Journal of
Applied Crystallography
ISSN 0021-8898

MASSHA – a graphics system for rigid-body modelling of macromolecular complexes against solution scattering data

P. V. Konarev, M. V. Petoukhov and D. I. Svergun

Journal of
Applied Crystallography
ISSN 0021-8898

Received 6 February 2001
Accepted 9 April 2001

MASSHA – a graphics system for rigid-body modelling of macromolecular complexes against solution scattering data

P. V. Konarev,[a,b] M. V. Petoukhov[a,b] and D. I. Svergun[a,b]*

[a]European Molecular Biology Laboratory, EMBL c/o DESY, Notkestrasse 85, D-22603 Hamburg, Germany, and [b]Institute of Crystallography, Russian Academy of Sciences, Leninsky pr. 59, 117333 Moscow, Russia. Correspondence e-mail: svergun@embl-hamburg.de

A program, *MASSHA*, for three-dimensional rendering and rigid-body refinement is presented. The program allows display and manipulation of high-resolution atomic structures and low-resolution models represented as smooth envelopes or ensembles of beads. *MASSHA* is coupled with a computational module to align structural models of different nature and resolution automatically. The main feature is the possibility for rigid-body refinement of the quaternary structure of macromolecular complexes. If high- or low-resolution models of individual subunits are available, the complex can be constructed on the computer display and refined to fit the experimental solution scattering data. *MASSHA* provides both interactive and automated refinement modes to position subunits in a heterodimeric complex (general case) and in a homodimeric complex with a twofold symmetry axis. Examples of application of the program (running on IBM PC compatible machines under Windows 9*x*/NT/2000) are given.

1. Introduction

A tremendous amount of information has been generated in recent decades about the structure and function of individual proteins. With the arrival of the 'post-genomics' era, thousands of new protein structures are expected to be produced using X-ray crystallography and nuclear magnetic resonance (NMR) (Burley, 2000; Edwards *et al.*, 2000). The most important cellular functions are accomplished, however, not by individual proteins but rather by macromolecular complexes. Such complexes are usually too large for NMR study and they often possess inherent structural flexibility, making them difficult to crystallize. Alternative experimental techniques and data analysis approaches are required for the analysis of large macromolecular assemblies to determine how the properties of individual macromolecules are linked to the overall structure and function of the macromolecular machines existing in the living cells.

Small-angle scattering of X-rays and neutrons (SAS) is widely used to analyse low-resolution structure and conformational changes of native biological macromolecules in solution (Feigin & Svergun, 1987). The scattering intensity $I(s)$ from a dilute monodisperse solution of macromolecules is proportional to the scattering from a single particle averaged over all orientations [here s denotes the modulus of the scattering vector $\mathbf{s} = (s, \Omega)$, where $s = (4\pi/\lambda)\sin\theta$, Ω is the solid angle in reciprocal space, λ is the wavelength and 2θ the scattering angle]. The main advantage of solution scattering is the possibility to study structure and structural dynamics of native particles under nearly physiological conditions. The main disadvantage of the technique is a significant loss of information caused by random orientation of particles in solution.

Solution scattering patterns are sensitive to changes in the quaternary structure of macromolecules; the method is thus particularly useful for the analysis of macromolecular complexes. Reliable methods for *ab initio* shape determination of particles at low resolution have recently been proposed and successfully implemented (Chacon *et al.*, 1998; Svergun, 1999; Svergun *et al.*, 1996; Walther *et al.*, 1999). Far more detailed models can be constructed when high-resolution structures of individual domains or subunits composing the complex are available. Assuming that the tertiary structure of the domains is largely preserved upon the complex formation, rigid-body modelling can be used to build the macromolecular complexes (Ashton *et al.*, 1997; Krueger *et al.*, 1997; Svergun, Aldag *et al.*, 1998). Thus, for an assembly of two subunits, the complex is constructed by varying six positional parameters describing the relative position and orientation of the second subunit with respect to the first one. The methods to compute solution scattering patterns accurately from atomic models and further to evaluate scattering from complex particles rapidly are well established (Svergun, 1991; Svergun *et al.*, 1995, 1994; Svergun, Richard *et al.*, 1998). These methods permit an exhaustive search of positional parameters to fit the experimental scattering from the complex. Such a straightforward search may, however, yield a model that perfectly fits the data but fails to display proper intersubunit contacts. Relevant biochemical information (*e.g.* on contacts between the specific residues) can be taken into account by using an interactive search based on visual criteria. An interactive modelling system, running on major Unix platforms (SUN, SGI, DEC), has been developed by us previously (Kozin & Svergun, 2000; Kozin *et al.*, 1997). The three-dimensional graphics program *ASSA*, coupled with computational modules, permits the manipulation of the subunits as rigid bodies while observing corresponding changes in the fit to the experimental data.

