. The unit cell contains three bacteriorhopsin molecules (contour lines) (Biou et al., 2005).

$$F(\omega) = \sum_{l=0}^{L} \sum_{m=-l}^{l} f_{l,m} Y_{l,m}(\theta, \varphi). \qquad (11)$$

The coefficients $f_{l,m}$ are determined by a non-linear optimization procedure that minimizes the R factor between the calculated shape scattering function and the experimental curves (Svergun & Stuhrmann, 1991; Svergun, 1997). Computer simulation on model bodies indicated that the low-resolution shape determination for error-free data is unique, even when very limited ranges are used in the simulated curves (Svergun *et al.*, 1996). Some tests on X-ray scattering from several proteins with known atomic structures in the crystal are shown in Fig. 6. This method is being used successfully (König *et al.*, 1998; Macheroux *et al.*, 1998).

The envelope representation has intrinsic limitations, *e.g.* it is impossible to account for holes inside the particle. A remedy for this shortcoming came from Monte Carlo type methods, which produce a more or less compact particle from a large number of connected scattering centres (spherical beads) in a confined volume (Chacon *et al.*, 1998; Svergun, 1999; Walther *et al.*, 2000; Heller *et al.*, 2002).

Another limitation of the *ab initio* shape-determination methods is the assumption of uniform scattering-length density, which restricts the usable portion of the scattering data (typically up to 20 Å resolution for proteins in the absence of solvent contrast variation). In an alternative approach (Svergun *et al.*, 2001; Svergun, 2007), the protein is represented by an assembly of dummy residues instead of beads, whereby each residue has a form factor equal to that


(a)

(b)
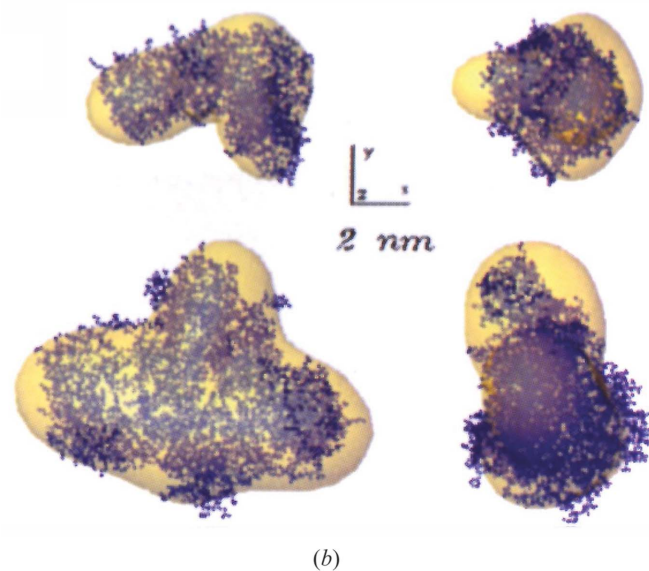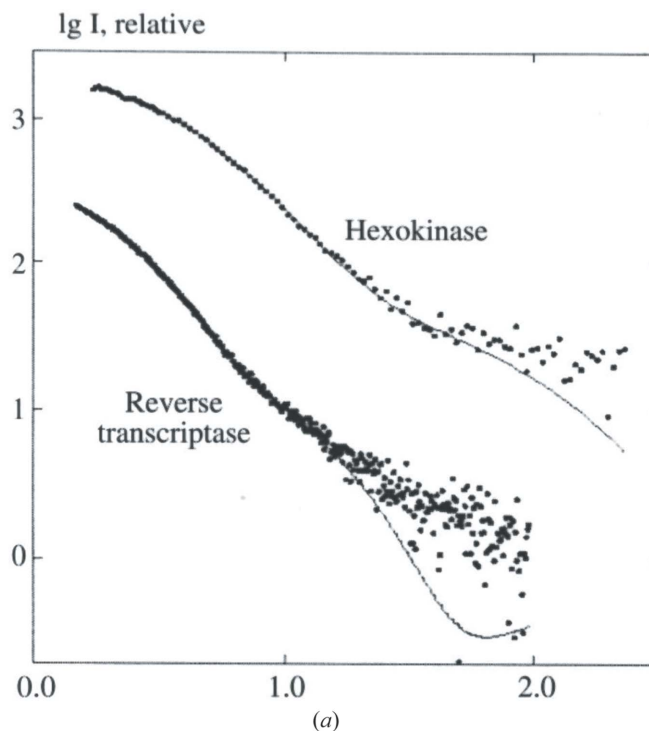
Figure 6
Shape determination of the hexokinase and HIV-1 reverse transcriptase. (*a*) Experimental X-ray scattering data (dots) and the curves calculated from the restored envelopes using $L = 4$ (solid lines). Q on the horizontal axis is given in nm^{-1}. (*b*) Comparison between the envelopes (transparent solids) and the atomic structures in the crystal. Top panel: hexokinase, bottom panel: reverse transcriptase. Right: pictures are rotated $90°$ counterclockwise around **y**. The atomic structures of hexokinase and of the transcriptase are from Bennet & Steitz (1980) and Wang *et al.* (1994), respectively. From Vachette & Svergun (2000). In *Structure and Dynamics of Biomolecules*, edited by Eric Fanchoon *et al.* (2000). By permission of Oxford University Press, http://www.oup.com.

