electronic reprint

Journal of
Applied Crystallography
ISSN 0021-8898

CRYSOL – a Program to Evaluate X-ray Solution Scattering of Biological Macromolecules from Atomic Coordinates

D. Svergun, C. Barberato and M. H. J. Koch

J. Appl. Cryst. (1995). **28**, 768–773

J. Appl. Cryst. (1995). **28**, 768–773

CRYSOL – a Program to Evaluate X-ray Solution Scattering of Biological Macromolecules from Atomic Coordinates

BY D. SVERGUN,* C. BARBERATO AND M. H. J. KOCH

European Molecular Biology Laboratory, Hamburg Outstation, Notkestrasse 85, *D*-22603 *Hamburg, Germany*

(*Received* 7 *March* 1995; *accepted* 24 *May* 1995)

Abstract

A program for evaluating the solution scattering from macromolecules with known atomic structure is presented. The program uses multipole expansion for fast calculation of the spherically averaged scattering pattern and takes into account the hydration shell. Given the atomic coordinates (*e.g.* from the Brookhaven Protein Data Bank) it can either predict the solution scattering curve or fit the experimental scattering curve using only two free parameters, the average displaced solvent volume per atomic group and the contrast of the hydration layer. The program runs on IBM PCs and on the major UNIX platforms.

Introduction

Comparisons between experimental X-ray solution scattering [small-angle X-ray scattering (SAXS)] curves and those evaluated from crystallographic structures have been widely used to validate theoretical models, to verify the structural similarity between proteins and nucleic acids in crystals and in solution and to predict quaternary structures (see *e.g.* Langridge *et al.*, 1960; Ninio, Luzzati & Yaniv, 1972; Fedorov, Ptitsyn & Voronin, 1972; Fedorov & Denesyuk, 1978; Müller, 1983; Pavlov, Sinev, Timchenko & Ptitsyn, 1986; Grossmann *et al.*, 1993). Moreover, for multisubunit macromolecules, accurate evaluation of the solution scattering amplitudes from the atomic coordinates of separate domains allows one to determine their relative positions from the SAXS data (Svergun, 1991, 1994).

The main problem in evaluating the solution scattering from atomic coordinates is to adequately take into account the solvent scattering. Several methods have been developed that basically differ in the representation of the particle volume inaccessible to the solvent. In the effective-atomic-scattering-factor method (*e.g.* Langridge *et al.*, 1960; Fraser, MacRae & Suzuki, 1978; Lattman, 1989), the excluded volume is built by dummy solvent atoms placed at the positions of the atoms in the macromolecule. This approach is well justified at resolutions down to 1–2 nm [i.e. in the range of momentum transfer $0 \lesssim s \lesssim 3$ nm^{-1}, $s = (4\pi/\lambda) \sin\theta$, 2θ is the scattering angle, λ the wavelength]. At higher resolution, the inhomogeneous filling of the excluded volume may introduce systematic deviations. The cube method (Fedorov, Ptitsyn & Voronin, 1972; Ninio, Luzzati & Yaniv, 1972) and its modifications (Müller, 1983; Pavlov & Fedorov, 1983) homogeneously fills this volume with cubic elements and thus provides better accuracy at higher scattering vectors ($s \gtrsim 3$ nm^{-1}).

The above methods do not take into account the hydration shell surrounding macromolecules in solution. Omission of this shell can lead to significant systematic errors even near the origin of the scattering curves and therefore to misinterpretation of the results. Attempts to include the hydration shell have been made, *e.g.*, by Hubbard, Hodgson & Doniach (1988) and Grossmann *et al.* (1993), but no general-purpose program has been developed. *CRYSOL*, the program described below, evaluates the SAXS profiles from crystallographic structures taking into account the scattering from the hydration shell.

