Published in final edited form as:

J Struct Biol. 2019 March 01; 205(3): 53–58. doi:10.1016/j.jsb.2019.01.006.

Enhancement of tomogram interpretability using the locked self-rotation function

Andrei Fokine^{a,*}, Baldeep Khare^a, Yingyuan Sun^a, and Michael G. Rossmann^{a,*}

^aDepartment of Biological Sciences, Purdue University, West Lafayette, IN 47907, United States

Abstract

The interpretation of cryo-electron tomograms of macromolecular complexes can be difficult because of the large amount of noise and because of the missing wedge effect. Here it is shown how the presence of rotational symmetry in a sample can be utilized to enhance the quality of a tomographic analysis. The orientation of symmetry axes in a sub-tomogram can be determined using a locked self-rotation function. Given this knowledge, the sub-tomogram density can then be averaged to improve its interpretability. Sub-tomograms of the icosahedral bacteriophage phiX174 are used to demonstrate the procedure.

Graphical Abstract

Keywords

cryo-electron tomography; locked self-rotation function; point group symmetry; macromolecular complexes; virus structure

1. Introduction

Cryo-electron tomography (cryo-ET) is a tool for studying structures of biological macromolecular complexes (Beck and Baumeister, 2016; Lucic et al., 2013) which may have

*Corresponding authors at: Department of Biological Sciences, Purdue University, 240 S. Martin Jischke Drive, West Lafayette, IN 47907, USA. mr@purdue.edu (M.G. Rossmann) and afokine@purdue.edu (A. Fokine). Telephone: (765) 494-4911; FAX: (765) 496-1189.

Author contributions:

A.F., B.K., and M.G.R. developed the method. Y.S. collected phiX174 tomograms. A.F. performed the averaging calculations. A.F. and M.G.R. wrote the paper. All authors reviewed and approved the manuscript.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

differences between individual particles. Cryo-electron tomography data are collected as a series of projection images of the sample acquired at different tilt angles. Tomographic tilt series are usually collected using low electron doses, to minimize radiation damage, resulting in a large amount of noise in the reconstructed three-dimensional maps. Furthermore, in a typical cryo-ET experiment, projection images cannot be collected for the full range of tilt angles, because at high tilts the electron path through the sample becomes long, resulting in a large number of electrons being scattered inelastically or lost. Furthermore, projections at a high-tilt angle can be obscured by the sample support. Therefore, in cryo-ET experiments, the tilt angle is usually limited to be within the range of about -70° to $+70^\circ$. Each projection image, recorded in a tomographic tilt series, corresponds to a central plane in Fourier space, perpendicular to the direction of the electron beam. Thus, the lack of high-tilt images results in the lack of a wedge-shaped region of data in reciprocal space, known as “the missing wedge”. The missing wedge results in distortion of the reconstructed map, which lacks resolution in the direction of the electron beam. As a result the map features appear to be “stretched” in this direction.

Because of noise and the missing-wedge, raw tomograms are usually difficult to interpret (Fig. 1). Therefore, averaging of structurally similar regions of a tomographic map should be beneficial to increase the signal to noise ratio and to reduce the missing wedge effect. If a biological complex has symmetry, such as the icosahedral symmetry of many viruses, the map of an individual particle (subtomogram) can be improved by symmetry averaging. Particles extracted from tomograms have random orientations. Therefore, to calculate such symmetry-averaged maps, the orientations of the symmetry axes and the position of the particles' centers will have to be determined prior to averaging. Real space tomogram-averaging approaches for viruses in cells had been previously reported. These employ icosahedral harmonics to “lock” the symmetry and maximum likelihood analysis to determine the particle center positions and orientations (Fu et al., 2010; Wang et al., 2011).

The self-rotation function is a protein crystallographic tool (Rossmann and Blow, 1962; Tong and Rossmann, 1990) which can be used to find the orientations of non-crystallographic symmetry axes. Here we use this technique to determine rotation symmetry operators in maps of individual particles, extracted from tomograms. In the rotation function calculations a volume, extracted from a tomogram, can be treated as a unit cell of an imaginary crystal with P1 symmetry. This map can be Fourier transformed to calculate Fourier coefficients $F(\vec{h})$ that correspond to the structure factors in a crystallographic analysis.