In the present paper, a new Wintel-based rigid-body modelling system is described, which offers enhanced possibilities for combining interactive and automated search strategies. The system includes a special modelling option for the practically important case of a homodimeric structure possessing a twofold symmetry axis. The

program is implemented on a PC under Win9*x*/NT/2000, currently the most easily accessible hardware platform for the majority of users.

2. Rigid-body modelling

Let us consider a complex consisting of two subunits (*A* and *B*). The scattering amplitudes from the subunits centred at the origin in reference orientations are denoted as $A(s)$ and $B(s)$, respectively. An arbitrary complex can be constructed by fixing the first subunit and rotating and moving the second one. The rotation is described by the Euler angles α, β and γ (Edmonds, 1957) and the shift by a vector $\mathbf{u} = (u_x, u_y, u_z)$, so that the entire operation is described by six parameters.

Denoting $C(s)$ the scattering amplitude from a displaced second subunit, the scattering from the complex is expressed as (Svergun, 1991, 1994)

$$I(s) = I_A(s) + I_B(s) + 2\langle A(\mathbf{s})\, C^*(\mathbf{s})\rangle_\Omega, \tag{1}$$

where $\langle \ldots \rangle_\Omega$ stands for the spherical average in the reciprocal space.

It is convenient to represent the scattering amplitudes using spherical harmonics $Y_{lm}(\Omega)$ as

$$A(\mathbf{s}) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} A_{lm}(s)\, Y_{lm}(\Omega), \tag{2}$$

$$B(\mathbf{s}) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} B_{lm}(s)\, Y_{lm}(\Omega), \tag{3}$$

$$C(\mathbf{s}) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} C_{lm}(s)\, Y_{lm}(\Omega). \tag{4}$$

Owing to the orthogonal properties of spherical harmonics, the spherical average in equation (1) leads to a closed expression

$$I(s) = 2\pi^2 \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \{|A_{lm}(s)|^2 + |B_{lm}(s)|^2 + 2\mathrm{Re}[A_{lm}(s)\, C^*_{lm}(s)]\}. \tag{5}$$

Assuming that the structures of the subunits are known, the scattering amplitudes and, further, the partial functions $A_{lm}(s)$ and $B_{lm}(s)$ can be computed from the atomic models as described previously (Svergun *et al.*, 1995; Svergun, Richard *et al.*, 1998). The partial functions $C_{lm}(s)$ of the rotated and moved second subunit are expressed analytically as (Svergun, 1991, 1994)

$$C_{lm}(s) = \sum_{k=-l}^{l} D^l_{mk}(\alpha, \beta, \gamma)\, C^0_{lk}(s, \mathbf{u}) \tag{6}$$

where $D^l_{mk}(\alpha, \beta, \gamma)$ denotes the elements of the finite rotation matrix (Edmonds, 1957). Assuming without loss of generality that the entire complex is rotated to make the direction $\mathbf{u}$ coincide with the Z axis, the functions $C^0_{lk}(s, \mathbf{u})$ are (Svergun *et al.*, 1997)

$$C^0_{lm}(s, \mathbf{u}) = (-1)^m \sum_{p=0}^{\infty} j_p(su) \sum_{k=|l-p|}^{l+p} d_{lm}(k, p) \sum_{j=-k}^{k} B_{kj}(s), \tag{7}$$

where $j_p(x)$ are the spherical Bessel functions and the coefficients $d_{lm}(k, p)$ are represented through the 3*j* Wigner symbols (Edmonds, 1957):

$$d_{lm}(k, p) = i^p (2p+1)[(2l+1)(2k+1)]^{1/2} \times \begin{pmatrix} l & p & k \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l & p & k \\ -m & 0 & m \end{pmatrix}. \tag{8}$$

The idea of rigid-body modelling against solution scattering data is to determine the six positional parameters of the second subunit, minimizing the discrepancy

$$\chi^2 = [1/(N-1)] \sum_{j=1}^{N} \{[I(s_j) - I_{\exp}(s_j)]/\sigma(s_j)\}^2, \tag{9}$$

where $I_{\exp}(s)$ is the experimental intensity specified at N points s_j, $j = 1, \ldots, N$, and $\sigma(s_j)$ is the corresponding standard deviation. The rigid-body modelling algorithm is implemented in the computer programs *ALM22INT* (computation of the fit to the experimental data for a given configuration) and *DIMREF* (refinement by a grid search in the vicinity of the current configuration).