Theory

Macromolecules in solution can be schematically represented as illustrated in Fig. 1. The particle with scattering density $\rho_a(\mathbf{r})$ is surrounded by a solvent with an average scattering density ρ_0. The hydration shell is approximated by a border layer of effective thickness Δ and density ρ_b that may differ from ρ_0. The SAXS intensity from such particles in dilute solution is proportional to the averaged scattering of a single particle:

$$I(s) = \langle |A_a(\mathbf{s}) - \rho_0 A_c(\mathbf{s}) + \delta\rho A_b(\mathbf{s})|^2 \rangle_\Omega, \qquad (1)$$

where $A_a(\mathbf{s})$ is the scattering amplitude from the particle *in vacuo*, $A_c(\mathbf{s})$ and $A_b(\mathbf{s})$ are, respectively, the scattering amplitudes from the excluded volume and the border layer, both with unitary density, $\delta\rho = \rho_b - \rho_0$, and $\langle\ \rangle_\Omega$ stands for the average over all particle orientations [Ω is the solid angle in reciprocal space, $\mathbf{s} = (s, \Omega)$].

* On leave from the Institute of Crystallography, Russian Academy of Sciences, Leninsky pr. 59, 117333 Moscow, Russia.

ISSN 0021-8898

Spherical averaging in (1) is greatly facilitated by use of the multipole expansion (Stuhrmann, 1970*a*; Lattman, 1989). For the atomic coordinates $\mathbf{r}_j = (r_j, \omega_j) = (r_j, \theta_j, \varphi_j)$ and the corresponding atomic form factors $f_j(s)$, the scattering amplitude *in vacuo* of a particle consisting of N atoms is

$$A_a(\mathbf{s}) = \sum_{j=1}^{N} f_j(s) \exp(i\mathbf{s}\mathbf{r}_j). \quad (2)$$

Substituting the relation (Edmonds, 1957)

$$\exp(i\mathbf{s}\mathbf{r}) = 4\pi \sum_{l=0}^{\infty} \sum_{m=-l}^{l} i^l j_l(sr) Y^*_{lm}(\omega) Y_{lm}(\Omega), \quad (3)$$

where the $j_l(sr)$ are the spherical Bessel functions and the $Y_{lm}(\Omega)$ are the spherical harmonics, into (2), one can write

$$A_a(\mathbf{s}) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} A_{lm}(s) Y_{lm}(\Omega), \quad (4)$$

where $A_{lm}(s)$ are the partial amplitudes:

$$A_{lm}(s) = 4\pi i^l \sum_{j=1}^{N} f_j(s)\, j_l(sr_j) Y^*_{lm}(\omega_j). \quad (5)$$

With the excluded volume represented as a superposition of dummy atoms with form factors $g_j(s)$ centered at the same coordinates $\mathbf{r}_j$, the amplitude $A_c(\mathbf{s})$ is expressed in the form of (4) with the partial amplitudes

$$C_{lm}(s) = 4\pi i^l \sum_{j=1}^{N} g_j(s)\, j_l(sr_j) Y^*_{lm}(\omega_j). \quad (6)$$

Following Stuhrmann (1970*b*), the border layer can be described by a two-dimensional angular function $F(\omega)$ (Fig. 1) as

$$\rho_b(\mathbf{r}) = \begin{cases} 1 & F(\omega) \le r \le F(\omega) + \Delta \\ 0 & 0 < r < F(\omega) \text{ or } r > F(\omega) + \Delta. \end{cases} \quad (7)$$

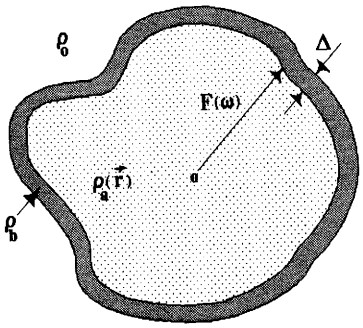

Fig. 1. Schematic representation of a macromolecule in solution. For explanations see text.