In crystallography the self-rotation function is usually employed when the structure factor phases are missing and only the amplitudes are available. Therefore the standard crystallographic self-rotation function was developed based on the Patterson function which can be calculated using only structure factor amplitudes (Rossmann and Blow, 1962). However, in the tomographic analysis presented here, the structure factor phases are available and, therefore, a more powerful real-space self-rotation function can be used (Tong and Rossmann, 1990). This function is defined as

$$\text{RF}([C], \vec{S}) = \int_V \rho(\vec{r}) \rho([C](\vec{r} - \vec{S}) + \vec{S}) dV \quad (1)$$

where $\rho(\vec{r})$ is the cryo-EM density map (Coulomb potential map), and $\rho([C](\vec{r} - \vec{S}) + \vec{S})$ represents the same map rotated around the center of the particle $\vec{S}$ using the rotation operator $[C]$ (which can be expressed as a 3×3 matrix or as a set of three angles). The integration is performed over a spherical volume V which has the center $\vec{S}$ at the center of the particle. The radius R of this sphere of integration is usually chosen such that it would encompass the whole particle. The electron density superimposes well onto a rotated version of itself when the rotation corresponds to a symmetry operator of the particle. Thus the rotation function should ideally have a maximum when $[C]$ corresponds to a symmetry operator of the studied object. It should be noted that contrary to the standard Patterson-based rotation functions, which are independent of the position of the particle's center, the real-space rotation function requires the center $\vec{S}$ of the particle to be known (or assumed).

The real-space self-rotation function can be calculated using the structure factors in the expression

$$\text{RF}([C]) = \sum_{\vec{p}} F(\vec{p}) e^{-2\pi i \vec{p} \cdot \vec{S}} \sum_{\vec{h}} F(\vec{h}) G_{\text{hp}} e^{-2\pi i \vec{h} \cdot \vec{S}} \quad (2)$$

where $G_{\text{hp}} = \frac{3(\sin 2\pi HR - 2\pi HR \cos 2\pi HR)}{(2\pi HR)^3}$ represents the Fourier transform of the sphere of integration with radius R (Rossmann and Blow, 1962). H is the length of the reciprocal space vector $\vec{h} + \vec{p}[C]$. The G_{hp} function has the maximum value of 1, when $H = |\vec{h} + \vec{p}[C]| = 0$, and it quickly approaches zero and then oscillates around zero as H increases (for example $G_{\text{hp}} < 0.086$ when $HR > 0.725$). Therefore, terms in Equation 2 for which HR is greater than a certain cutoff may be neglected to speed up the calculation.

Rotation functions have proven to be useful for determining orientations of non-crystallographic symmetry axes in macromolecular crystals. However in the case of particles extracted from tomograms, the correct rotation function peak may be hidden in noise. If the particle has a point group symmetry, the power of the rotation search can be enhanced by searching for all symmetry axes simultaneously as is possible with the locked self-rotation function (LSRF) (Rossmann et al., 1972; Tong and Rossmann, 1990). The robustness of the LSRF has been proven in crystallographic structure determinations using low-completeness diffraction datasets (Plevka et al., 2012).

The LSRF determines the orientation of a particle relative to a pre-defined “standard” orientation. An example of a standard orientation would be the orientation of an icosahedral particle, with its 2-fold axes along the coordinate axes X , Y and Z , and a 5-fold axis in the XZ plane at 31.7° to Z .

