UROX 1.0: An interactive tool for fitting atomic models into Electron Microscopy reconstructions

Xavier Siebert, Jorge Navaza

UROX 1.0 : An interactive tool for fitting atomic models into Electron Microscopy reconstructions

Xavier Siebert* and Jorge Nazava*

* *Institut de Biologie Structurale* J.P. Ebel, *41 rue Jules Horowitz, F-38027 Grenoble; CEA; CNRS; Université Joseph Fourier, France*

Electron microscopy of a macromolecular structure can lead to three-dimensional reconstructions whose resolution is typically in the 30-10 Å range. Fitting atomic models of the macromolecular structure's individual components (e.g., obtained by X-ray crystallography or nuclear magnetic resonance) into an electron microscopy map allows an interpretation of the latter at nearly atomic resolution, providing insights into the interactions between the components. We present a graphical software designed for interactive fitting and refinement of atomic models into electron microscopy reconstructions. Several characteristics make it applicable over a wide range of cases and resolutions. First, calculations are performed in reciprocal space, which results in fast algorithms. This allows big volumes of the map to be used by taking into account the symmetry of the reconstruction (both in the calculations and in the graphical display) while working only with independent variables. Second, atomic models can be placed graphically in the map while the correlation between the model-based electron density and the electron microscopy reconstruction is computed and displayed in real-time. The positions and orientations of the models are refined by a least-squares minimization.

We illustrate these features on three practical cases with different symmetries and resolutions. The software as well as examples and user instructions are available free of charges at `http://mem.ibs.fr/UROX/`.

I INTRODUCTION

The three-dimensional structure of a macromolecular complex provides important information about the intricate interactions between its components. Some macromolecular complexes have been produced in homogeneous form, crystallized and analyzed at high resolution (3-2 Å or better) by X-ray crystallography (XR). However, most of them are too large or too unstable to crystallize, in which case only individual components of such complexes could be analyzed. On the other hand, electron microscopy (EM) allows 3D reconstructions of whole macromolecular complexes under conditions closer to in-vivo (or at least to in-vitro, but is limited to relatively low resolutions. By fitting atomic models of individual components into the EM reconstruction the latter can be interpreted at a resolution higher than its nominal one, thereby effectively bridging the different resolution ranges (for recent reviews see (*1*, *2*)). The first combination of EM and XR relied simply on visual inspection of the EM map and manual docking of the XR models. Despite the subjectivity inherent to such a procedure it lead to significant results such as the identification of several components of the adenovirus (*3*) and its binding footprint (*4*).

Recent methodological developments improved the quality of the fitting procedure. A variety of softwares are currently available, such as COAN (*5*), DOCKEM (*6*), EMFIT (*7*), FOLDHUNTER (*8*), SITUS (*9*), 3SOM (*10*), URO (*11*). Careful use of such softwares enhances the information that can be gained from the fitting compared to a manual docking procedure and allows errors to be estimated with other criteria than the human eye. However, important issues have to be treated with caution, defining criteria that a fitting software should fulfill:

1. *No masking*: At relatively low resolution (30-20 Å or lower) the electron density often does not contain enough distinctive features to unambiguously place an atomic model in the EM map. Therefore, fitting algorithms which require as a first step to carve out a piece of the EM reconstruction (a *mask* around putative positions of the models) would inevitably bias the solution.

2. *Speed*: Ideally one should use all the experimental information contained in an EM map without applying filtering techniques or simplified representations of the data, such as the

Reduced Vector Representations (*9*). This may result in prohibitively long fitting procedures, especially if the calculations are performed in real-space and/or if the structure is sizeable, such as an icosahedral virus.

3. *Symmetry*: The reconstruction often contains a particular symmetry (e.g. helicoidal, icosahedral) which should be taken into account adequately.

4. *Anisotropy*: There can be anisotropy in the EM data, as in the case of tomographic data with a missing wedge region. The fitting procedure must account for this anisotropy in order to avoid the introduction of a bias in the solution.

5. *Occupancies*: In the case of partial occupancy of some molecules it is necessary to assign a different weight (occupancy) to them.

6. *Graphics*: Fast visual inspection of putative solutions is necessary, in particular when numerical criteria do not discern the correct one. Mismatches between the EM map and the fitted molecules can point to inaccuracies or plausible modifications of the models, as those produced by flexible fitting (*9*, *12–14*).

Among the above-mentioned softwares only URO satisfies the first four criteria thanks to its reciprocal-space formulation (*11*). However, it relies on an external graphics software such as O (*15*), Coot (*16*) or Pymol (*17*), to place the atomic models in the EM map and to inspect the solutions. Switching back and forth between the computations and the graphics is time-consuming, and becomes cumbersome when several molecules and their symmetry mates have to be placed in the EM map.

The need for a close interaction between the graphics and the calculations is at the origin of the development of UROX, which aims at treating all the above-mentioned issues efficiently while providing a user-friendly interface.

II UROX DESIGN AND METHODOLOGY

We start by a summary of the reciprocal-space formalism (*11*) and then describe how this formalism is integrated with the graphics.