As the partial amplitudes are the Hankel transforms of the real-space radial functions:

$$B_{lm}(s) = i^l (2/\pi)^{1/2} \int_0^{\infty} \rho_{lm}(r) j_l(sr) r^2\, dr, \quad (8)$$

where

$$\rho_{lm}(r) = \int_{\omega} \rho_b(\mathbf{r}) Y^*_{lm}(\omega)\, d\omega, \quad (9)$$

it is readily shown that

$$B_{lm}(s) = i^l (2/\pi)^{1/2} \int_{\omega} Y^*_{lm}(\omega)\, d\omega \int_{F(\omega)}^{F(\omega)+\Delta} j_l(sr) r^2\, dr. \quad (10)$$

By (4)–(6) and (10), the three amplitudes entering into (1) are represented *via* their multipole components. Owing to the orthogonal properties of the spherical harmonics, all cross terms cancel out in the average over Ω, leading to a simple expression for the SAXS intensity:

$$I(s) = \sum_{l=0}^{L} \sum_{m=-l}^{l} |A_{lm}(s) - \rho_0 C_{lm}(s) + \delta\rho B_{lm}(s)|^2, \quad (11)$$

where the truncation value L defines the resolution of the representation of the particle.

Evaluation of the partial amplitudes

The form factors $f_j(s)$ and $g_j(s)$ necessary for the evaluation of the partial amplitudes $A_{lm}(s)$ and $C_{lm}(s)$ are specified as follows. Since usually only the coordinates of the non-H atoms are available in the Protein Data Bank (Bernstein *et al.*, 1977), the summation in (2) runs only over these atoms, but those having covalently bound hydrogens are replaced by the corresponding atomic groups according to Table 1. A similar approach was used earlier by Pavlov & Fedorov (1983) and Lattman (1989). The spherically averaged form factors $f_i(s)$ of each atomic group are tabulated on the interval $0 \le s \le 10$ nm^{-1} using a five-Gaussian approximation of the form factors of the individual atoms (*International Tables for X-ray Crystallography*, 1974) and the interatomic distances.

The form factors of the dummy atoms entering (6) are expressed as (Fraser, MacRae & Suzuki, 1978)

$$g_j(s) = G(s) V_j \exp(-\pi s^2 V_j^{2/3}), \quad (12)$$

where $V_j = (4\pi/3)\, r_{wj}^3$ is the solvent volume displaced by the jth atom, or atomic group, represented by the Gaussian sphere of radius r_{gj} (see Table 1), and $G(s)$ is an overall expansion factor

$$G(s) = (r_0/r_m)^3 \exp[-(4\pi/3)^{3/2} \pi s^2 (r_0^2 - r_m^2)]. \quad (13)$$

Here, $r_m = N^{-1} \sum_{j=1}^{N} r_{gj}$ is the actual average radius of the atomic group (for proteins, r_m is normally around

Table 1. *Parameters of the atomic groups and heteroatoms*

Atom or atomic group	Displaced solvent volume (nm^3)	Radius (nm)
H*	0.00515	0.107
C*	0.01644	0.158
CH†	0.02159	0.173
CH_2†	0.02674	0.185
CH_3†	0.03189	0.197
N*	0.00249	0.084
NH†	0.00764	0.122
NH_2†	0.01279	0.145
NH_3	0.01794	0.162
O*	0.00913	0.130
OH†	0.01428	0.150
S‡	0.01986	0.168
SH†	0.02510	0.181
Mg‡	0.01716	0.160
P‡	0.00573	0.111
Ca‡	0.03189	0.197
Mn‡	0.00920	0.130
Fe‡	0.00799	0.124
Cu‡	0.00878	0.128
Zn‡	0.00985	0.133

* Observed displaced volumes according to Fraser, MacRae & Suzuki (1978).

† Evaluated by adding the displaced volume of corresponding hydrogens.

‡ Radii taken from *International Tables for X-ray Crystallography* (1968).

0.162 nm), and r_0, the effective atomic radius, is a variable parameter that can be used to change the displaced volume per atomic group and thus to adjust the total excluded volume. Note that the expansion factor (13) does not depend on the atomic positions and can be taken out of the summation in (6).