Figure 5
Argand diagram of the $1, 0, -1$ reflection of trypsin. The root of the intensity measured at five wavelengths near the K-edge of sulfur is presented by radii intersected by a short section of a circle. The anomalous dispersion of X-ray scattering calculated from the model is shown by dots, which are blue near the edge of sulfate, and red near the edge of S atoms of trypsine. More widely spaced points correspond to strong anomalous dispersion. The calculated values to be compared with the experimental ones are the large green spheres. $E_1 = 2462$ eV (5.035 Å), $E_2 = 2471$ eV (5.017 Å), $E_3 = 2474$ eV (5.012 Å), $E_4 = 2482$ eV (4.995 Å), $E_5 = 2485$ eV (4.989 Å) (Stuhrmann *et al.*, 1997).

residue in water. The method starts from a randomly distributed gas of dummy residues. The number of dummy residues is usually known *a priori* from the sequence and the method employs simulated annealing to find the coordinates of the residues fitting the experimental data and building a protein-like structure (Davies *et al.*, 2005; Hough *et al.*, 2004; Shi *et al.*, 2005).

As a model consisting of dummy residues is more detailed than a shape of uniform density, the uniqueness of structure determination from SAS data is lost, at least to some extent. Running the Monte Carlo methods from random starts produces somewhat different models yielding nearly identical scattering patterns. These models can be superimposed and averaged to obtain the most probable model (Volkov & Svergun, 2003).

5. The internal structure

The internal structure is more complicated than its envelope. Its elucidation is what is usually called structure determination. It requires the analysis of all accessible basic scattering functions in (6). Measurement of neutron scattering in H_2O/D_2O mixtures provides the easiest access to the basic scattering functions. A more pinpointed attack will rely on labelling an interesting region of a particle either by resonant X-ray scattering or by isotopic substitution in neutron scattering (Table 1). In the latter case, the contrast of the label can be enhanced by nuclear polarization (§6).
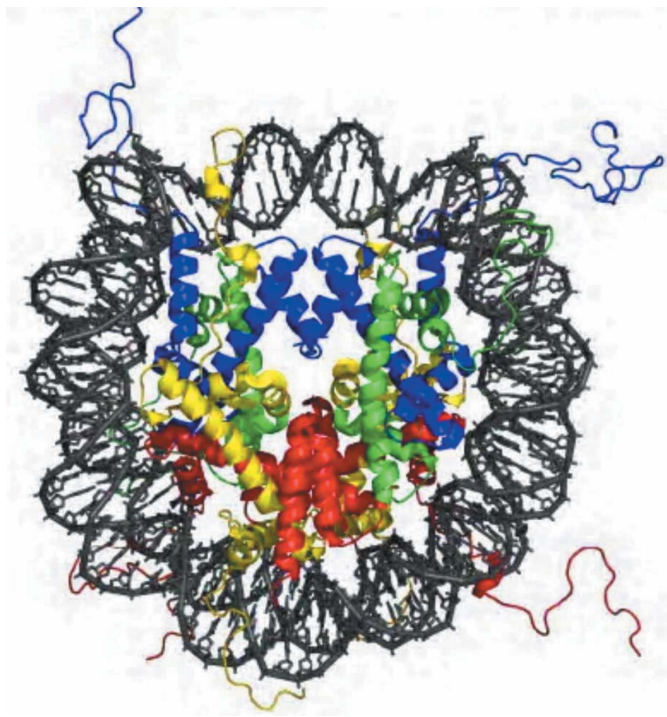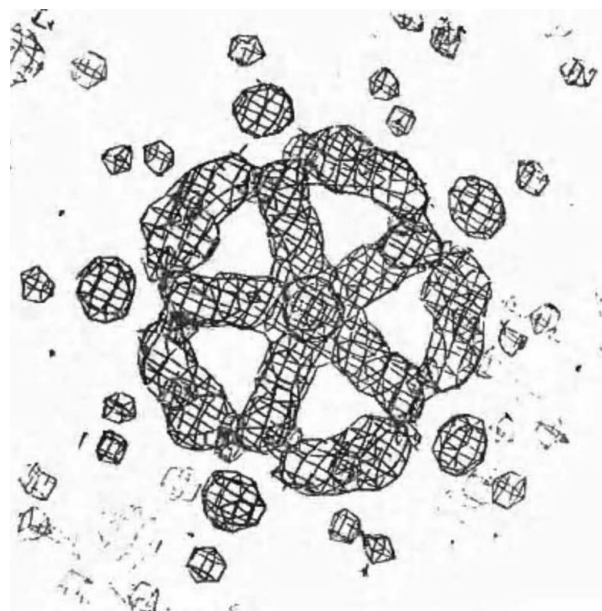
5.1. The nucleosome core particle

The neutron scattering study of the nucleosome core particle by Hjelm *et al.* (1977) is one of the first applications of solvent contrast variation. The particle contains 195 base pairs of DNA and histones H2A, H2B, H3 and H4 (Fig. 7). Measurement of the intensity of neutron small-angle scattering (SANS) in various H_2O/D_2O solvents allowed the contrast dependence of R_g to be examined giving ρ_p, a positive α and $\beta = 0$. As the scattering density of proteins is lower than that of DNA, there is an increased probability of finding the histones in the core region of the particle. The analysis of the basic scattering functions defined by (7) allows the modelling of the particle shape and its internal structure. The spherically averaged structure contains most of its histones in a core of 32 Å radius surrounded by a loosely packed DNA-rich shell resulting in a particle of 52 Å radius on average.