A Reciprocal-space fitting

The fitting problem is formulated in reciprocal space as the minimization of the so-called "quadratic misfit" (Q) between the electron density based on the molecules (including their symmetry mates) and the EM map (*11*). In real space Q is expressed as

$$Q = \frac{\int |\rho^{em}(\mathbf{r}) - \lambda\rho^{mod}(\mathbf{r})|^2 d^3\mathbf{r}}{\int |\rho^{em}(\mathbf{r})|^2 d^3\mathbf{r}}\,, \tag{1}$$

where $\rho^{em}(\mathbf{r})$ is the electron density of the EM map, $\rho^{mod}(\mathbf{r})$ the electron density of the independent molecules and their symmetry mates, and λ the relative scale between these two densities. The integral in Eq. 1 is performed on a volume containing the EM map. On the other hand, in reciprocal space Q is expressed as

$$Q = \frac{\int |F^{em}(\mathbf{s}) - \lambda F^{mod}(\mathbf{s})|^2 d^3\mathbf{s}}{\int |F^{em}(\mathbf{s})|^2 d^3\mathbf{s}}\,, \tag{2}$$

where $F^{em}(\mathbf{s})$ and $F^{mod}(\mathbf{s})$ are the Fourier transforms of $\rho^{em}(\mathbf{r})$ and $\rho^{mod}(\mathbf{r})$, respectively. Explicitly, $F^{mod}(\mathbf{s})$ is expressed in terms of the molecular scattering factors f_m of the independent molecules as (*11*)

$$F^{mod}(\mathbf{s}) = \sum_{m\in M}\sum_{g\in G} f_m(\mathbf{s}\mathbf{M_g}\mathbf{R_m})\exp[2\pi i\mathbf{s}(\mathbf{M_g}\mathbf{X_m} + \mathbf{T_g})]\,, \tag{3}$$

where m refers to one of the M independent molecules, located at the position $\mathbf{X_m}$ in the orientation $\mathbf{R_m}$ with respect to a reference position (as detailed in (*18*)), while g refers to the symmetry operator represented by the translation $\mathbf{T_g}$ and the rotation $\mathbf{M_g}$. The F^{mod} are thus functions of the positional variables of the independent molecules.

It is worth noticing that minimizing Eq. 2 amounts to maximizing the correlation coefficient

$$CC = \frac{\int \overline{F^{em}(\mathbf{s})} F^{mod}(\mathbf{s}) d^3\mathbf{s}}{\sqrt{\int |F^{em}(\mathbf{s})|^2 d^3\mathbf{s}}\sqrt{\int |F^{mod}(\mathbf{s})|^2 d^3\mathbf{s}}}\,. \tag{4}$$

In practice, integrals are calculated on discrete regularly spaced grids, which amounts to substitute the integrals on the continuous variable $\mathbf{s}$ by summations on the discrete variable $\mathbf{h}$. Eq. 1 and Eq. 2 are strictly equivalent, both for continuous and for discrete Fourier transforms. This is not a "superficial invocation of Parseval's theorem" as stated in (*19*), but its rigorous application.

The first three points raised in the introduction are addressed naturally when fitting is performed in reciprocal space. Indeed, one can use the entire EM reconstruction, or at least a substantial part of it containing all the independent molecules and several of their symmetry mates. Moreover, the least-squares refinement algorithm used to minimize Q, where the symmetry of the reconstruction is efficiently taken into account by Eq 3, has proven to be fast and to have a large convergence radius (*20*). This formalism is implemented in URO and has been successfully applied to several fitting problems (*21–43*).

The reciprocal-space formulation is general in that it can be applied to low-resolution maps derived from other experimental sources than EM (e.g., Small Angle X-ray Scattering) and it can use electron densities or ensembles of NMR structures instead of an atomic model. Also, we found that the so-called "R-factor" widely used in crystallography and specific to the reciprocal space,

$$R = \frac{\sum_{\mathbf{h}} ||F_{\mathbf{h}}^{em}| - |F_{\mathbf{h}}^{mod}||}{\sum_{\mathbf{h}} |F_{\mathbf{h}}^{em}|}, \tag{5}$$

helps assessing the resolution of the EM reconstruction.

B Interaction with the graphics

The interaction between the calculations and the graphics is made possible through the use of graphical libraries from the Visual ToolKit (VTK, (*44*)) and through a substantial re-writing of the reciprocal-space code. The Python language is used to wrap together the computational and graphical subroutines. The position and orientation of all molecules are extracted by graphical subroutines each time a molecule is moved by the user. This information is given to a subroutine that computes a correlation coefficient (Eq. 4) which is then returned to the display. The reciprocal-space formulation is fast enough to allow this interaction to occur in real-time, as detailed below.

A Graphical User Interface (GUI) is also provided, with a modular architecture so that the user

can add or modify components if necessary. The issues mentioned in the introduction are treated as follows:

1. *No masking*: By default, calculations are performed with the entire (unmasked) EM map provided by the user. An option to restrict the *field of view* is available, using the "BoxWidget" tool from the VTK libraries, which provides a mask that can be manipulated interactively to inspect local portions of the EM map. Contrary to real-space fitting procedures this mask is not used in the computations but serves only for visual purposes.