To evaluate the envelope function $F(\omega)$, the particle is moved so that its geometrical center coincides with the origin. A quasiuniform grid of N_g angular directions is evaluated using Fibonacci numbers as described by Svergun (1994). Each non-H atom in the macromolecule updates the envelope function in the direction ω_i if the minimum distance between the atom and this direction is less than the sum of the atomic radius r_{gj} and the radius of the water molecule, $r_w = 0.15$ nm. The updated value is

$$F(\omega_i) = \max\{F'(\omega_i), (r_{ji} + 0.5r_{gj})\}, \qquad (14)$$

where $F'(\omega_i)$ is the current value of the envelope function and r_{ji} is the projection of $\mathbf{r}_j$ onto the direction ω_i (Fig. 2).

After all atoms have been sorted, $F(\omega)$ contains the distances between the origin and the particle surface for each ω_i. The amplitudes $C_{lm}(s)$ are then evaluated by numerical integration of (10). The integral over r does not depend on the envelope function and can be tabulated in advance.

The density of the bound solvent can differ noticeably from that of the bulk solvent within only a few ångströms distance from the surface (see *e.g.* Cheng & Schoenborn, 1990; Badger, 1993), *i.e.* the condition $F(\omega) \gg \Delta$ is usually fulfilled. This means that the contribution from the border layer depends mostly on the product $\delta\rho\Delta$ and without loss of generality one of the two parameters can be fixed. In *CRYSOL*, the effective width of the border layer is taken to be 0.3 nm to simulate the most ordered first hydration layer. The SAXS intensity (11) depends on two parameters: (i) the average displaced volume per atomic group and (ii) the contrast of the border layer. The former parameter is expressed *via* the effective atomic radius, which should not differ much from r_m (we found $0.96r_m \le r_0 \le 1.04r_m$). The upper limit of the latter is $(\delta\rho)_{max} \simeq 70$ e nm^{-3} and corresponds to the difference in the electron density between free and bound water molecules (Perkins, 1986).

Program implementation

The above algorithms are implemented in the interactive Fortran77 program *CRYSOL*, which performs the following steps:

(1) The atomic coordinates are read from the data file in the Brookhaven Protein Data Bank (PDB) format into a temporary heap storage by blocks of a thousand to determine the geometrical center of the macromolecule and the origin is shifted to this point. The use of the heap storage allows one to avoid limitations on the number of atoms in the input file.

(2) A quasiuniform grid of angular directions ω_i, $i = 1 \ldots N_g$ is evaluated ($N_g \le 4185$), the form factors and the integrals of Bessel functions (10) are tabulated in the range of momentum transfer and at the resolution level ($L \le 15$) specified by the user.

(3) The atomic types and coordinates are read from the PDB file into a temporary heap storage once again. For each atom or heteroatom, the atomic group and the displaced volume are identified according to Table 1 and the contributions to the partial amplitudes $A_{lm}(s)$, $C_{lm}(s)$ and the envelope function $F(\omega)$ are evaluated. O atoms belonging to water molecules are ignored.

(4) The amplitudes $B_{lm}(s)$ are evaluated from the function $F(\omega)$.

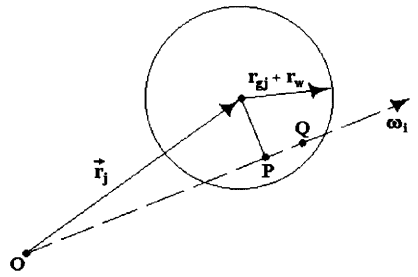

Fig. 2. Evaluation of the envelope function. The current atomic coordinate relative to the origin O is $\mathbf{r}_j$, the current direction ω_i, $|OP| = r_{ji}$, $|PQ| = 0.5r_{gj}$, $F(\omega_i) = |OQ|$. For details see text.

(5) The SAXS intensity, $I(s, r_0, \delta\rho)$ is calculated using (11) for a value of $\rho_0 = 334$ e nm^{-3} corresponding to the bulk water with the default parameters, $r_0 = r_m$ and $\delta\rho = 30$ e nm^{-3}.