Let $[I_i]$ ($i=1, \dots, N$) be symmetry operators of the point group defined in the standard orientation (the first operator being the identity matrix). The matrix $[E]$ defines the rotation required to rotate the assumed point group from the standard orientation to a new orientation in the unit cell. The point group symmetry operators are then expressed as $[\alpha_i] [E][I_i][E]^{-1}$. Let RF_i be the ordinary rotation function values obtained when using the matrices $[\alpha_i]$. The RF_i values should be large if $[E]$ corresponds to the correct rotation of the point group into the unit cell. The LSRF is then defined as

$$LSRF([E]) = \frac{1}{N-1} \sum_2^N RF_i \quad (3)$$

The LSRF samples the rotation function values for all $N - 1$ symmetry operators. Therefore the signal is expected to be about $\sqrt{N-1}$ times stronger compared to the ordinary rotation function values.

Here we present the rotation function-based averaging method tested on sub-tomograms of the icosahedral bacteriophage phiX174 (McKenna et al., 1992).

2. Methods

2.1. Algorithm for sub-tomogram averaging

The sub-tomogram averaging algorithm includes the following steps:

1. Maps (sub-tomograms) corresponding to individual particles, are extracted from reconstructed tomograms. The centers of the extracted maps correspond approximately to the centers of the particles.
2. Each extracted map is Fourier transformed to calculate a set of structure factors.
3. The real-space self-rotation functions are calculated for each individual particle using the program GLRF (Tong and Rossmann, 1997; Tong and Rossmann, 1990) assuming different positions of the particle's center. The array of possible centers may be defined by a grid which covers the central region of the map.
4. The LSRF is calculated from each self-rotation function by averaging values at appropriate grid points (as shown in Equation 3). The local symmetry operators, used in this calculation, are defined according to the assumed point group symmetry of the particle given its standard orientation.
5. The highest value among all the calculated LSRF values is selected.
6. The rotation operator and the particle's center corresponding the highest LSRF value is used to calculate the symmetry-averaged map of the particle.
7. Finally, the symmetry-averaged maps of individual particles are averaged together, rejecting particles that are outliers.

To perform steps 2 and 4 – 7 we developed programs which use the EMAN2 software library (Tang et al., 2007). These programs are available from the authors on request. This

procedure represents a six-dimensional search because the value of the LSRF is calculated for each combination of the particle's center and rotation.

2.2. Preparation of the bacteriophage phiX174 sample and tomography data collection

Bacteriophage phiX174 has an icosahedral capsid with a diameter of about 320 Å (Fig. 2) (McKenna et al., 1992). The capsid has characteristic spikes at each 5-fold vertex consisting of G protein pentamers. The phiX174 sample was purified as previously described (Sun et al., 2017). PhiX174 solution with a concentration of 10^{12} pfu/ml was mixed with BSA-conjugated 10 nm gold particle solution (Aurion, Wageningen, Netherlands) in a 5:1 volume ratio. This mixture (3 µl) was applied to C-flat CF-2/2-4C grids (Electron Microscopy Sciences), blotted, and then rapidly frozen in liquid ethane. Tilt-series data were collected using an FEI Titan Krios microscope, operated at 300 kV, and equipped with a K2 Gatan Summit detector. The SerialEM software (Mastronarde, 2005) was used to collect data at appropriate grid positions. A magnification of 11000 was used with a pixel size of 1.3 Å. Tilt series were collected from -60° to $+60^\circ$ in 3° angular steps with a defocus of $-3\text{ }\mu\text{m}$. The tilt axis was named Z and the direction of the electron beam Y^1 . The total electron dose applied to the grid was $\sim 71\text{ e}^-/\text{\AA}^2$. The tilt series were aligned using the gold fiducial markers and the tomographic reconstruction was performed using the IMOD software (Kremer et al., 1996; Mastronarde, 1997) (Supplementary Fig. 1). Individual phiX174 particles were extracted from the tomograms using the program Dynamo (Castaño-Diez et al., 2017) as cubic boxes with edges of 416 Å. The manually selected centers of the boxes corresponded approximately to the centers of the particles. The maps of the extracted particles were four times binned, so that the resultant pixel size was 5.2 Å. Ten extracted particles were used in the calculations.

3. Results

The structure factors were calculated for each individual particle by a fast Fourier transform of the map. The structure factors corresponding to the missing wedge were removed from further calculations.