In the next step, neutron diffraction data from crystals of the nucleosome core particle were measured to a resolution of 25 Å (Bentley *et al.*, 1981). Using H_2O/D_2O mixtures containing 39% D_2O and 65% D_2O, the scattering density of the histones and of the DNA, respectively, had been matched. The data were consistent with a model in which 1.8 turns of a DNA superhelix of pitch 27.5 Å and radius 42 Å are wound around the protein core. X-ray diffraction data at 7 Å resolution corroborated this model (O'Halloran *et al.*, 1987). The features of the early low-resolution model derived from neutron scattering remain most striking, even at atomic resolution (Fig. 7) (Davey *et al.*, 2002).

5.2. RNA in satellite tobacco necrosis virus

A number of high-resolution X-ray crystallographic studies of icosahedral viruses has given atomic models of the protein

Figure 7
The model of the nucleosome core particle obtained from X-ray diffraction data at 1.9 Å resolution (Davey *et al.*, 2002, picture from Wikipedia) (http://en.wikipedia.org/wiki/Nucleosomes). The histones are surrounded by a DNA superhelix.

Figure 8
The positive density at 40% D_2O looking down the fivefold axis from the centre of the virus. For the main features of the RNA, see text. Reprinted from Bentley *et al.* (1987). *J. Mol. Biol.* **194**, 129–141. Copyright (1987), with permission from Elsevier.

coat. These results have revealed a remarkable structural similarity in those domains of the protein which are responsible for organizing the subunits on an icosahedral lattice to form the outer protective coat of the virus. By contrast, details of the internal structure, *i.e.* the RNA and the regions of the proteins which interact with the RNA, are lacking. Neutron diffraction in H_2O/D_2O mixtures is the method of choice to complete a virus structure.

The satellite tobacco necrosis virus (STNV) has 60 identical subunits arranged on a $T = 1$ icosahedral lattice to form a shell which encloses a single strand of RNA. In the X-ray crystal structure analysis, nearly all main-chain atoms of the protein coat and most of the side-chain atoms are well ordered in the electron-density map (Jones & Liljas, 1984).

Neutron diffraction from a STNV crystal soaked with an H_2O/D_2O mixture containing 40% D_2O shows essentially the positive density of RNA. A view along the fivefold axis from the virus centre is shown in Fig. 8. Two different structural motifs can be distinguished. The RNA density (motif 1) lies along the edges of each triangle, thus forming the edges of an icosahedron. The RNA density of motif 2 lies on a fivefold axis at a distance of 67 Å from the virus centre. Minor regions of density seen at larger distances correspond to positive fluctuations of the protein density (Bentley *et al.*, 1987).

5.3. The ribosome

There is probably no other particle of which the structure has been studied so intensely as ribosome. Various methods were used to get a microscopic picture of the mRNA controlled protein synthesis.

The shape of the *E. coli* ribosome ($M = 2.4 \times 10^6$ Da) and that of its two unequal subunits with $M = 0.9 \times 10^6$ and $M = 1.5 \times 10^6$ Da, respectively, were obtained by electron microscopy. Quite a number of the over 50 different ribosomal proteins could be visualized by their antibodies using the

method of immune electron microscopy (Stöffler *et al.*, 1988). Neutron scattering in H_2O/D_2O mixtures distinguished clearly between the ribosomal proteins and the ribosomal RNA which comprises two thirds of the particle (Serdyuk & Grenader, 1974; Stuhrmann *et al.*, 1978).

It was in the early 70's when Hoppe (1972) and Engelman & Moore (1972) presented a method for the determination of the architecture of multi-protein complexes, the essential prerequisite being that each of their components would be amenable to selective deuteration. A sufficiently large number of interprotein distances determined by neutron scattering allows the reconstruction of a three-dimensional model by triangulation. The result of 15 years neutron scattering from the small subunit of *E. coli* ribosomes is shown in Fig. 9 (Capel *et al.*, 1988). The model compares well with the more recent crystallographic model of the small ribosomal subunit at atomic resolution (Wimberly *et al.*, 2000).

A more efficient way of contrast variation by proton spin polarization (Fig. 1) turned out to be the more appropriate choice for the study of the large ribosomal subunit with its 34 proteins and of the even larger functional complex of the ribosome. As there is a strong variation of the intensity of incoherent neutron scattering from the protons with nuclear polarization, the presence of protons in the sample has been restricted to a minimum. Except for the region of interest (= label), the ribosomal particle and the solvent were perdeuterated.