2. *Speed*: The computation of the correlation coefficient (CC, Eq. 4) takes about $10^{-7}s$ per Fourier coefficient and per symmetry operation on a single-processor machine, which makes it possible to compute CC in real-time. However, a practical limitation arises: if the number of Fourier coefficients multiplied by the number of symmetry operators becomes greater than 10^5 (e.g., the Rotavirus example below), the computations are too slow to allow real-time interactions if all coefficients are used. Nevertheless, if N independent molecules have to be placed in the map, only $6N$ parameters have to be determined, corresponding to the positions and orientations of the molecules. Even though all Fourier coefficients are not independent from each other, the fitting problem is widely over-determined. Several options are available to limit the number of coefficients, thereby allowing real-time computations:

   - Perform the calculations at a lower resolution.
   - Extract a subset of Fourier coefficients by a decimation procedure.
   - Select only those coefficients belonging to the asymmetric unit in reciprocal space.

   The first two options are applicable in a general case while the latter is particularly useful in the case of high point-group symmetry (e.g., icosahedral symmetry). Note that the loss of high resolution is not critical for the real-time computation since its goal is to provide a suitable starting point for the subsequent least-squares minimization procedure which uses the whole resolution range of the data.

Moreover, the display can also be accelerated either by decimating the EM map, by using the above-mentioned "BoxWidget" and/or by taking advantage of the symmetry of the reconstruction.

3. *Symmetry*: Several built-in symmetries are available: icosahedral, tetrahedral, helicoidal, dihedral (D_n), cyclic (C_n), including of course the case of no symmetry (called C_1 or P_1). These symmetries were chosen to cover most of the practical cases in EM but the user has the possibility to add another one if it is not in the provided set. The symmetry is included in the calculations (Eq. 3), with the option of defining a different set of operators for each molecule. This is useful for example in the case of a trimeric protein lying on a 3-fold symmetry axis, as in the icosahedral rotavirus described in the *Applications* section. Additionally, when the user moves one independent molecule in the map the symmetry mates have the ability to move along in real-time.

4. *Anisotropy*: Anisotropy of the data can be taken into account. For example, Fourier coefficients falling into a missing wedge region can be detected and naturally excluded from the summation in Eq. 2 or Eq. 4; this would not be possible in the real space formulation (Eq. 1).

5. *Occupancies*: It is possible to attribute to each atomic model a partial occupancy by accordingly weighting its contribution to the term $F_{\mathbf{h}}^{mod}$ in the discrete form of Eq. 2.

6. *Graphics*: A close connection between the graphics and the computations is provided. In particular, each time a molecule is moved on the display, the new position and orientation of this molecule are used to compute an updated CC. This feature is particularly useful to place the molecules at initial positions in the EM map and to analyze the results. Several other options are available via the GUI such as map manipulations, least-squares refinement, exhaustive searches, refinement of the relative geometric scale between the map and the models.

C Flowchart

Practically, our software consists of five main steps. First, an EM map is loaded and its resolution limits are defined. Because the EM map magnification can be inaccurate by as much as 5%, a scale factor is optionally introduced when calculating the Fourier transform of the map. This scale factor can be refined, as detailed in the section "Strategy and Optimization". Second, atomic models are either loaded from a file or downloaded from the Protein Data Base (PDB). The Fourier transform of each model is calculated at the resolution of the EM map with the model placed at a reference position (*18*). It may happen that several independent molecules correspond to the same atomic model. For example, the helical assembly formed by VP6, the major capsid protein of Rotavirus (*45*), contains two independent copies of the same VP6 trimeric model (see *Applications* section). In such a case the molecular scattering factors f_m (Eq. 3) are calculated only once for the unique model at a reference position, and used for all the independent molecules corresponding to this model. Third, the symmetry of the reconstruction is introduced, which defines a set of symmetry operators to be applied to the independent molecules. Only the operators that keep the center of mass of each symmetry mate within the limits of the EM map are relevant for the fitting. Fourth, the independent molecules are placed at an initial position in the EM map using the real-time CC as a guide. The set of relevant symmetry operators is updated automatically along the procedure, with the option of using either the same set for all independent molecules, or a different set for each one. Finally, the molecular positions and orientations are refined with a least-squares minimization. The overall strategy and the optimization options to enhance the results of this last step are described in the next section. At any point of the procedure, the positions and orientations of any molecule (including any symmetry mate) can be saved.

D Strategy and Optimization

The procedure leading to a converged fitting solution should only take a few minutes of CPU time on a single-processor computer (one cycle of least-squares minimization effectively takes about $2.5 \times 10^{-7} s$ per reflection, per symmetry operator and per molecule). The radius of con-

vergence of the least-squares minimization is roughly proportional to the resolution of the data. For example, data up to 20 Å typically lead to a radius of convergence of about 30 Å (*11*). As the resolution gets higher, the convergence radius gets smaller, enhancing the dependence on the initial positioning of the molecules.