3. Three-dimensional graphics

The three-dimensional display and manipulation of macromolecular models is performed by a Fortran program *MASSHA* ('modelling of atomic structures and shape analysis') running under Windows NT, Windows95/98 and Windows2000. It uses the Win32 API and QuickWin graphics libraries of Compaq Visual Fortran for the three-dimensional rendering and for the menu-driven user interface. *MASSHA* can represent up to ten different macromolecular structures simultaneously, each body being provided with its own drawing features, such as colour, brightness and type of representation.

The major items in the user interface include 'Add/Remove body' menus for loading and removing graphic objects, 'Zoom in/out' menus for scaling the graphics window, a 'Commands' menu to perform operations with the objects (*e.g.* coordinate transformations, changing attributes, saving current positions), and a 'Modelling' menu, for opening an interface to interactive and automatic rigid-body modelling. On-line help about the main commands and options is available. *MASSHA* represents three-dimensional objects either as solid bodies or as collections of points. The first option is usually employed to visualize low-resolution models and displays the particle surface with the help of an angular envelope function $r = F(\omega)$, where (r, ω) are spherical coordinates in real space. The information is stored in an ASCII file containing multipole coefficients of the surface representation with the spherical harmonics (Stuhrmann, 1970)

$$F(\omega) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} f_{lm}(s)\, Y_{lm}(\omega). \tag{10}$$

The files containing the coefficients f_{lm} (`*.flm`) are generated, in particular, by the *ab initio* shape-determination program *SASHA* (Svergun *et al.*, 1996) and by the program *CRYSOL* (Svergun *et al.*, 1995). For the display, the `*.flm` files are transformed into binary surface-representation files (`*.sld`) by the program *FLM2SLD* (invoked by *MASSHA* automatically). The triangulation algorithm used to represent the surface is described elsewhere (Kozin *et al.*, 1997). The solid-body objects are represented either as wire frames or as Gouraud shaded surfaces (Foley *et al.*, 1990).

For the second type of objects, *MASSHA* reads the coordinates in the standard Protein Data Bank (PDB) format (Bernstein *et al.*, 1977). The graphic entry is an atomic model represented by either a C_α backbone or a P backbone. The PDB format (`*.pdb`) can also be used to read and display electron microscopic models or bead models provided, for example, by *ab initio* shape determination (Svergun, 1999). Interface programs are available to convert common electron microscope formats [*e.g.* contour lines or volumetric *SPIDER* format (Frank *et al.*, 1996)] into bead models in the PDB format. The beads can be displayed as open circles or as shaded spheres with variable radii. Fig. 1 illustrates different modes of shape/atomic structure representation using high- and low-resolution models of lysozyme [PDB entry 6lyz (Diamond, 1974)].

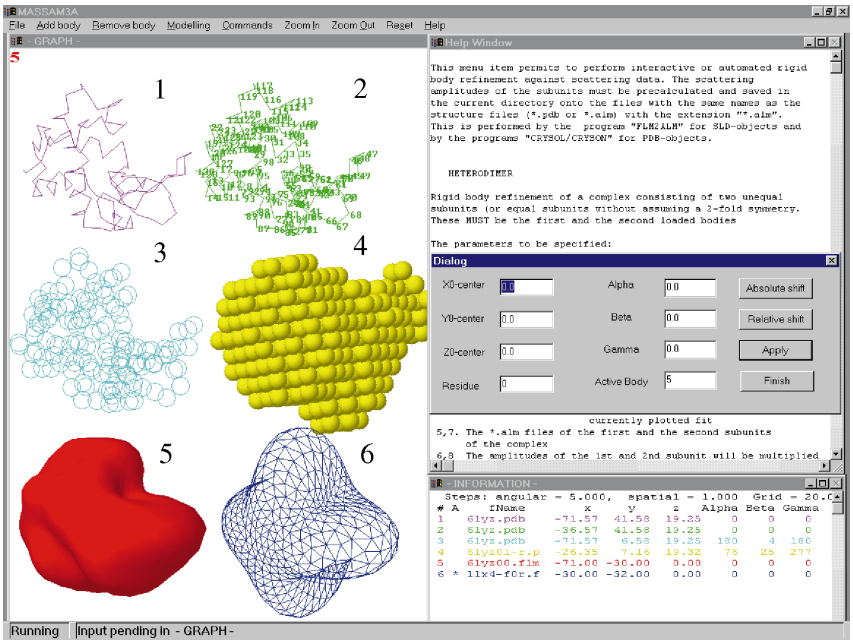

Figure 1
User interface of *MASSHA* illustrating the different modes of shape/atomic structure representation for the example of lysozyme: (1) and (2) C_α chain without and with residue numbering, respectively; (3) C_α chain as a collection of circles; (4) *ab initio* bead model obtained by the program *DAMMIN* (Svergun, 1999); (5) Gouraud shaded envelope computed by *CRYSOL* (Svergun *et al.*, 1995); wire-frame *ab initio* envelope from the program *SASHA* (Svergun *et al.*, 1996).