The analysis of the neutron scattering data in terms of (6) starts from the amplitude $U(\mathbf{Q})$ of the particle shape known from electron microscopy. The amplitude $V(\mathbf{Q})$ of the label is determined by a non-linear optimization procedure that minimizes the R factor between the calculated basic scattering functions and the experimental curves. The result is a low-resolution model of the *in situ* structure of the label and of its coordinates with respect to the ribosomal shape. A number of proteins of the large ribosomal subunit have been studied in this way (Fig. 10). Many of the protein positions agree with
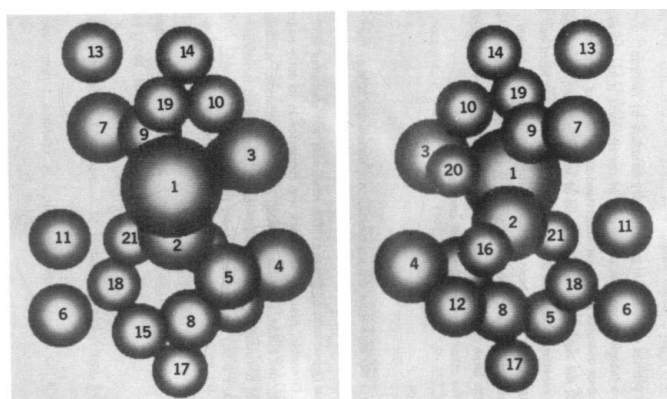
Figure 9
Two views of the neutron map of the 30S subunit of *E. coli* ribosomes. Each protein is represented by a sphere whose volume is the same as that of the protein. The maximum linear dimension of the array is about 190 Å. The two views are related to each other by a 180° rotation about an axis oriented vertically in the plane of the figure. From Capel *et al.* (1988). *Makromol. Chem. Macromol. Symp.* **15**, 123–130. Copyright Wiley-VCH Verlag GmbH & Co. KGaA. Reproduced with permission.
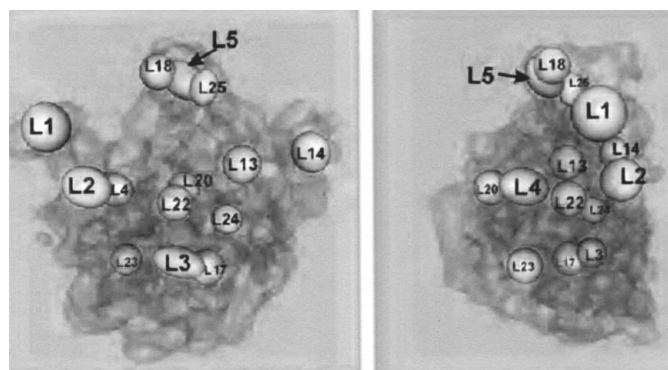
Figure 10
The positions of ribosomal proteins of the large subunit of *E. coli* ribosomes determined by nuclear spin contrast variation. The ribosome structure (shadowed) is courtesy of J. Frank. The figure on the right side is rotated by 90°. Reprinted from Willumeit *et al.* (2001). *Biochim. Biophys. Acta*, **1520**, 7–20. Copyright (2001), with permission from Elsevier.

those of the crystallographic model from Ban *et al.* (1998) but some are at variance (Willumeit *et al.*, 2001).

There is another method aiming at the *in situ* structure of regions inside a larger particle using selective gradual deuteration, *i.e.* 'triple isotopic substitution' (TIS) (Serdyuk & Pavlov, 1988). The gist of the TIS method is that it eliminates any interference of the amplitude of the label with that of the host particle at any concentration of the solute, a property which it shares with the triangulation method (Serdyuk & Zaccai, 1996).

6. Nuclear spin contrast

The history of nuclear spin contrast, and in particular of proton spin contrast, started with polarized neutron diffraction from dynamically polarized proton spins of a crystal (Hayter *et al.*, 1974; Leslie *et al.*, 1980). After an interruption of several years, nuclear spin contrast variation became a new technique of neutron small-angle scattering and now there are reasons to believe that it will be appreciated by crystallographers again in the not too distant future. What were and what still are the driving forces of this development?

Most importantly, there is a huge change of the scattering length of polarized coherent neutron scattering from polarized protons (Fig. 1). Incoherent scattering vanishes when the direction of proton polarization coincides with the polarization direction of the neutron beam.

High nuclear polarization in insulators – most hydrogenous substances belong to this class – is achieved by the method of dynamic nuclear spin polarization (Abragam & Goldman, 1978). A small amount of paramagnetic centres is added to the sample. In a moderately strong magnetic field of 2.5 T at temperatures $T \leq 1$ K, microwave irradiation will polarize the nuclear spins. The direction of nuclear spin polarization P depends on the choice of the microwave frequency. Slightly below the electron paramagnetic resonance (EPR), P will be positive, and, slightly above the EPR, P will be negative with respect to the direction of the external magnetic field. The polarization of the polarized nuclei is measured by NMR. NMR techniques are also used to selectively change the polarization of different isotopes. As already pointed out by Hayter *et al.* (1974), a possible discrimination between atoms of the same isotope differing in their Larmor frequency due to their distance from a paramagnetic centre by NMR techniques would open new prospects in structural studies.

It is for this reason that dynamically polarized protons of a crystal of lanthanum magnesium nitrate (LMN) $La_2Mg_3(NO_3)_{12} \cdot 26H_2O$ doped with $^{142}Nd^{3+}$ were studied by polarized neutron diffraction (Hayter *et al.*, 1974). The positions of the H atoms had been obtained from a limited set of reflections. The non-uniform proton polarization was one of the reasons for the unexpectedly poor quality of the polarized proton-density map (Leslie *et al.*, 1980). Later, John Hayter advised the author of this paper not to resume this line of research but rather to use nuclear polarization by brute force, *i.e.* at thermal equilibrium.