Therefore, the usual strategy is to perform several cycles of minimization, starting from low resolution to high resolution, while monitoring in real-time the CC as well as the positions and orientations of the molecules. In the reciprocal-space formalism, changing resolution is straightforward because the Fourier transforms of the EM map and of the models (F^{em} and F^{mod}, respectively) are calculated only once at the resolution of the EM reconstruction.

Additionally, exhaustive searches (translational and/or rotational) can be performed in a user-defined region of the map. A full 6-dimensional search (3 rotations and 3 translations) is generally not useful (and fairly time-consuming) to perform since the positions of the molecules are found quite accurately by the least-squares algorithm, especially when symmetry is present. However, if all data up to high resolution are included at early stages of refinement, molecules may be trapped in false positions corresponding to a local maximum of CC (Eq. 4). In such a case, a rotational exhaustive sampling proves quite useful, as illustrated in the GroEl example in the "Applications" section. To accelerate the rotational sampling a Burdina-Lattman parameterization is used (*46*, *47*).

Other optimizations options are available. For example, because the EM map magnification can be inaccurate by as much as 5%, the absolute scale of the reconstruction is determined by automatically performing the least-squares refinement at several magnifications. Overall isotropic temperature factors (B-factors) of the molecules can also be refined. If the absolute handedness of the EM map is unknown, fitting can be performed with left- and right-handed maps, and the correlation coefficient can be used to discriminate between them.

Finally, the algorithm purposely does not prevent clashes between the different molecules placed in the EM map, again with the idea of using at best the experimental data and avoid biases. Indeed, a model may not exactly fit the EM map because the molecule underwent modifications, some of which may be taken into account by normal modes analysis. In the current implementation each independent molecule is considered as a rigid body, but inclusion of normal modes

analysis to allow for flexible docking (*12*) is underway. The amount of overlap between different molecules can therefore serve as a rough indication of the quality of the fit. In addition to CC other figures of merit are optionally computed to help assessing the quality of the solution.

III APPLICATIONS

A Rotavirus Capsid Proteins

The X-ray crystallographic structure of VP6, the major capsid protein of rotavirus (PDB code 1QHD (*48*)) was fitted into EM reconstructions corresponding to assemblages of different symmetries.

1 Helical VP6 assemblies

The helical high-pH VP6 assembly (referred to as "small tubes" (*45*)) was reconstructed to a resolution of 20 Å. This reconstruction was chosen to illustrate the difficulty in carving out a volume of density around one molecule (Figure 1). Indeed, although several VP6 trimers can be distinguished by eye, the density is continuous between them and it would be difficult to decide where to delineate the contour of a monomer. This problem is circumvented through the reciprocal space formulation by using an EM map containing several symmetry-related molecules (44 VP6 monomers, more than 20,000 Fourier coefficients to 20 Å). After optimization of the scale factor corresponding to the magnification of the EM map, we obtained a correlation of 94.1% and an R-factor of 33.4% (Figure 2), in agreement with a previously obtained result (*11*).

2 Icosahedral VP6 assemblies

We fitted the atomic model of VP6 into double- and triple-layer assemblies (DLP and TLP, respectively (*49*)). Both DLP and TLP are icosahedral (with a triangulation number (*50*) T=13 for the VP6 layer (*51, 52*)) and contain five independent VP6 molecules (four trimers and a monomer). The Fourier transform of each EM map leads to more than 650,000 coefficients at 20 Å resolution.

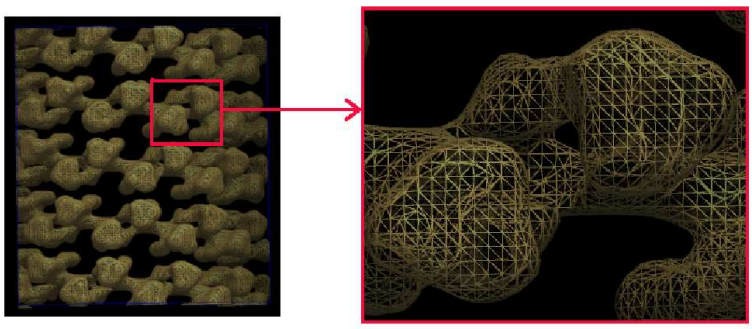

FIGURE 1. Electron Density map of part of a helical VP6 assembly ("small tubes") contoured at 1.5 σ. The zoom on a VP6 dimer reveals the contiguous density between VP6 monomers. The figure is a snapshot of UROX.