Each graphic object (shape or structure) has its own graphic features and current transformation matrix (the latter is changed upon rotations and shifts). The information about the current location of an object is available *via* a 'hotkey'. Relative shifts and rotations can be performed either by menu dialog or by hotkeys. The default angle increment for the hotkey transformations is $\Delta\varphi = 5^\circ$ and the default space increment ($\Delta X = \Delta Y = \Delta Z$) is defined by *MASSHA* to be about 1/20 of the screen window (both increments can be changed by the user). A transformation menu dialog enables the positioning of an object in the absolute coordinate system. *MASSHA* permits PDB objects to be saved at their current positions, which is convenient for storing the results of interactive or automated model building. An option 'Align PDB' automatically superimposes two arbitrary low-resolution or high-resolution objects. This option invokes the program *SUPCOMB* (Kozin & Svergun, 2001), aligning two bodies represented by ensembles of points without information about the correspondence between the points. After *SUPCOMB* is terminated, the second object is moved to the position that provides the best alignment with the first one.

MASSHA is oriented toward simultaneous display of high- and low-resolution models and toward rigid-body modelling. Compared with specialized graphics systems like *O* (Jones *et al.*, 1991), *RasMol* (Sayle & Milner-White, 1995) and *Chime* (Horton, 1999), *MASSHA* uses simplified representations of atomic structures. However, it provides all the necessary manipulation tools and a user interface for rigid-body modelling, as described below.

4. Heterodimer modelling

To perform rigid-body modelling of a heterodimeric structure, the multipole components of the scattering amplitudes of the two subunits have to be pre-computed and stored as separate files. This is performed by the programs *FLM2ALM* for the envelope models and *CRYSOL* (Svergun *et al.*, 1995) for atomic models. The two subunits are loaded as separate bodies and the modelling is started by selecting the 'Heterodimer' option from the 'Modelling' menu (Fig. 2). The following parameters have to be specified in the 'Heterodimer' dialog box.

(i) Input file containing experimental data to be fitted. The name may be selected with a file chooser ('Select' option).

(ii) Scale for **s** vector, *i.e.* abscissa units in the input file.

(iii) Fitting range (as a fraction of the maximum s value).

(iv) Output file containing the currently plotted fit.

(v) The pre-computed `*.alm` files of the first and second subunits.

(vi) The pre-multipliers (*e.g.* contrasts) of the two subunits.

MASSHA warns the user if the information is incorrect or incomplete. The rigid-body refinement menu contains two modelling buttons: interactive ('Compute') and automatic ('Refine'). The button 'Compute' runs the program *ALM22INT*, to calculate the complex scattering intensity, and displays the fit to the experimental data in a separate 'Curve fit' window. The user may change the positions of the two subunits by manipulating the objects in the three-dimensional window and can recalculate the fit by clicking 'Compute'. The button 'Refine' runs an automated refinement program *DIMREF*. The second subunit is moved around its current position by $\pm\Delta(XYZ)$ (*i.e.* three points along each direction X, Y and Z are probed). Its orientation is varied around the current orientation by rotating around X, Y and Z by $\pm 2\Delta\varphi$, *i.e.* five rotations around each axis are probed. The user can choose the residue number to define the centre of rotation (by default the second body is rotated around its geometric centre). After the refinement, the second body is moved to the position providing the best fit and its initial position is added as the third body to visualize the change of the model. The fit to the experimental data is displayed in the 'Curve fit' window. The user can further adjust the solution manually or continue the fitting procedure starting from the new position of the second body. To speed up the computations, the amplitudes of the subunits submitted to *ALM22INT* and *DIMREF* are always rotated to direct the vector **u** along the Z axis (Svergun *et al.*, 1997). This permits the use of equations (6), (7) and (8); the response in the 'Compute' mode is then a matter of fractions of a second on a Pentium III 600 MHz machine. A default refinement cycle takes about 1 min.