In the mid 80's, several groups embarked on experiments of polarized SANS from dynamically polarized proton spins in crown ethers (Kohgi *et al.*, 1987), polymers (Glättli *et al.*, 1989) and biological macromolecules (Knop *et al.*, 1986, 1992). All of them took advantage of a new type of polarized target material which had been optimized for high-energy-physics experiments. Its preparation for the purpose of neutron scattering from hydrophilic macromolecules is simple. A small amount of an organic radical, *e.g.* the sodium salt of bis(2-hydroxy-2-ethylbutyrato)-oxochromate (abbreviated as EHBA-Cr^V) (Krumpolc & Rožek, 1979) together with the biological macromolecule of interest is added to a glycerol/water mixture and rapidly frozen to a glassy platelet in a liquid-nitrogen-cooled copper mould. The data shown in Fig. 10 have been obtained from samples that had been prepared in this way.

6.1. Dynamic nuclear spin contrast

In a simple microscopic picture of dynamic nuclear polarization (DNP), the nuclear polarization develops near the paramagnetic centre through electron nuclear dipolar interaction decreasing with the third power of the distance between the electron and nuclear moments (solid effect). More distant bulk nuclei are polarized by dipolar interactions between nuclei (spin diffusion). Would there be a way to use the directly polarized close protons as a label in macromolecular structure research?

The formation of selectively polarized proton spin domains can best be observed at the onset of DNP. Time-resolved polarized neutron scattering from a solution of EHBA-Cr^V in a deuterated solvent showed that protons of the solute were polarized much more rapidly than those of the deuterated solvent (van den Brandt *et al.*, 2002) (Fig. 11). Clearly, the low concentration of protons in the deuterated solvent reduces
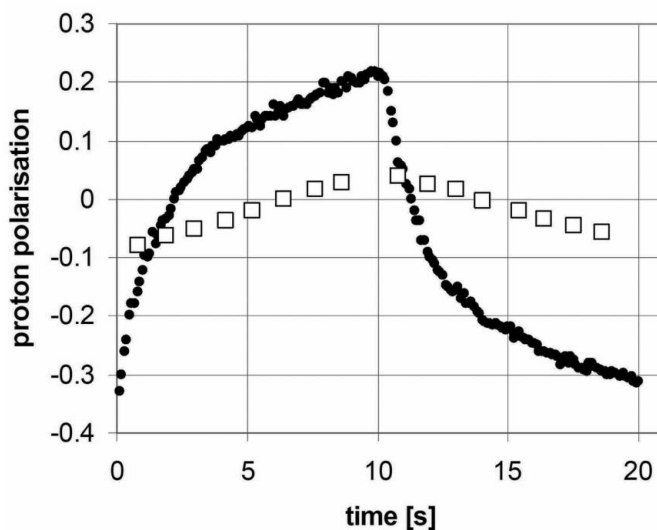
Figure 11
Polarization of close protons (H atoms of EHBA-Cr^V): circles from neutron scattering and polarization of the bulk protons (residual H atoms of deuterated solvent molecules); squares from NMR. The direction of DNP has been changed each 10 s (11 s) (van den Brandt *et al.*, 2002).

considerably proton-spin diffusion into the bulk. But, lowering the deuteration of the solvent stepwise to 80% led to a surprisingly modest decrease of the proton polarization gradient at the surface of the EHBA-Cr^V molecule (van den Brandt *et al.*, 2006). Unfortunately, a further decrease of deuteration of the solvent renders the measurement of the coherent scattering intensity on an increasing background of incoherent scattering very difficult. Time-resolved neutron scattering from radicals of larger size, which are presently analysed, seems to be more appropriate for the elucidation of the size and life-time of polarized proton domains in the absence of deuteration. From studies on tyrosine radicals of catalase, it appears that about four protons per radical are polarized in 5 s (Stuhrmann, 2004).

There is another method that may turn out to be useful for the characterization of polarized proton spin domains. The probability W of dynamic nuclear polarization by the solid effect (well resolved or not) (Abragam & Goldman, 1978) varies with the angle θ (Fig. 12).

$$W_{\text{solid effect}}(\mathbf{r}) \propto \left(\frac{\sin\theta\cos\theta}{r^3} \right)^2. \tag{12}$$

From (12), it appears that those nuclear spins near a paramagnetic centre that happen to be close to an angle of $\theta = 45°$ with respect to the external magnetic field are more susceptible to DNP (Fig. 12). The anisotropy of the intensity distribution increases with the angle ε between the direction of the magnetic field at the sample and the neutron beam. Fig. 12 shows the anisotropic part of the scattering pattern at $\varepsilon = 90°$. Data of time-resolved neutron scattering from EHBA-Cr^V taken with $\varepsilon = 7°$ confirm the very small predicted anisotropy of less than 0.001 (Stuhrmann, 2007). Experiments at larger ε, *e.g.* at $\varepsilon = 40°$, which would be perfectly feasible with the polarized target facility of the Paul-Scherrer Institut (PSI), would increase the effect by one to two orders of magnitude.

7. Outlook

Contrast variation is an expanding field of research, and there is continuing cross-fertilization between SAS and crystallography. To give an example, there are reasons to believe that further progress in the field of polarized neutron scattering from dynamically polarized nuclei will be facilitated by moving from SAS to crystallography. The most evident reason, at least for neutron protein crystallography, is the suppression of incoherent scattering from H atoms by high proton polarization.