A set of real-space self-rotation functions was calculated for each particle using the program GLRF (Tong and Rossmann, 1997; Tong and Rossmann, 1990) assuming different positions of the particle's center. A step size of 10.4 Å (2 pixels) was used to test different centers systematically. The search was limited to the range of $\pm 20.8\text{ }\text{\AA}$ in the X and Z directions, and to $\pm 31.2\text{ }\text{\AA}$ in the Y direction. The larger Y range was used because the resolution along the Y axis is lower which lowers the accuracy of particle center picking along this direction. Only structure factors in the resolution range from infinity to 25 Å were used in the LSRF calculations, because we assumed that the higher-resolution structure factors could be largely affected by noise. Also, it may sometimes be beneficial to remove very low resolution Fourier terms from the LSRF calculation because they mainly describe the spherical shape of the virus particle and are insensitive to rotation. The radius of integration was set to 170 Å, which is slightly larger than the $\sim 160\text{ }\text{\AA}$ radius of the particle. The rotation

¹Our software also allows the more common axis setting with the direction of the beam named Z and the tilt axis Y.

function was calculated using Eulerian angles (Rossmann and Blow, 1962), where θ_1 gives the rotation about the Z axis, θ_2 gives the rotation about new Y, and θ_3 gives the rotation about new Z. The rotation functions were calculated for the entire Eulerian space with a 4° angular step. Taking into account the ~ 160 Å radius of the phiX174 particle this angular step results in about $4 \times 160 \times \pi / 180 = \sim 11$ Å step on the capsid surface, which would be sufficiently small for ~ 22 Å resolution.

It should be noted that to speed up calculations it is sufficient to calculate the rotation functions only for a few sections of constant angle kappa in the polar angle space, with the kappa values corresponding to the rotation angles around the symmetry axes.

The LSRFs were calculated from each self-rotation function in the angular range of $0 - 90^\circ$ for θ_1 , θ_2 , and θ_3 . This range included more than one icosahedral asymmetric unit of the Eulerian rotation space. The angular step of the LSRF calculation was 4° .

For each particle the center and orientation were determined based on the highest value obtained in the LSRF calculations (Fig. 3, Table 1). Then each particle was rotated to the standard orientation and icosahedrally-averaged maps of individual particles were calculated.

The averaged map of each particle showed a capsid that had the characteristic 5-fold spikes associated with the phiX174 structure (Fig. 4), showing that the LSRF analyses were able to determine the orientations and centers of each particle.

At this stage of the analysis the symmetry-averaged maps can be compared with each other, in order to remove those particles whose structure is different from other particles. If more than one structural conformation is present among the particles, then the structures can be classified using, for example, hierarchical clustering and different clusters can be averaged separately. For phiX174 the correlation coefficients between pairs of averaged maps were in the range of 0.8 to 0.9, showing good agreement between particles' structures. Therefore the inter-particle average structure was calculated by averaging of the icosahedrally-averaged maps of all the particles (Fig. 5). To estimate the resolution of the overall averaged structure, the maps were separated into two equal sets, and the averaged map was calculated for each of these half-datasets. The Fourier shell correlation (FSC) between these two averaged maps was calculated and the resolution was estimated to be 26 Å using the FSC cutoff of 0.143 (Supplementary Fig. 2).

Although the original crystallographic determination of the phiX174 structure (McKenna et al., 1992) had a resolution of 3 Å it required numerous crystals with each crystal containing about 10^{12} particles. Another structure determination technique would be single particle cryo-electron microscopy (cryo-EM) as was used for determining the structure of the closely homologous ST1 phage (Chen et al., manuscript in preparation) that accomplished 2.3 Å resolution using 39,000 particles. In comparison only 10 icosahedral particles were needed for the present analysis to obtain a resolution of 26 Å showing that the tomographic method can be reasonably successful when only very small sample is available or the available sample is rather heterogeneous.

The capability of the LSRF to determine particles' orientations was further tested assuming lower-symmetry point group D5 for the phiX174 particles. D5 has 2-fold axes perpendicular to a 5-fold axis. This point group describes, for example, the symmetry of the major capsid protein shell in the prolate phage capsids, like T4 (Fokine et al., 2004).