Fig. 2 presents an example of rigid-body modelling of the structure of *A. hydrophila* zinc β-lactamase using *MASSHA*. This protein, with molecular mass 25 kDa, belongs to the family of β-lactamases typically exhibiting activity against a broad range of β-lactamase antibiotics, but, in contrast to other known proteins of this family, the β-lactamase from *A. hydrophila* shows a narrow substrate profile (Payne, 1993). Moreover, conserved residues attributed to metal binding differ from other known zinc β-lactamases. According to the sequence alignment (30% homology), the tertiary structure of *A. hydrophila* β-lactamase should be similar to that of *B. cereus* β-lactamase [the latter is known from an X-ray crystallographic study

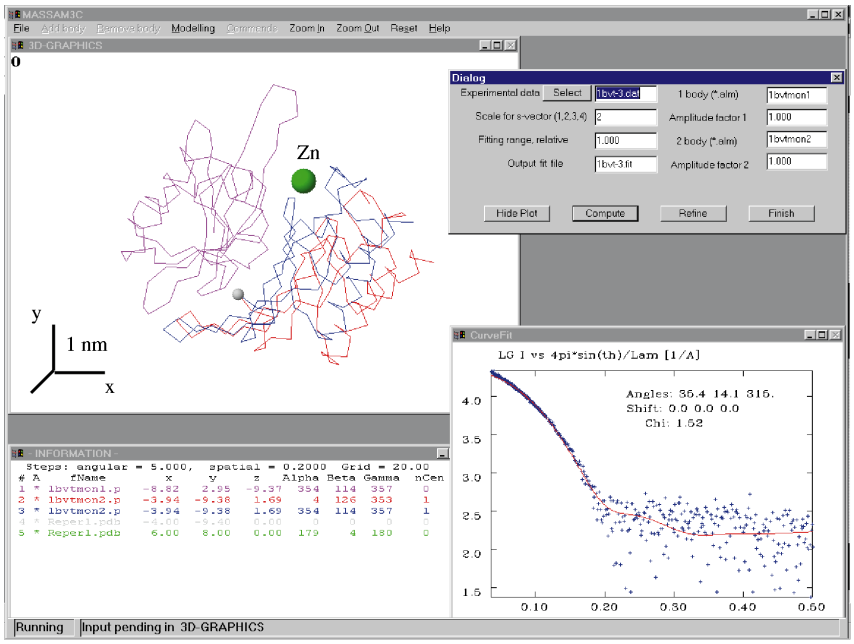

Figure 2
User interface of *MASSHA* for rigid-body modelling of the zinc β-lactamase. The domain structure of *B. cereus* β-lactamase is displayed in magenta and blue chains; the refined position of the second domain in *A. hydrophila* β-lactamase is shown in red. The pivotal residue is shown as a grey sphere and the active centre as a green sphere. The best fit to the experimental data is displayed in the bottom right corner.