The build up of polarization domains in space and time can be studied in more detail in single crystals of macromolecules, also because the polarization-dependent intensity is relatively stronger in crystal diffraction than in solution scattering. Two main lines may be envisaged: (i) the determination of a spin-density map in dilute paramagnets, *e.g.* of radicals in polymers or biological macromolecules, at medium resolution, and (ii) the distinction between protons with different Larmor

frequencies near chemically interesting radical sites at high resolution.

Another example is anomalous X-ray diffraction from native labels like sulfur and phosphorus in protein crystallography which is still in its infancy (Djinović Carugo *et al.*, 2005). As the concentration of phosphorus in RNA/DNA is quite high, its anomalous contrast in X-ray scattering reaches half of that which could be obtained by substitution of ^{1}H by ^{2}H in neutron scattering. The foreseeable role of this
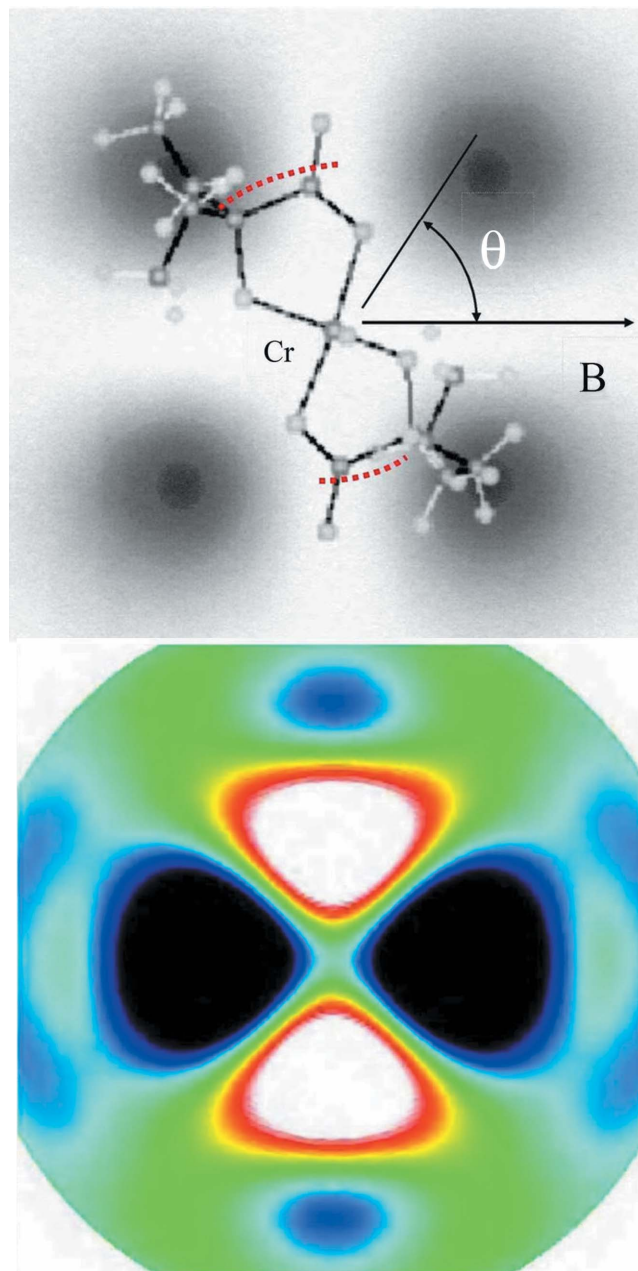
Figure 12
Top: The orientation of the EHBA-Cr^V molecule is favourable for dynamic polarization of its protons. The calculated anisotropic distribution of the intensity of polarized neutrons scattered by dynamically polarized protons of EHBA-Cr^V is shown below ($\varepsilon = 90°$). $Q_{\max} = 0.8$ Å^{-1} (Stuhrmann, 2007).

technique is similar to that of neutron protein crystallography. Soft X-ray diffraction will be used in special cases, *e.g.* anomalous dispersion of phosphorus for the study of lipid–protein interaction (Biou *et al.*, 2005), or for the RNA/DNA interaction with proteins in viruses and phages.

While high-performance neutron diffractometers for protein crystallography exist at all larger neutron scattering facilities, present technical developments for soft X-ray diffraction with synchrotron radiation will have to go on to meet the standards of modern protein crystallography (Biou *et al.*, 2005).