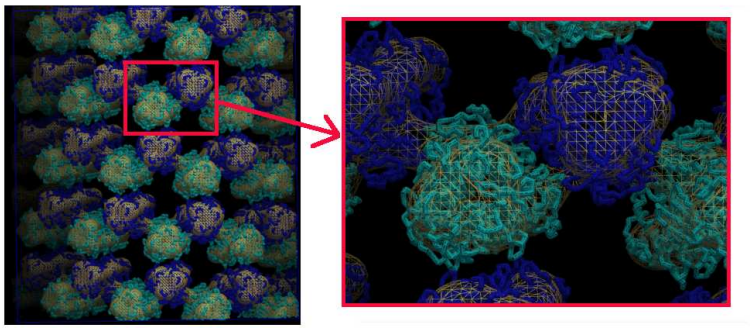

FIGURE 2. Electron Density map of part of a helical VP6 assembly ("small tubes"), contoured at 1.5 σ, with the result of the fit after refinement (CC=94.1%, R=33.4%). There are two independent VP6 trimers, colored in cyan and in blue; molecules related by symmetry are in the same color. The figure is a snapshot of UROX.

We used only about 11,000 coefficients belonging to the asymmetric unit of the icosahedron to reduce the computational cost to one second per refinement cycle. The resulting fit is shown in Figure 3. As described elsewhere (*49*) we fitted the VP6 atomic model in six reconstructions of viral particles containing different layers of capsid proteins from the rotavirus. The handedness of each reconstruction was checked by fitting in a left-handed and in a right-handed map. The EM magnification was estimated by fitting into a series of reconstructions with different scales (e.g., from 0.9 to 1.1). This example illustrates that the speed of the algorithm is instrumental, considering the number of fits to be performed.

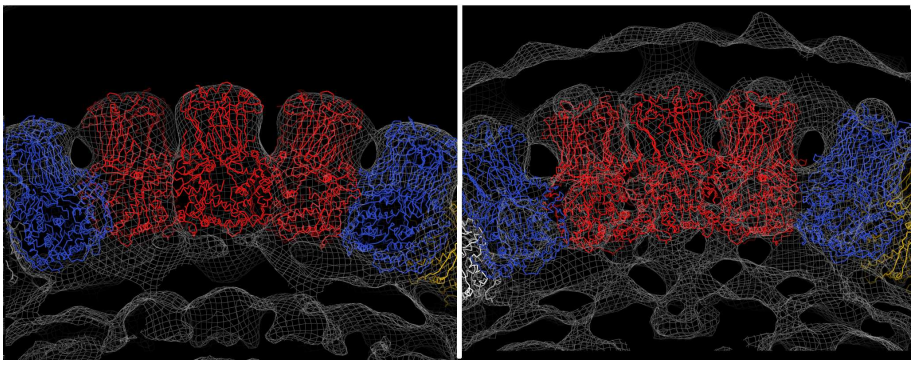

FIGURE 3. Electron Density map and results of the fitting of VP6 into double-layer (DLP, left) and triple-layer (TLP, right) assemblies. The VP6 molecules related by symmetry are in the same color. The figure was prepared with Pymol.

B GroEl

We fitted the crystallographic structure of GroEl (PDB code 1OEL, (*53*)) in a cryo-negative stain EM reconstruction of GroEl (*54*), which possesses a D_7 symmetry (one independent molecule). Using data up to a resolution of 20 Å (about 5000 Fourier coefficients), one cycle of least-squares refinement takes about $10^{-2}s$. To illustrate the problem of a molecule being trapped in an wrong position, the least-squares refinement was launched from an arbitrary position with data up to 20 Å. The least-squares minimization converged to a wrong "upside-down" local optimum (Figure 4). Although the difference between the good and the bad solutions may not be immediately obvious to the eye, the correlation coefficient allows to distinguish them clearly (good CC=89.6%; bad CC=74.0%). Starting from the bad solution, an exhaustive search for the rotational parameters allows one to recover the correct solution in a few minutes.

IV CONCLUSION

UROX is a user-friendly software for fitting atomic models into Electron Microscopy reconstructions. It uses a reciprocal-space formulation (*11*), which was adapted for interactive initial positioning of the molecules in the EM map, with real-time calculation and display of the correlation between them. The symmetry of the EM reconstruction is used both in the calculations and