(6) If the experimental curve $I_e(s)$ is given, the parameters are adjusted so as to fit the experimental data. A plain grid search is made for $0.96\, r_m \leq r_0 \leq 1.04\, r_m$ and $0 \leq \delta\rho \leq 60$ e nm^{-3} to minimize the functional

$$\chi^2(r_0, \delta\rho) = \frac{1}{N_p} \sum_{i=1}^{N_p} \left[\frac{I_e(s_i) - cI(s_i, r_0, \delta\rho)}{\sigma(s_i)} \right]^2, \quad (15)$$

where N_p is the number of experimental points, the $\sigma(s_i)$ are the experimental errors and

$$c = \left[\sum_{i=1}^{N_p} \frac{I_e(s_i) I(s_i, r_0, \delta\rho)}{\sigma(s_i)^2} \right] \left[\sum_{i=1}^{N_p} \frac{I(s_i, r_0, \delta\rho)^2}{\sigma(s_i)^2} \right]^{-1} \quad (16)$$

is the scale factor. The fit is presented on a graphic display and the parameters can also be changed manually.

(7) The results (integral parameters of the particle, its envelope function, partial amplitudes, intensities and the fit to the experimental data) are stored in the corresponding ASCII and binary files. The data can be retrieved for further calculations with other parameters and/or experimental data sets.

The program is compiled on IBM PC computers using the Microsoft Fortran PowerStation 1.0 with the Phar Lap MS-DOS extender and requires DOS version 3.3 or later, 2 Mbytes of free memory (conventional + extended) and EGA/VGA/SVGA video display. Versions for the major UNIX platforms (Sun, Silicon Graphics, DEC Alpha), which use the public domain graphical package *GNUPLOT*, are also available.

Examples of applications

To illustrate the use of *CRYSOL*, we present the results obtained for hen egg white lysozyme (molecular weight 14 KDa), which has already been used for illustrative purposes by various authors (Pickover & Engelman, 1982; Pavlov & Fedorov, 1983; Lattman, 1989). The X-ray scattering curve from a lysozyme solution (Fig. 3) was recorded using standard procedures on the X33 camera of the EMBL in HASYLAB at the Deutsches Electronen Synchrotron (DESY) in Hamburg. The coordinates of the lysozyme (Diamond, 1974) were taken from the PDB file pdb6lyz.ent. Fig. 4 displays the scattering curves $I_a(s) = \langle A_a(\mathbf{s}) \rangle^2$, $I_c(s) = \langle \rho_0 A_c(\mathbf{s}) \rangle^2$ and $I_b(s) = \langle \delta\rho A_b(\mathbf{s}) \rangle^2$ evaluated with $L = 12$ and $N_g = 2585$. In Fig. 3, the best fit to the experimental data ($\chi = 0.477$) obtained at $r_0 = 0.161$ nm (total excluded volume 17.4 nm^3) and $\delta\rho = 25$ e nm^{-3} corresponding to a hydration of 0.4 g g^{-1} (gram of H_2O per gram of protein) is presented. The experimental radius of gyration is $R_g = 1.52$ (2) nm; the theoretical value is $R_{gt} = 1.48$ nm. Note that when the hydration shell (its radius of gyration is 1.88 nm) is not taken into account by fixing of $\delta\rho = 0$, the fit to the experimental data is poorer ($\chi = 0.765$, $R_{gt} = 1.43$ nm).