References

Abragam, A. & Goldman, M. (1978). *Rep. Prog. Phys.* **41**, 395–467.
Ban, N., Freeborn, B., Nissen, P., Penczek, P., Grassuci, R. A., Sweet, R., Frank, J., Moore, P. B. & Steitz, T. A. (1998). *Cell*, **93**, 1105–1115.
Bennet, W. S. Jr & Steitz, T. A. (1980). *J. Mol. Biol.* **140**, 183–209.
Bentley, G. A., Finch, J. T. & Lewit-Bentley, A. (1981). *J. Mol. Biol.* **145**, 771–784.
Bentley, G. A., Lewit-Bentley, A., Liljas, L., Skoglund, U., Roth, M. & Unge, T. (1987). *J. Mol. Biol.* **194**, 129–141.
Biou, V., Bösecke, P., Bois, J.-M., Brandolin, G., Kahn, R., Mas, C., Nauton, L., Nury, H., Pebay-Peyroula, E., Vicat, J. & Stuhrmann, H. (2005). *J. Synchrotron Rad.* **12**, 402–409.
Bragg, W. L. & Perutz, M. F. (1952). *Proc. R Soc. London Ser. A*, **213**, 425–435.
Brandt, B. van den, Glättli, H., Grillo, I., Hautle, P., Jouve, H., Kohlbrecher, J., Konter, J. A., Leymarie, E., Mango, S., May, R. P., Michels, A., Stuhrmann, H. B. & Zimmer, O. (2006). *Eur. Phys. J.* B**49**, 157–165.
Brandt, B. van den, Glättli, H., Grillo, I., Hautle, P., Jouve, H., Kohlbrecher, J., Konter, J. A., Leymarie, E., Mango, S., May, R. P., Stuhrmann, H. B. & Zimmer, O. (2002). *Europhys. Lett.* **59**, 62–67.
Bricogne, G. (1993). *Acta Cryst.* D**49**, 37–60.
Capel, M. S., Engelman, D. M., Freeborn, B. R., Kjeldgard, M., Langer, J. A., Ramakrishnan, V., Schindler, D. G., Schneider, D. K., Schoenborn, B. P., Sillers, I.-Y., Yabuki, S. & Moore, P. B. (1988). *Makromol. Chem. Macromol. Symp.* **15**, 123–130.
Chacon, P., Moran, F., Diaz, J. F., Pantos, E. & Andreu, J. M. (1998). *Biophys. J.* **74**, 2760–2775.
Crick, F. H. C. & Magdoff, B. S. (1956). *Acta Cryst.* **9**, 901–908.
Davey, C. A., Sargent, D. F., Luger, K., Maeder, A. W. & Richmond, T. J. (2002). *J. Mol. Biol.* **319**, 1097–1113.
Davies, J. M., Tsuruta, H., May, A. P. & Weis, W. I. (2005). *Structure (Camb.)*, **13**, 183–195.
Djinović Carugo, K., Helliwell, J. R., Stuhrmann, H. & Weiss, M. S. (2005). *J. Synchroton Rad.* **12**, 410–419.
Engelman, D. M. & Moore, P. B. (1972). *Proc. Natl Acad. Sci. USA*, **69**, 1997–1999.
Finch, J. T. & Holmes, K. C. (1967). *Methods in Virology*, edited by K. Maramorosch & H. Koprowski, Vol. 3, pp. 351–474. New York: Academic Press.
Fourme, R., Shepard, W., Kahn, R., l'Hermite, G. & Li de La Sierra, I. (1995). *J. Synchrotron Rad.* **2**, 36–48.

Glättli, H., Fermon, C. & Eisenkremer, M. (1989). *J. Phys.* **50**, 2375–2388.
Goerigk, G., Schweins, R., Huber, K. & Ballauf, M. (2004). *Europhys. Lett.* **66**, 331–337.
Hayter, J. B., Jenkin, G. T. & White, J. W. (1974). *Phys. Rev. Lett.* **33**, 696–699.
Heller, W. T., Abusamhadneh, E., Finley, N., Rosevaer, V. R. & Trewhella, J. (2002). *Biochemistry*, **41**, 15654–15663.
Henderson, R. & Unwin, P. N. T. (1975). *Nature (London)*, **257**, 28–31.
Hendrickson, W. (1991). *Science*, **254**, 51–58.
Hjelm, R. P., Kneale, G. G., Suau, P., Baldwin, J. P., Bradbury, E. M. & Ibel, K. (1977). *Cell*, **10**, 139–151.
Hoppe, W. (1972). *J. Mol. Biol.* **78**, 581–585.
Hough, M. A., Grossmann, J. G., Antonyuk, S. V., Strange, R. W., Doucette, P. A., Rodriguez, J. A., Whitson, L. J., Hart, P. J., Hayward, L. J., Valentine, J. S. & Hasnain, S. S. (2004). *Proc. Natl Acad. Sci. USA*, **101**, 5976–5981.
Ibel, K. & Stuhrmann, H. B. (1975). *J. Mol. Biol.* **93**, 255–265.
Jones, T. A. & Liljas, R. (1984). *J. Mol. Biol.* **177**, 735–767.
Kendrew, J. C., Dickerson, R. E., Strandberg, B. E., Hart, R. G., Davies, D. R., Phillips, D. C. & Shore, V. C. (1960). *Nature (London)*, **185**, 422–427.
Kirste, R. G. & Stuhrmann, H. B. (1967). *Z. Phys. Chem. Neue Folge*, **56**, 338–341.
Knop, W., Hirai, M., Schink, H.-J., Stuhrmann, H. B., Wagner, R., Zhao, J., Schärpf, O., Crichton, R. R., Krumpolc, M., Nierhaus, K. H., Rijllart, A. & Niinikoski, T. O. (1992). *J. Appl. Cryst.* **25**, 155–165.
Knop, W., Nierhaus, K. H, Novotny, V., Niinikoski, O., Krumpolc, M., Rieubland, J.-M., Rijllart, A., Schärpf, O., Schink, H. J., Stuhrmann, H. B. & Wagner, R. (1986). *Helv. Phys. Acta*, **59**, 741–746.
Kohgi, M., Ishida, M., Ishikawa, Y., Ishimoto, S., Kanno, Y., Masaike, A., Masuda, Y. & Morimoto, K. (1987). *J. Phys. Soc. Jpn*, **56**, 2681–2688.
König, S., Svergun, D. I., Volkov, V. V., Feigin, L. A. & Koch, M. H. J. (1998). *Biochemistry*, **37**, 5329–5334.
Krumpolc, M. & Rožek, J. (1979). *J. Am. Chem. Soc.* **101**, 3206–3209.
Kühnholz, O. (1991). *J. Appl. Cryst.* **24**, 811–814.
Leslie, M., Jenkin, G. T., Hayter, J. B., White, J. W., Cox, S. & Warner, G. (1980). *Philos. Trans. R. Soc. London Ser. B*, **290**, 497–503.
Macheroux, P., Schönbrunn, E., Svergun, D. I., Volkov, V. V., Koch, M. H. J., Bornemann, S. & Thorneley, R. N. F. (1998). *Biochem. J.* **335**, 319–327.
O'Halloran, T. V., Lippard, S. J., Richmond, J. T. & Klug, A. (1987). *J. Mol. Biol.* **194**, 705–712.
Ramin, M., Shepard, W., Fourme, R. & Kahn, R. (1999). *Acta Cryst.* D**55**, 157–167.
Seiff, F., Wallat, I., Ermann, P. & Heyn, M. P. (1985). *Proc. Natl Acad. Sci. USA*, **82**, 3227–3231.
Serdyuk, I. N. & Grenader, A. K. (1974). *FEBS Lett.* **59**, 133–136.
Serdyuk, I. N. & Pavlov, M. Y. (1988). *Makromol. Chem. Macromol. Symp.* **15**, 167–184.
Serdyuk, I. N. & Zaccai, G. (1996). *J. Mol. Struct.* **383**, 197–200.
Shi, Y. Y., Hong, X. G. & Wang, C. C. (2005). *J. Biol. Chem.* **280**, 22761–22768.
Simon, J.-P. (2007). *J. Appl. Cryst.* **40**, s1–s9.
Stöffler, G., Redl, B., Walleczek, J. & Stöffler-Meilicke, M. (1988). *Methods Enzymol.* **164**, 64–76.
Stuhrmann, H. B. (1970*a*). *Acta Cryst.* A**26**, 297–306.
Stuhrmann, H. B. (1970*b*). *Z. Phys. Chem. (Frankfurt am Main)*, **72**, 177–184.
Stuhrmann, H. B. (1970*c*). *Z. Phys. Chem. (Frankfurt am Main)*, **72**, 185–198.
Stuhrmann, H. B. (2004). *Rep. Prog. Phys.* **67**, 1073–1115.
Stuhrmann, H. B. (2007). *Acta Cryst.* A**63**, 455–464.
Stuhrmann, H. B. & Kirste, R. (1965). *Z. Phys. Chem. (Frankfurt am Main)*, **46**, 247–250.