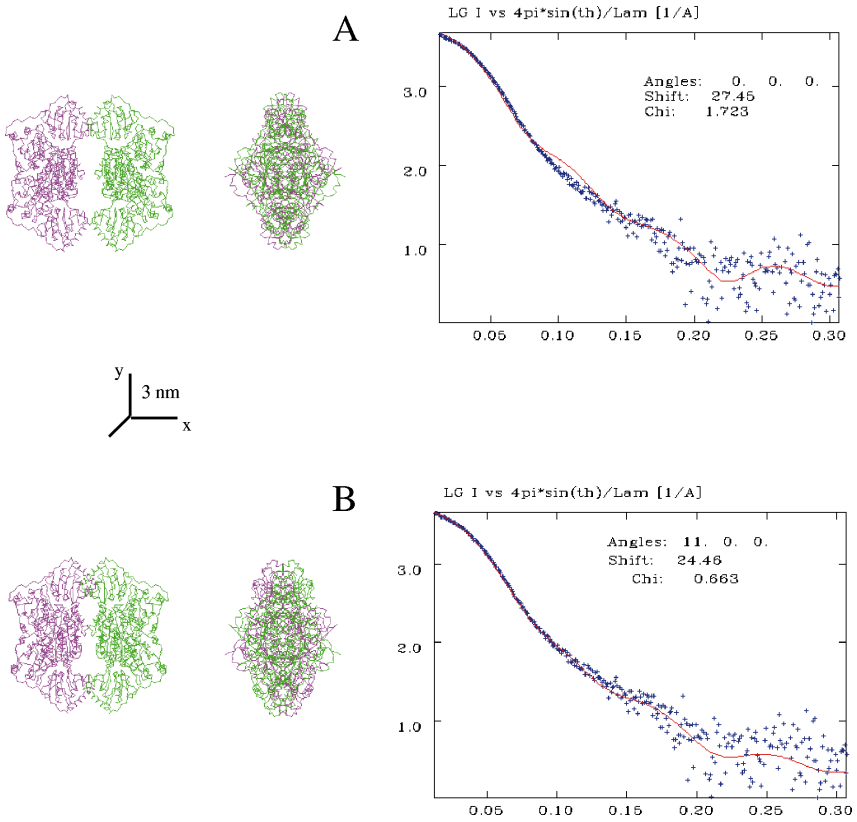

Figure 3
Rigid-body modelling of the structure of tetrameric yeast PDC in terms of two dimers: (*a*) original crystallographic structure; (*b*) model provided by *MASSHA* using the 'Homodimer' option. The diagrams of the atomic models and of the fits to the experimental data were prepared using the copy/paste utility of *MASSHA*. Side views (middle column) are rotated by 90° around *Y*.

(Carfi *et al.*, 1998)]. The aim of the study was to verify the quaternary structure similarity between the two enzymes.

The synchrotron radiation scattering data were collected on the EMBL beamline at HASYLAB (DESY, Hamburg) (Boulin *et al.*, 1988) in the range of momentum transfer $0.2 < s < 5.5$ nm^{-1}. Solutions with protein concentrations of 2, 4, 8 and 16 mg ml^{-1} were measured and a composite scattering curve was obtained using standard procedures. The experimental radius of gyration of 2.00 (3) nm noticeably exceeded the value computed from the crystallographic model of *B. cereus* β-lactamase (1.80 nm). A comparison of the experimental and theoretical curves (χ = 2.82) suggested that *A. hydrophila* β-lactamase in solution is significantly more anisometric than *B. cereus* β-lactamase. As the crystal structure of the latter displays two distinct domains (shown in Fig. 2 in magenta and blue), rigid-body refinement against experimental data was performed. The centre of rotation of the moving domain was set to the residue in the hinge region (grey sphere). The refined position of the second domain (red in Fig. 2) yields a good fit to the experimental data, with χ = 1.52. The zinc-containing active centre of the protein (displayed as a green sphere) becomes more exposed to the solvent in the refined model and this may play an important role for the catalytic activity of the β-lactamase in solution. One should, of course, bear in mind that the structures constructed by rigid-body refinement reflect overall positions of subunits and this modelling may distort the tertiary structure of the interface (*e.g.* the active-site geometry). Building of biochemically reasonable models based on the results of rigid-body refinement may therefore require more sophisticated treatment, in particular, appropriate alterations of the tertiary structure, using specialized software packages like *O* (Jones *et al.*, 1991).

5. Homodimer modelling

A special mode is provided by *MASSHA* for the practically important case of a homodimeric structure possessing a twofold symmetry axis. Given the structure of a monomer, the homodimer is defined by the orientation of the monomer and the distance between the centre of the latter and the twofold axis, *i.e.* by four parameters instead of six as for a heterodimer. The particle symmetry thus imposes significant restrictions on the possible model and the scattering intensity from the complex, equation (5), depends, in this case, only on the amplitudes of the first body (monomer). Assuming that the twofold axis coincides with *Y* and the monomers are separated along *Z* by $2u_z$, the intensity is expressed as (Svergun *et al.*, 1997)

$$I(u_z, s) = 2\sum_{l=0}^{\infty}\sum_{m=-l}^{l}\left|(-1)^m\sum_{p=0}^{\infty}\mathrm{Re}[i^p j_p(su_z)(2p+1)\times\sum_{k=|l-p|}^{l+p}[(2l+1)(2k+1)]^{1/2}\times\begin{pmatrix} l & p & k\\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix} l & p & k\\ -m & 0 & m\end{pmatrix}\sum_{j=-k}^{k}A_{kj}(s)\right|^2, \quad (11)$$

where $A_{kj}(s)$ are the amplitudes of the rotated monomer computed as in (6).

The 'Homodimer' option of *MASSHA* is used after loading the first body (monomer). By pressing 'Homodimer', the second monomer is generated automatically, assuming Y to be the twofold symmetry axis. One should specify the parameters as in the 'Heterodimer' modelling, with the exception of the second-body amplitude and factor. The button 'Compute' runs the interactive modelling program *ALM2INT2* and the button 'Refine' starts the automated refinement program *DIMREF2*. In the latter mode, the separation between the monomers is varied by $\pm 4\Delta X$, probing a total of five displacements, and the monomer is rotated around X, Y and Z by $\pm 5\Delta\varphi$ (*i.e.* eleven rotations around each axis). After the refinement, the homodimer providing the best fit is displayed along with the first monomer in the initial position, added as the third body. The user can further adjust the solution manually or continue the automated refinement from the new configuration. In the 'Homodimer' mode, movements and/or rotations of the first body lead to appropriate operations with the symmetrically related second body. The user may also apply operations preserving the symmetry (rotation and movements along the Y axis) of the entire homodimer.