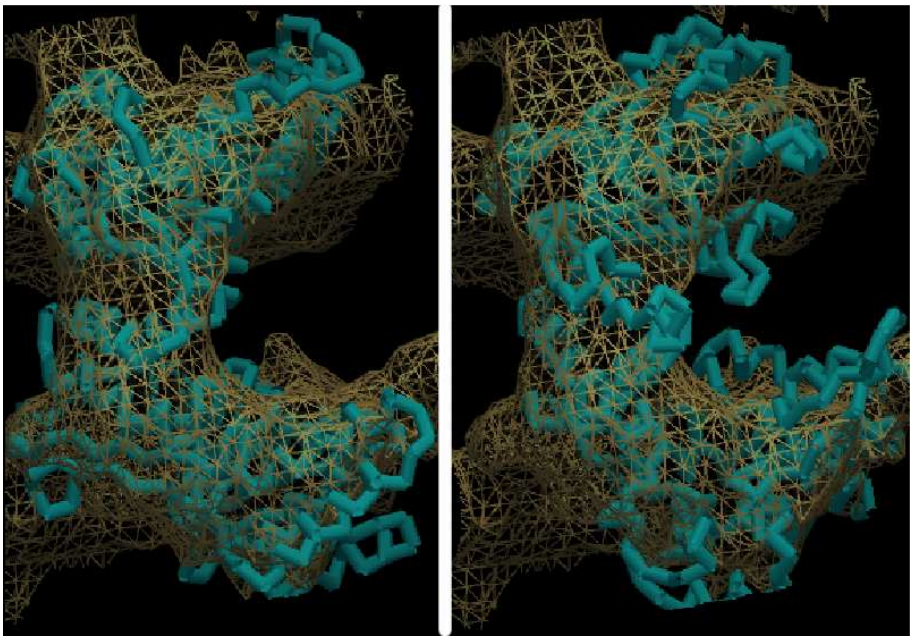

FIGURE 4. Fitting solution of the crystal structure of GroEl into a cryo-negative stain EM reconstruction. The correct solution (left) has a CC=89.6% while the upside-down solution (right) has a CC=74.0%. The figure is a snapshot of UROX.

in the graphics.

A user-friendly graphical interface is provided, with a variety of options for the fitting. Additionally, regions of the map can be hidden from the display to simplify the visualization while effectively using the whole EM map (or a substantial part of it) in the calculations. This combines advantages of real-space visualization and reciprocal-space fast computation.

Future developments are under way, notably to include interactive Normal-Mode flexible fitting (*12*). As the main programs for the graphical interface are written in a modular way using Python, additional user scripts can be easily incorporated. The UROX software package is freely available at `http://mem.ibs.fr/UROX`. This site also provides detailed installation instructions including a user's manual and several solved examples.

ACKNOWLEDGEMENTS

We are grateful to Jean Lepault for providing the VP6 reconstructions. We thank Leandro Estrozi and Stefano Trapani for critical reading of the manuscript. XS was supported by a Marie Curie International Reintegation Grant (IRG-021715).

REFERENCES

(1) M. G. Rossmann, M. C. Morais, P. G. Leiman, W. Zhang, Combining X-ray crystallography and electron microscopy, Structure 13 (3) (2005) 355–362.

(2) N. Volkmann, D. Hanein, Docking of atomic models into reconstructions from electron microscopy, Methods Enzymol. 374 (2003) 204–225.

(3) P. L. Stewart, S. D. Fuller, R. M. Burnett, Difference imaging of adenovirus: bridging the resolution gap between X-ray crystallography and electron microscopy, EMBO J 12 (7) (1993) 2589–2599, comparative Study.

(4) G. J. Wang, C. Porta, Z. G. Chen, T. S. Baker, J. E. Johnson, Identification of a Fab interaction footprint site on an icosahedral virus by cryoelectron microscopy and X-ray crystallography, Nature 355 (6357) (1992) 275–278.

(5) N. Volkmann, D. Hanein, Quantitative fitting of atomic models into observed densities derived by electron microscopy, J. Struct. Biol. 125 (2-3) (1999) 176–184.

(6) A. M. Roseman, Docking structures of domains into maps from cryo-electron microscopy using local correlation, Acta Cryst. D56 (10) (2000) 1332–1340.

(7) M. G. Rossmann, Fitting atomic models into electron-microscopy maps, Acta Cryst. D56 (10) (2000) 1341–1349.

(8) W. Jiang, M. L. Baker, S. J. Ludtke, W. Chiu, Bridging the information gap: computational tools for intermediate resolution structure interpretation, J. Mol. Biol. 308 (5) (2001) 1033–1044.

(9) W. Wriggers, R. A. Milligan, J. A. McCammon, Situs: A package for docking crystal structures into low-resolution maps from electron microscopy, J. Struct. Biol. 125 (2-3) (1999) 185–195.

(10) H. Ceulemans, R. B. Russell, Fast fitting of atomic structures to low-resolution electron density maps by surface overlap maximization, J. Mol. Biol. 338 (4) (2004) 783–793.

(11) J. Navaza, J. Lepault, F. A. Rey, C. Alvarez-Rua, J. Borge, On the fitting of model electron densities into EM reconstructions: a reciprocal-space formulation, Acta Cryst. D58 (10) (2002) 1820–1825.

(12) K. Suhre, J. Navaza, Y. Sanejouand, NORMA: a tool for flexible fitting of high-resolution protein structures into low-resolution electron-microscopy-derived density maps, Acta Cryst. D62 (9) (2006) 1098–1100.

(13) K. Hinsen, N. Reuter, J. Navaza, D. L. Stokes, J. J. Lacapere, Normal mode based fitting of atomic structure into electron density maps: application to SR Ca-ATPase, Biophys. J.

(14) M. Delarue, P. Dumas, On the use of low-frequency normal modes to enforce collective movements in refining macromolecular structural models, Proc. Natl. Acad. Sci. U. S. A. 101 (18) (2004) 6957–6962.

(15) T. A. Jones, Z.-Y. Zou, S. W. Cowan, M. Kjeldgaard, Improved methods for the building of protein models in electron density maps and the location of errors in these models, Acta Cryst. A47 (1991) 110–119.