The symmetry-averaged map for each particle was calculated with the same procedure as described above, but the icosahedral symmetry operators were replaced by D5 operators in the LSRF calculations and in the averaging procedure (Fig. 6). As expected, the D5 symmetry-averaged maps are noisier, compared to the icosahedrally-averaged maps, because the number of symmetry operators in D5 is less than those in the icosahedral point group (10 and 60 operators, respectively). Nevertheless, the D5-symmetry averaged maps are more interpretable than the unaveraged subtomograms. The averaged maps of 8 out of 10 particles (Fig. 6) and the corresponding inter-particle averaged map (Fig. 7) showed recognizable 5-fold spikes, characteristic of the phiX174 structure.

3.2. Conclusion

This study demonstrated that the real-space LSRF can be used to determine orientations of symmetric particles extracted from cryo-EM tomograms of biological samples. The advantage of the self-rotation function is that it does not require any model (or reference density), and therefore the results are not biased by the model used. The averaged maps of individual particles can be used to detect and filter out particles that are outliers because of distortions or lack of homogeneity. The symmetry-averaged maps of individual particles, placed in the standard orientation, were averaged together to produce the final inter-particle averaged map. The particle orientations and centers could have been further refined by aligning the initial maps to the inter-particle averaged map. Then the new inter-particle averaged map can be re-calculated in an iterative procedure for map refinement. The tomographic method described here can be used for any biological assembly that has some symmetry and can be used even if only one particle is available.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements:

We thank Thomas Klose for help with the tomographic data collection. We also thank Sheryl Kelly for help with the preparation of the manuscript. This work was conducted relative to grants awarded to M.G.R. by the National Science Foundation (MCB-1515260) and the National Institutes of Health (AI076331 and AI081726).

References

- Beck M, Baumeister W, 2016 Cryo-electron tomography: can it reveal the molecular sociology of cells in atomic detail? *Trends Cell Biol* 26, 825–837. [PubMed: 27671779]
- Castaño-Diez D, Kudryashev M, Stahlberg H, 2017 Dynamo Catalogue: Geometrical tools and data management for particle picking in subtomogram averaging of cryo-electron tomograms. *J. Struct. Biol* 197, 135–144. [PubMed: 27288866]
- Fokine A, Chipman PR, Leiman PG, Mesyanzhinov VV, Rao VB, Rossmann MG, 2004 Molecular architecture of the prolate head of bacteriophage T4. *Proc. Natl. Acad. Sci. U.S.A* 101, 6003–6008. [PubMed: 15071181]