An example of the homodimer modelling is given by the rigid-body refinement of tetrameric yeast pyruvate decarboxylase (PDC) in Fig. 3. Yeast PDC is a thiamin diphosphate-dependent enzyme involved in some steps of alcoholic fermentation. PDC is a tetramer of molecular weight 236 kDa at low pH and dissociates into dimers at high pH (Koenig *et al.*, 1993). The theoretical scattering pattern computed from the crystallographic model [PDB entry 1pvd (Arjunan *et al.*, 1996), Fig. 3*a*] deviates significantly from the experimental data (Fig. 3*a*, right, χ = 1.723). Svergun *et al.* (2000) performed a time-consuming interactive rigid-body refinement of the structure in terms of two dimers. Using the automatic refinement in *MASSHA* and starting from the crystallographic model, nearly the same model as obtained by Svergun *et al.* (2000) is obtained in a matter of minutes (Fig. 3*b*). The dimers in the final model are tilted by 11° around the X axis (left view) and the distance between them is diminished by 0.4 nm. The model provides a good fit to the experimental data, with χ = 0.66. The root-mean-square displacement of atoms in the refined model is 0.57 nm relative to the structure in the crystal. This large conformational change can be explained by weak contacts between the dimers in the crystallographic model (Fig. 3*a*, left panel) so that the crystallographic environment is able to influence the quaternary structure of the enzyme significantly.

6. Conclusions

Rigid-body modelling permits the useful combination of high-resolution crystallographic or NMR information and low-resolution information obtained by solution scattering. Interactions between subunits in macromolecular complexes are non-covalent in nature and this often permits one to move and rotate the subunits as rigid bodies. In particular, the energy of such interactions is comparable with that between neighbouring molecules in the crystals. The rigid-body refinement technique is thus also a useful tool for the analysis of differences in the quaternary structure of multisubunit enzymes caused by crystal packing forces (Svergun *et al.*, 2000).

There are many excellent three-dimensional graphics programs available for the representation and analysis of high-resolution structures. Our aim was not to compete with these programs, but to create a complementary tool for rigid-body modelling against solution scattering data. The three-dimensional display functionality of *MASSHA* is comparable with that provided by specialized packages like *SITUS* (Wriggers *et al.*, 1999) or *MASKIT* (Kleywegt & Jones, 1999). We wrote a relatively simple three-dimensional graphics module coupled with computational programs, in the frame of an easy-to-use graphical user interface. Not much attention was paid to sophisticated presentations, like ribbons, ball-and-stick diagrams, *etc*., irrelevant for rigid-body modelling. Instead, the three-dimensional module provides full control over the absolute and relative positions of the objects. Similarly, *MASSHA* does not offer an automatically coloured space-filling display of atomic structures, but provides an option to select the bead radius, which is essential for the display of low-resolution models.

In contrast to the earlier Unix-based system (Kozin & Svergun, 2000; Kozin *et al.*, 1997), *MASSHA* runs on a Wintel platform, which makes the rigid-body modelling available for numerous PC users. Furthermore, *MASSHA* provides enhanced options for interactive modelling and the possibility of an automated refinement in the vicinity of the current configuration. A practically very important option to account for the twofold symmetry in homodimeric complexes is included. A unique feature of *MASSHA* is the option for automated alignment of the objects by calling the program *SUPCOMB* (Kozin & Svergun, 2001).

The program package *MASSHA*, along with the computational modules and full instructions, is available from the EMBL Web page http://www.embl-hamburg.de/ExternalInfo/Research/Sax/index.html.

This work was supported by the Russian President's Grant for post-graduate students training abroad (1999–2000) and by an EMBL fellowship for PVK, and by the INTAS grants YSF 00-50, YSF 00-53 and 00-243. The authors thank U. Heinz for providing the protein samples, S. Koenig for providing the experimental data and M. Malfois for help with the synchrotron X-ray scattering measurements.

References

Arjunan, P., Umland, T., Dyda, F., Swaminathan, S., Furey, W., Sax, M., Farrenkopf, B., Gao, Y., Zhang, D. & Jordan, F. (1996). *J. Mol. Biol.* **256**, 590–600.

Ashton, A. W., Boehm, M. K., Gallimore, J. R., Pepys, M. B. & Perkins, S. J. (1997). *J. Mol. Biol.* **272**, 408–422.

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.

Boulin, C. J., Kempf, R., Gabriel, A. & Koch, M. H. J. (1988). *Nucl. Instrum. Methods A*, **269**, 312–320.