The results of *CRYSOL* were compared to those of the program of Pavlov & Fedorov (1983), which uses the modified cube method and the numerical average in reciprocal space. The excluded volume 16.8 nm^3

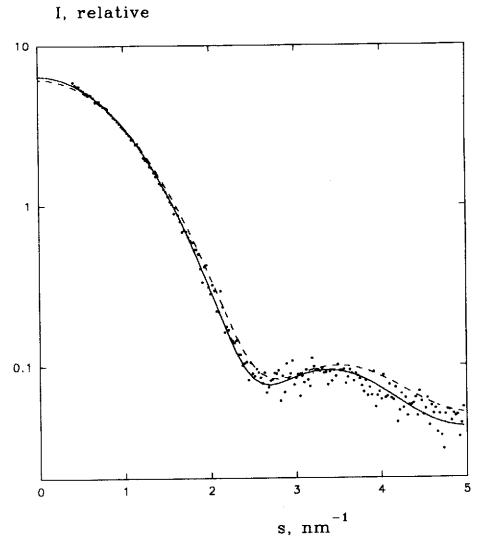

Fig. 3. Experimental solution scattering from lysozyme (dots) and the best fit by *CRYSOL* (solid line). The dashed curve indicates the best fit at $\delta\rho = 0$.

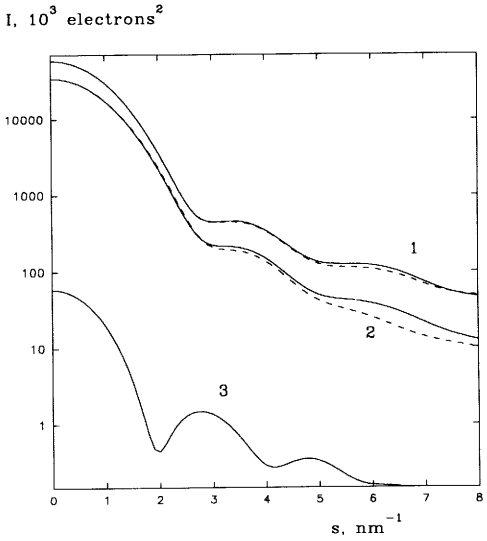

Fig. 4. Scattering from the atomic structure *in vacuo* $I_a(s)$ (1), shape scattering $I_s(s)$ (2) and scattering from the border layer $I_b(s)$ for lysozyme (solid lines). Dashed lines: curves evaluated by Pavlov's program with the shape scattering scaled to the same excluded volume.

corresponds well to the value obtained by *CRYSOL* and the curves $I_a(s)$ and $I_c(s)$ are in good agreement up to $s \simeq 4$ nm^{-1}. For higher angles, there are deviations in the shape scattering from *CRYSOL* owing to the inhomogeneously filled excluded volume. The scattering curve calculated by Pavlov's program, which does not take the hydration shell into account, has noticeable systematic deviations at small angles ($\chi = 0.687$, $R_{gt} = 1.47$ nm). The deviations can be reduced by artificial changing of the solvent density to $\rho_0 = 310$ e nm^3, which gives $\chi = 0.581$ and $R_{gt} = 1.45$ nm. The total CPU time required by *CRYSOL* on an IBM AT/486 DX50 was 310 s. In comparison, Pavlov's program, for which the user has to run three separate executable modules, requires a total of 660 s of CPU time.

We have also attempted to make a comparison with the program of Lattman (1989) which uses the effective atomic factors method and the multipole expansion. This comparison failed, apparently owing to software limitations encountered in running this program with our experimental data.

Fig. 5 illustrates the use of CRYSOL for the *E. coli* aspartate transcarbamylase (ATCase), which is a dodecamer with a molecular weight of 303 KDa. The coordinates of the *T* form of the ATCase (Kantorowitz & Lipscomb, 1988) were generated using the appropriate symmetry operations from the PDB file pdb6at1.ent. The experimental curve recorded at the small-angle scattering installation of the synchrotron-radiation laboratory LURE in Orsay, France (Herve *et al.*, 1985) yields $R_g = 4.68$ (3) nm. *CRYSOL* provides the best fit ($\chi = 1.16$, $R_{gt} = 4.64$ nm) at $r_0 = 0.168$ nm and $\delta\rho = 58$ e nm^{-3} corresponding to a hydration of 0.21 g g^{-1}. The best fit achieved without the hydration shell at $r_0 = 0.167$ is poor ($\chi = 4.89$, $R_{gt} = 4.43$ nm) and displays a significant shift of the subsidiary maxima (such a shift has already been reported by Altman, Ladner & Lipscomb, 1982).