Stuhrmann, H. B., Koch, M. H. J., Parfait, R., Haas, J., Ibel, K. & Crichton, R. R. (1978). *J. Mol. Biol.* **119**, 203–212.

Stuhrmann, S., Bartels, K. S., Braunwarth, W., Doose, R., Dauvergne, F., Gabriel, A., Knöchel, A., Marmotti, M., Stuhrmann, H. B., Trame, C. & Lehmann, M. S. (1997). *J. Synchrotron Rad.* **4**, 298–310.

Svergun, D. I. (1997). *J. Appl. Cryst.* **30**, 792–797.

Svergun, D. I. (1999). *Biophys. J.* **76**, 2879–2886.

Svergun, D. I. (2007). *J. Appl. Cryst.* **40**, s10–s17.

Svergun, D. I. & Stuhrmann, H. B. (1991). *Acta Cryst.* A**47**, 736–744.

Svergun, D. I., Volkov, V. V., Kozin, M. B. & Stuhrmann, H. B. (1996). *Acta Cryst.* A**52**, 419–426.

Svergun, D. I., Petoukhov, M. V. & Koch, M. H. J. (2001). *Biophys. J.* **80**, 2946–2953.

Vachette, P. & Svergun, D. (2000). *Structure and Dynamics of Biomolecules, HERCULES*, Vol. IV, edited by E. Fanchon, E. Geissler, J.-L. Hodeau, J.-R. Regnard & P. A. Timmins, pp. 199–237. Oxford University Press.

Volkov, V. V. & Svergun, D. I. (2003). *J. Appl. Cryst.* **36**, 860–864.

Walther, D., Cohen, F. E. & Doniach, S. (2000). *J. Appl. Cryst.* **33**, 350–363.

Wang, J., Smerdon, S. J., Jaeger, J., Kohlstaedt, J. A., Friedman, J., Rice, P. A. & Steitz, T. A. (1994). *Proc. Natl Acad. Sci. USA*, **91**, 7242–7246.

Willumeit, R., Diedrich, G., Forthmann, S., Beckmann, J., Stuhrmann, H. B. & Nierhaus, K. H. (2001). *Biochim. Biophys. Acta*, **1520**, 7–20.

Wimberly, B. T., Brodersen, D. E., Clemons, W. M., Morgan-Warren, R. J., Carter, A. P., Vonrhein, C., Hartsch, Th. & Ramakrishnan, V. (2000). *Nature (London)*, **407,** 327–339.

Worcester, D. L. (1975). Brookhaven Symposium on Neutrons in Biology, Vol. III, pp. 37–57.

Zaccai, G. & Gilmore, D. J. (1979). *J. Mol. Biol.* **132**, 181–191.