(16) P. Emsley, K. Cowtan, Coot: Model-building tools for molecular graphics, Acta Cryst. D60 (2004) 2126–2132.

(17) W. L. DeLano, The pymol molecular graphics system, http://www.pymol.org, W.L. DeLano Scientific, San Carlos, CA, USA. (2002).

(18) J. Navaza, On the computation of structure factors by FFT techniques, Acta Cryst. A58 (6) (2002) 568–573.

(19) F. Fabiola, M. S. Chapman, Fitting of high-resolution structures into electron microscopy reconstruction images, Structure 13 (3) (2005) 389–400.

(20) E. E. Castellano, G. Oliva, J. Navaza, Fast rigid-body refinement for molecular-replacement

techniques, J. Appl. Cryst 25 (1992) 281–284.

(21) D. L. Gibbons, I. Erk, B. Reilly, J. Navaza, M. Kielian, F. A. Rey, J. Lepault, Visualization of the target-membrane-inserted fusion protein of Semliki Forest virus by combined electron microscopy and crystallography, Cell 114 (5) (2003) 573–583.

(22) C. Chevalier, M. Galloux, J. Pous, C. Henry, J. Denis, B. Da Costa, J. Navaza, J. Lepault, B. Delmas, Structural peptides of a non-enveloped virus are involved in assembly and membrane translocation, J. Virol. 79 (19) (2005) 12253–12263.

(23) A. Krebs, K. N. Goldie, A. Hoenger, Complex formation with kinesin motor domains affects the structure of microtubules, J. Mol. Biol. 335 (1) (2004) 139–153.

(24) F. Coulibaly, C. Chevalier, I. Gutsche, J. Pous, J. Navaza, S. Bressanelli, B. Delmas, F. A. Rey, The birnavirus crystal structure reveals structural relationships among icosahedral viruses, Cell 120 (6) (2005) 761–772.

(25) C. K. Kennaway, J. L. P. Benesch, U. Gohlke, L. Wang, C. V. Robinson, E. V. Orlova, H. R. Saibil, N. H. Keep, Dodecameric structure of the small heat shock protein Acr1 from Mycobacterium tuberculosis, J. Biol. Chem. 280 (39) (2005) 33419–33425.

(26) A. Krebs, K. N. Goldie, A. Hoenger, Structural rearrangements in tubulin following microtubule formation, EMBO Rep. 6 (3) (2005) 227–232.

(27) C. M. S. Fabry, M. Rosa-Calatrava, J. F. Conway, C. Zubieta, S. Cusack, R. W. H. Ruigrok, G. Schoehn, A quasi-atomic model of human adenovirus type 5 capsid, EMBO J. 24 (9) (2005) 1645–1654.

(28) E. J. Evans, R. M. Esnouf, R. Manso-Sancho, R. J. C. Gilbert, J. R. James, C. Yu, J. A. Fennelly, C. Vowles, T. Hanke, B. Walse, T. Hunig, P. Sorensen, D. I. Stuart, S. J. Davis, Crystal structure of a soluble CD28-Fab complex, Nat. Immunol. 6 (3) (2005) 271–279.

(29) J. Pous, C. Chevalier, M. Ouldali, J. Navaza, B. Delmas, J. Lepault, Structure of birnavirus-like particles determined by combined electron cryomicroscopy and X-ray crystallography, J. Gen. Virol. 86 (8) (2005) 2339–2346.

(30) A. L. Okorokov, M. B. Sherman, C. Plisson, V. Grinkevich, K. Sigmundsson, G. Selivanova, J. Milner, E. V. Orlova, The structure of p53 tumour suppressor protein reveals the basis for

its functional plasticity, EMBO J. 25 (21) (2006) 5191–5200.

(31) D. K. Clare, P. J. Bakkes, H. van Heerikhuizen, S. M. van der Vies, H. R. Saibil, An expanded protein folding cage in the GroEL-gp31 complex, J. Mol. Biol. 358 (3) (2006) 905–911.

(32) P. Fuschiotti, G. Schoehn, P. Fender, C. M. S. Fabry, E. A. Hewat, J. Chroboczek, R. W. H. Ruigrok, J. F. Conway, Structure of the dodecahedral penton particle from human adenovirus type 3, J. Mol. Biol. 356 (2) (2006) 510–520.

(33) S. L. Greig, J. A. Berriman, J. A. O'Brien, J. A. Taylor, A. R. Bellamy, M. J. Yeager, A. K. Mitra, Structural determinants of rotavirus subgroup specificity mapped by cryo-electron microscopy, J. Mol. Biol. 356 (1) (2006) 209–221.

(34) N. A. Ranson, D. K. Clare, G. W. Farr, D. Houldershaw, A. L. Horwich, H. R. Saibil, Allosteric signaling of ATP hydrolysis in GroEL-GroES complexes, Nat. Struct. Mol. Biol. 13 (2) (2006) 147–152.