- Fu CY, Wang K, Gan L, Lanman J, Khayat R, Young MJ, Jensen GJ, Doerschuk PC, Johnson JE, 2010 *In vivo* assembly of an archaeal virus studied with whole-cell electron cryotomography. *Structure* 18, 1579–1586. [PubMed: 21134637]
- Kremer JR, Mastronarde DN, McIntosh JR, 1996 Computer visualization of three-dimensional image data using IMOD. *J. Struct. Biol* 116, 71–76. [PubMed: 8742726]
- Lucic V, Rigort A, Baumeister W, 2013 Cryo-electron tomography: the challenge of doing structural biology *in situ*. *J. Cell Biol* 202, 407–419. [PubMed: 23918936]
- Mastronarde DN, 1997 Dual-axis tomography: an approach with alignment methods that preserve resolution. *J. Struct. Biol* 120, 343–352. [PubMed: 9441937]
- Mastronarde DN, 2005 Automated electron microscope tomography using robust prediction of specimen movements. *J. Struct. Biol* 152, 36–51. [PubMed: 16182563]
- McKenna R, Xia D, Willingmann P, Ilag LL, Krishnaswamy S, Rossmann MG, Olson NH, Baker TS, Incardona NL, 1992 Atomic structure of single-stranded DNA bacteriophage ΦX174 and its functional implications. *Nature* 355, 137–143. [PubMed: 1370343]
- Plevka P, Perera R, Cardosa J, Kuhn RJ, Rossmann MG, 2012 Crystal structure of human enterovirus 71. *Science* 336, 1274. [PubMed: 22383808]
- Rossmann MG, Blow DM, 1962 The detection of sub-units within the crystallographic asymmetric unit. *Acta Crystallogr* 15, 24–31.
- Rossmann MG, Ford GC, Watson HC, Banaszak LJ, 1972 Molecular symmetry of glyceraldehyde-3-phosphate dehydrogenase. *J. Mol. Biol* 64, 237–245. [PubMed: 5015398]
- Sun Y, Roznowski AP, Tokuda JM, Klose T, Mauney A, Pollack L, Fane BA, Rossmann MG, 2017 Structural changes of tailless bacteriophage ΦX174 during penetration of bacterial cell walls. *Proc. Natl. Acad. Sci. U.S.A* 114, 13708–13713. [PubMed: 29229840]
- Tang G, Peng L, Baldwin PR, Mann DS, Jiang W, Rees I, Ludtke SJ, 2007 EMAN2: an extensible image processing suite for electron microscopy. *J. Struct. Biol* 157, 38–46. [PubMed: 16859925]
- Tong L, Rossmann MG, 1997 Rotation function calculations with GLRF program. *Methods Enzymol* 276, 594–611.
- Tong LA, Rossmann MG, 1990 The locked rotation function. *Acta Crystallogr A* 46 (Pt 10), 783–792.
- Wang K, Fu CY, Khayat R, Doerschuk PC, Johnson JE, 2011 *In vivo* virus structures: simultaneous classification, resolution enhancement, and noise reduction in whole-cell electron tomography. *J. Struct. Biol* 174, 425–433. [PubMed: 21396453]

HIGHLIGHTS

- Self-rotation function can be used to find orientations of particles in tomograms
- Once the orientations have been determined the symmetry-averaged maps can be produced
- The symmetry-averaged maps may be helpful to reject particles that are outliers
- Finally, the inter-particle averaged map can be calculated to improve interpretability

Fig. 1.
Surface representation of a sub-tomogram of one phiX174 particle. The outline of the particle is barely visible because of noise and stretching of the map in the vertical direction because of the missing wedge.

Fig. 2. Structure of bacteriophage phiX174 determined by X-ray crystallography (McKenna et al., 1992). The major coat protein F is shown in blue and the pentameric vertex protein G is shown in green.

Fig. 3.

Section of the LSRF calculated for a subtomogram of one phiX174 particle. The rotation function was calculated in the Eulerian angular space (θ_1 – rotation around Z, θ_2 – rotation around new Y, θ_3 – rotation around new Z). This section corresponds to $\theta_2 = 72^\circ$. The peak, corresponding to the correct orientation of the particle is indicated by the black arrow.

Fig. 4. Icosahedrally-averaged maps of individual phiX174 particles. The views are perpendicular to one out of the six 5-fold axes in each capsid.

Fig. 5.
(A) The inter-particle averaged map calculated by averaging of the maps of the ten individual particles shown in Fig. 4. (B) The crystal structure of phiX174 fitted into the inter-particle averaged map. The vertex protein G is shown in green and the major capsid protein F is shown in gold.

Fig. 6. Symmetry-averaged maps of individual phiX174 particles obtained by using the D5 point group symmetry. The views are perpendicular to one out of the six 5-fold axes in each capsid.

Fig. 7.
The inter-particle averaged map calculated by averaging of the maps of the eight individual particles shown in Fig. 6.

Table 1.

Results of the locked self-rotation search using the icosahedral and D5 symmetry.

Particle number	The highest LSRF peak (in sigma units)	
	Search assuming the icosahedral symmetry	Search assuming the D5 symmetry
1	7.6	3.3
2	8.1	6.2
3	7.4	5.2
4	6.9	4.9
5	7.3	5.3
6	7.8	5.1
7	7.6	7.0
8	6.4	4.0
9	6.3	4.5
10	6.8	4.3