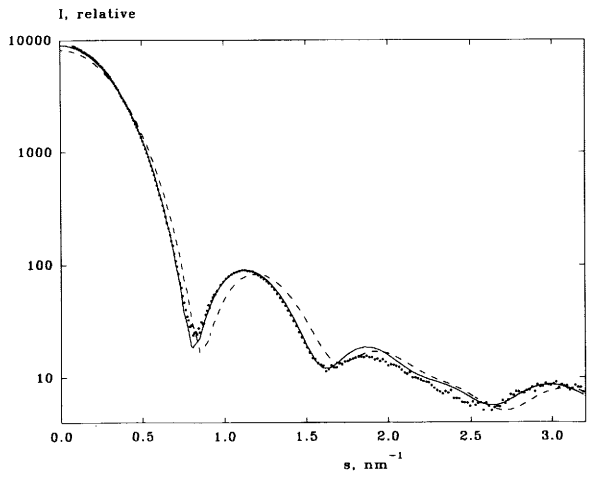

Fig. 5. Experimental solution scattering from ATCase and the fits with and without solvation shell. Notation is as in Fig. 3.

CRYSOL has been successfully used on a number of protein structures in ongoing projects at the EMBL Outstation (*e.g.* hexokinase, ribonucleotide reductase proteins *R*1 and *R*2 *etc.*). A beta-release of the program was also tested at the Stanford Synchrotron Radiation Laboratory (Stanford University, USA).

Concluding remarks

The importance of the contribution of the hydration shell to the scattering has been discussed by various authors (Hubbard, Hodgson & Doniach, 1988; Schoenborn, 1988; Badger, 1993; Grossmann *et al.*, 1993). The structures available from the Protein Data Bank usually contain less than 50% of the bound waters and these can hardly be used to represent the hydration shell in solution. In fact, attempts to include these waters in the calculations on lysozyme with Pavlov's program degraded the results.

The border layer introduced in *CRYSOL* is, of course, a simplified model of the hydration shell. For macromolecules with a complicated shape, the envelope function may not be single valued and the use of $F(\omega)$ would fill the inner cavities. Although $F(\omega)$ is thus not suitable to evaluate the shape scattering itself, one is still well justified to use it for the description of the outer hydration shell. By use of the shell of a constant density and the fixed thickness of 0.3 nm, the primary solvation layer is approximated. The primary layer is known to contain the most ordered waters (see, *e.g.*, Thanki, Thornton & Goodfellow, 1988; Cheng & Schoenborn, 1990; Badger, 1993) and thus dominates the scattering from the solvation shell. For all proteins studied up to now, we found that the contribution from the border layer significantly improved the fit to the experimental data (the hydration ratio was normally 0.2–0.3 g g^{-1}; the value 0.4 g g^{-1} for lysozyme was exceptionally high).

CRYSOL has been proven to adequately evaluate the SAXS profiles up to $s \lesssim 4$ nm^{-1} (*i.e.* a resolution of about 1.5 nm), where the deviations due to the inhomogeneously filled excluded volume and the finite number of multipoles are negligible. At higher angles, the cube methods are expected to be more accurate.

The executable code of the program for IBM PCs and UNIX workstations and a user manual are available from the authors (e-mail svergun@embl-Hamburg.de).

The authors thank Drs M. Pavlov and E. Lattman for providing their programs, Dr A. Semenyuk for his help at the early stage of this project and Dr P. Vachette for providing the data on ATCase. This work was supported by the NATO Linkage Grant LG 921231, INTAS grant 93-645 and the CNPq (Conselho Nacional de Desenvolvimento Cientifico e Tecnologico) fellowship of C. Barberato.

References

ALTMAN, R. B., LANDER, J. E. & LIPSCOMB, W. N. (1982). *Biochem. Biophys. Res. Commun.* **108**, 592–595.