Burley, S. K. (2000). *Nature Struct. Biol.* **7**(Suppl.), 932–934.

Carfi, A., Duée, E., Galleni, M., Frère, J.-M. & Dideberg, O. (1998). *Acta Cryst.* D**54**, 313–323.

Chacon, P., Moran, F., Diaz, J. F., Pantos, E. & Andreu, J. M. (1998). *Biophys. J.* **74**, 2760–2775.

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

Edmonds, A. R. (1957). *Angular Momentum in Quantum Mechanics. Investigations in Physics*, Vol. 4. Princeton University Press.

Edwards, A. M., Arrowsmith, C. H., Christendat, D., Dharamsi, A., Friesen, J. D., Greenblatt, J. F. & Vedadi, M. (2000). *Nature Struct. Biol.* **7**(Suppl.), 970–972.

Feigin, L. A. & Svergun, D. I. (1987). *Structure Analysis by Small-Angle X-ray and Neutron Scattering*. New York: Plenum Press.

Foley, J. D., van Dam, A., Feiner, S. K. & Hughes, J. F. (1990). *Computer Graphics. Principles and Practice*. New York: Addison-Wesley.
Frank, J., Radermacher, M., Penczek, P., Zhu, J., Li, Y., Ladjadj, M. & Leith, A. (1996). *J. Struct. Biol.* **116**, 190–199.
Horton, R. M. (1999). *Biotechniques*, **26**, 874–876.
Jones, T. A., Zou, J.-Y. & Cowan, S. W. (1991). *Acta Cryst.* A**47**, 110–119.
Kleywegt, G. J. & Jones, T. A. (1999). *Acta Cryst.* D**55**, 941–944.
Koenig, S., Svergun, D., Koch, M. H., Hubner, G. & Schellenberger, A. (1993). *Eur. Biophys. J.* **22**, 185–194.
Kozin, M. B. & Svergun, D. I. (2000). *J. Appl. Cryst.* **33**, 775–777.
Kozin, M. B. & Svergun, D. I. (2001). *J. Appl. Cryst.* **34**, 33–41.
Kozin, M. B., Volkov, V. V. & Svergun, D. I. (1997). *J. Appl. Cryst.* **30**, 811–815.
Krueger, J. K., Olah, G. A., Rokop, S. E., Zhi, G., Stull, J. T. & Trewhella, J. (1997). *Biochemistry*, **36**, 6017–6023.
Payne, D. J. (1993). *J Med. Microbiol.* **39**, 93–99.
Sayle, R. A. & Milner-White, E. J. (1995). *Trends Biochem Sci.* **20**, 374–376.
Stuhrmann, H. B. (1970). *Z. Phys. Chem. Neue Folge*, **72**, 177–198.
Svergun, D. I. (1991). *J. Appl. Cryst.* **24**, 485–492.
Svergun, D. I. (1994). *Acta Cryst.* A**50**, 391–402.
Svergun, D. I. (1999). *Biophys. J.* **76**, 2879–2886.
Svergun, D. I., Aldag, I., Sieck, T., Altendorf, K., Koch, M. H., Kane, D. J., Kozin, M. B. & Grueber, G. (1998). *Biophys. J.* **75**, 2212–2219.
Svergun, D. I., Barberato, C. & Koch, M. H. J. (1995). *J. Appl. Cryst.* **28**, 768–773.
Svergun, D. I., Pedersen, J. S., Serdyuk, I. N. & Koch, M. H. (1994). *Proc. Natl Acad. Sci. USA*, **91**, 11826–11830.
Svergun, D. I., Petoukhov, M. V., Koch, M. H. & Konig, S. (2000). *J. Biol. Chem.* **275**, 297–302.
Svergun, D. I., Richard, S., Koch, M. H., Sayers, Z., Kuprin, S. & Zaccai, G. (1998). *Proc. Natl Acad. Sci. USA*, **95**, 2267–2272.
Svergun, D. I., Volkov, V. V., Kozin, M. B. & Stuhrmann, H. B. (1996). *Acta Cryst.* A**52**, 419–426.
Svergun, D. I., Volkov, V. V., Kozin, M. B., Stuhrmann, H. B., Barberato, C. & Koch, M. H. J. (1997). *J. Appl. Cryst.* **30**, 798–802.
Walther, D., Cohen, F. E. & Doniach, S. (1999). *J. Appl. Cryst.* **33**, 350–363.
Wriggers, W., Milligan, R. A. & McCammon, J. A. (1999). *J. Struct. Biol.* **125**, 185–195.