(35) G. Schoehn, F. M. D. Vellieux, D. M. A., V. Receveur-Brechot, C. M. S. Fabry, R. W. H. Ruigrok, C. Ebel, A. Roussel, B. Franzetti, An archaeal peptidase assembles into two different quaternary structures: A tetrahedron and a giant octahedron, J. Biol. Chem. 281 (47) (2006) 36327–36337.

(36) O. Namy, S. J. Moran, D. I. Stuart, R. J. C. Gilbert, I. Brierley, A mechanical explanation of RNA pseudoknot function in programmed ribosomal frameshifting, Nature 441 (7090) (2006) 244–247.

(37) E. Neumann, I. Garcia-Saez, S. DeBonis, R. H. Wade, F. Kozielski, J. F. Conway, Human kinetochore-associated kinesin CENP-E visualized at 17 A resolution bound to microtubules, J. Mol. Biol. 362 (2) (2006) 203–211.

(38) G. Zanetti, J. A. G. Briggs, K. Grunewald, Q. J. Sattentau, S. D. Fuller, Cryo-electron tomographic structure of an immunodeficiency virus envelope complex in situ, PLoS Pathog. 2 (8) (2006) e83.

(39) H. E. White, E. V. Orlova, S. Chen, L. Wang, A. Ignatiou, B. Gowen, T. Stromer, T. M. Franzmann, M. Haslbeck, J. Buchner, H. R. Saibil, Multiple distinct assemblies reveal conformational flexibility in the small heat shock protein Hsp26, Structure 14 (7) (2006) 1197–

1204.

(40) K. Wyatt, H. E. White, L. Wang, O. A. Bateman, C. Slingsby, E. V. Orlova, G. Wistow, Lengsin is a survivor of an ancient family of class I glutamine synthetases re-engineered by evolution for a role in the vertebrate lens, Structure 14 (12) (2006) 1823–1834.

(41) R. J. C. Gilbert, Y. Gordiyenko, T. von der Haar, A. F.-P. Sonnen, G. Hofmann, M. Nardelli, D. I. Stuart, J. E. G. McCarthy, Reconfiguration of yeast 40S ribosomal subunit domains by the translation initiation multifactor complex, Proc. Natl. Acad. Sci. U. S. A. 104 (14) (2007) 5788–5793.

(42) D. Luque, I. Saugar, J. Rodriguez, N. Verdaguer, D. Garriga, C. Martin, J. Velazquez-Muriel, B. Trus, J. Carrascosa, J. Caston, Infectious Bursal Disease Virus capsid assembly and maturation by structural rearrangements of a transient molecular switch, J. Virol.

(43) R. Matadeen, W.-C. Hon, J. K. Heath, E. Y. Jones, S. Fuller, The dynamics of signal triggering in a gp130-receptor complex, Structure 15 (4) (2007) 441–448.

(44) VTK, The Visualization ToolKit: www.vtk.org.

(45) J. Lepault, I. Petitpas, I. Erk, J. Navaza, D. Bigot, M. Dona, P. Vachette, J. Cohen, F. A. Rey, Structural polymorphism of the major capsid protein of rotavirus, EMBO J. 20 (7) (2001) 1498–1507.

(46) V. I. Burdina, Symmetry of the rotation function, Soviet Phys. Crystallogr. 15 (1971) 545–550.

(47) E. E. Lattman, Optimal sampling of the rotation function, Acta Cryst. B28 (4) (1972) 1065–1068.

(48) M. Mathieu, I. Petitpas, J. Navaza, J. Lepault, E. Kohli, P. Pothier, B. V. Prasad, J. Cohen, F. A. Rey, Atomic structure of the major capsid protein of rotavirus: implications for the architecture of the virion, EMBO J. 20 (7) (2001) 1485–1497.

(49) S. Libersou, X. Siebert, M. Ouldali, L. F. Estrozi, A. Charpilienne, P. Garnier, D. Poncet, J. Lepault, Symmetry mismatch within the rotavirus capsid layers: a regulatory switch of the transcription activity of rotavirus, submitted.

(50) D. L. Caspar, A. Klug, Physical principles in the construction of regular viruses., Cold Spring

Harb. Symp. Quant. Biol. 27 (1962) 1–24.

(51) J. E. Ludert, F. Gil, F. Liprandi, J. Esparza, The structure of the rotavirus inner capsid studied by electron microscopy of chemically disrupted particles, J. Gen. Virol. 67 (8) (1986) 1721–1725.

(52) A. Roseto, J. Escaig, E. Delain, J. Cohen, R. Scherrer, Structure of rotaviruses as studied by the freeze-drying technique, Virology 98 (2) (1979) 471–475.

(53) K. Braig, P. Adams, A. Brunger, Conformational variability in the refined structure of the chaperonin groel at 2.8 a resolution., Nat. Struct. Biol. 2 (12) (1995) 1083–1094.

(54) S. De Carlo, C. El-Bez, C. Alvarez-Rua, J. Borge, J. Dubochet, Cryo-negative staining reduces electron-beam sensitivity of vitrified biological particles, J. Struct. Biol. 138 (3) (2002) 216–226.