BADGER, J. (1993). *Biophys. J.* **65**, 1656–1659.

BERNSTEIN, F. C., KOETZLE, T. F., WILLIAMS, G. J. B., MEYER, E. F. JR, BRICE, M. D., RODGERS, J. R., KENNARD, O., SHIMANOUCHI, T. & TASUMI, M. (1977). *J. Mol. Biol.* **112**, 535–542.

CHENG, X. & SCHOENBORN, B. P. (1990). *Acta Cryst.* B**46**, 195–208.

DIAMOND, R. (1974). *J. Mol. Biol.* **82**, 371–391.

EDMONDS, A. R. (1957). *Angular Momentum in Quantum Mechanics*. Princeton Univ. Press.

FEDOROV, B. A. & DENESYUK, A. I. (1978). *J. Appl. Cryst.* **11**, 473–477.

FEDOROV, B. A., PTITSYN, O. B. & VORONIN, L. A. (1972). *FEBS Lett.* **28**, 188–190.

FRASER, R. D. B., MACRAE, T. P. & SUZUKI, E. (1978). *J. Appl. Cryst.* **11**, 693–694.

GROSSMANN, G., ABRAHAM, Z. H. L., ADMAN, E. T., NEU, M., EADY, R. R., SMITH, B. E. & HASNAIN, S. S. (1993). *Biochemistry*, **32**, 7360–7366.

HERVÉ, G., MOODY, M. F., TAUÇ, P., VACHETTE, P. & JONES, P. T. (1985). *J. Mol. Biol.* **185**, 189–199.

HUBBARD, S. R., HODGSON, K. O. & DONIACH, S. (1988). *J. Biol. Chem.* **263**, 4151–4158.

International Tables for X-ray Crystallography (1968). Vol. III. Birmingham: Kynoch Press. (Present distributor Kluwer Academic Publishers, Dordrecht.)

International Tables for X-ray Crystallography (1974). Vol. IV. Birmingham: Kynoch Press. (Present distributor Kluwer Academic Publishers, Dordrecht.)

KANTOROWITZ, E. R. & LIPSCOMB, W. N. (1988). *Science*, **241**, 669–674.

LANGRIDGE, R., MARVIN, D. A., SEEDS, W. E., WILSON, H. R., HOOPER, C. W., WILKINS, M. H. F. & HAMILTON, L. D. (1960). *J. Mol. Biol.* **2**, 38–64.

LATTMAN, E. E. (1989). *Proteins*, **5**, 149–155.

MÜLLER, J. J. (1983). *J. Appl. Cryst.* **16**, 74–82.

NINIO, J., LUZZATI, V. & YANIV, M. (1972). *J. Mol. Biol.* **71**, 217–229.

PAVLOV, M. YU. & FEDOROV, B. A. (1983). *Biopolymers*, **22**, 1507–1522.

PAVLOV, M. YU., SINEV, M. A., TIMCHENKO, A. A. & PTITSYN, O. B. (1986). *Biopolymers*, **25**, 1385–1397.

PERKINS, S. J. (1986). *Eur. J. Biochem.* **157**, 169–180.

PICKOVER, C. A. & ENGELMAN, D. M. (1982). *Biopolymers*, **21**, 817–831.

SCHOENBORN, B. P. (1988). *J. Mol. Biol.* **201**, 741–749.

STUHRMANN, H. B. (1970*a*). *Acta Cryst.* A**26**, 297–306.

STUHRMANN, H. B. (1970*b*). *Z. Phys. Chem. Frankfurt*, **72**, 177–184, 185–198.

SVERGUN, D. I. (1991). *J. Appl. Cryst.* **24**, 485–492.

SVERGUN, D. I. (1994). *Acta Cryst.* A**50**, 391–402.

THANKI, N., THORNTON, J. M. & GOODFELLOW, J. M. (1988). *J. Mol. Biol.* **202**, 637–657.