Journal of Structural Biology xxx (2016) xxx–xxx

Contents lists available at ScienceDirect

Journal of Structural Biology

journal homepage: www.elsevier.com/locate/yjsbi

Hierarchical detection and analysis of macromolecular complexes in cryo-electron tomograms using Pyto software

Vladan Lučić *, Rubén Fernández-Busnadiego, Ulrike Laugks, Wolfgang Baumeister

Max-Planck-Institute of Biochemistry, Am Klopferspitz 18, 82152 Martinsried, Germany

ARTICLE INFO

Article history:
Received 31 May 2016
Received in revised form 15 September 2016
Accepted 6 October 2016
Available online xxxx

Keywords:
Electron tomography
Image processing
Cryo
Segmentation
Synapses

ABSTRACT

Molecular complexes, arguably the basic units carrying cellular function, can be visualized directly in their native environment by cryo-electron tomography. Here we describe a procedure for the detection of small, pleomorphic membrane-bound molecular complexes in cryo-tomograms by a hierarchical connectivity segmentation. Validation on phantom and real data showed above 90% true positive rates. This segmentation procedure is implemented in the Pyto software package, together with methods for quantitative characterization and classification of complexes detected by our segmentation procedure and for statistical analysis between experimental conditions. Therefore, the methods presented provide a means for the detection and quantitative interpretation of structures captured in cryo-electron tomograms, as well as for the elucidation of their cellular function.

1. Introduction

Molecular complexes, arguably the basic units carrying cellular function, can be visualized directly in their native environment by cryo-electron tomography (cryo-ET) (Harapin et al., 2013; Koning and Koster, 2013; Lucic et al., 2013). Thanks to the faithful sample preservation and the direct imaging of proteins and lipids, these tomograms provide an accurate representation of heterogeneous cellular constituents such as macromolecular complexes. However, the detection of specific structures in comprehensive cellular pictures obtained by this method is aggravated by the crowded nature of the cell interior. Large molecular complexes, as well as those having distinct shape can be visually identified, but the majority of cellular complexes require detection by image processing methods. Furthermore, a quantitative assessment of structures of interest, and the comparison over different experimental conditions necessitates the use of automated detection methods.

Image processing methods are used for different cryo-ET post-processing tasks (reviewed in Fernandez, 2012). Image segmentation separates and partitions the relevant information that is represented by one or more segments (foreground) from the image background. Several segmentation software packages for cryo-tomograms were developed, but they generally failed to gain prominence in comparison to the manual segmentation (reviewed in Sandberg, 2007; Volkmann, 2010; Fernandez, 2012). Automated segmentation procedures for cryo-tomograms have the advantage that they avoid tedious and error-prone manual segmentation of noisy 3D images. Such procedures were developed for segmentation of more prominent cellular structures such as lipid membranes (Nguyen and Ji, 2008; Moussavi et al., 2010; Martinez-Sanchez et al., 2011; Martinez-Sanchez et al., 2014; Tasel et al., 2016) and cytoskeletal filaments (Rigort et al., 2012; Rusu et al., 2012).

Among the more general algorithms used for segmentation, various types of density thresholding and watershed are commonly used. While these were used successfully on cryo-tomograms (Volkmann, 2002; Cyrklaff et al., 2005), they have considerable limitations. Simple thresholding requires that the optimal threshold is determined, which is often not possible. Hysteresis thresholding solves this problem by using a threshold range, but it can not generate segments entirely composed of voxels that have values above the lower threshold (lower voxel values correspond to higher electron density). On the other hand, watershed segmentation takes all greyscale levels into account, but it often leads to oversegmentation, especially on noisy images (Soille, 2003).

Our main biological goal was to develop a procedure for detection and analysys of pleomorphic, membrane-bound molecular complexes, such as those that interlink synaptic vesicles or link them to the plasma membrane (termed connectors). To this end, we developed the hierarchical connectivity segmentation, an auto-

* Corresponding author.
E-mail address: vladan@biochem.mpg.de (V. Lučić).

http://dx.doi.org/10.1016/j.jsb.2016.10.004
1047-8477/

mated segmentation procedure for cryo-tomograms that detects molecular complexes (called segments in image processing) that contact larger, clearly detectable structures (boundaries). These boundaries are typically lipid membranes, but organelles and cytoskeletal filaments can also be used. This segmentation procedure combines thresholding at multiple levels with the detection of connected components and organizes the segments into a hierarchical tree. It is particularly suited for complexes that are bound to two different boundaries, such as opposing membranes. This procedure is implemented in the Pyto package written in Python programming language. This package also contains methods for determination of various properties of complexes, as well as those for the statistical analysis of these properties between datasets obtained under different experimental conditions. The validation and usefulness of these procedures are shown on cryo-tomograms of neuronal synapses, but the whole package is developed in the way that allows straightforward application to other biological systems. Here we first define the segmentation procedure, then present the Pyto software design and finally show the validation results obtained on both phantom and real datasets.

2. Material and methods

2.1. Sample preparation

Cerebrocortical synaptosomes were extracted from 6–8 week old male Wistar rats as described previously (Dunkley et al., 1988; Godino et al., 2007; Fernández-Busnadiego et al., 2013) in accordance with the procedures accepted by the Max Planck Institute for Biochemistry. In brief, anesthetized animals were sacrificed, and the cortex was extracted and homogenized in homogenization buffer (HB; 0.32 M sucrose, 50 mM EDTA, 20 mM DTT, and one tablet of Complete mini EDTA-free protease inhibitor cocktail (Roche; 10 ml, pH 7.4) with up to seven strokes at 700 rpm in a Teflon glass homogenizer. The homogenate was centrifuged for 2 min at 2000 g, and the pellet was resuspended in HB and centrifuged for another 2 min at 2 000 g. Supernatants from both centrifugations were combined and centrifuged for 12 min at 9 500 g. The pellet was resuspended in HB and loaded onto a three-step Percoll gradient (3%, 10%, and 23%; Sigma–Aldrich) in HB without protease inhibitor cocktail. The gradients were spun for 6 min at 25 000 g, and the material accumulated at the 10/23% interface was recovered and diluted to a final volume of 100 ml in Hepes-buffered medium (HBM; 140 mM NaCl, 5 mM KCl, 5 mM $NaHCO_3$, 1.2 mM Na_2HPO_4, 1 mM $MgCl_2$, 10 mM glucose, and 10 mM Hepes, pH 7.4). Percoll was removed by an additional washing step with HBM by centrifugation for 10 min at 22,000 g, and the pellet was resuspended in HBM and immediately used in the experiments. All steps were performed at 4 °C.

2.2. Vitrification, imaging and 3D reconstruction

For vitrification, a 3-μl drop of 10-nm colloidal gold (Sigma–Aldrich) was deposited on plasma-cleaned, holey carbon copper EM grids (Quantifoil) and allowed to dry. A 3-μl drop of synaptosomes was placed onto the grid, blotted with filter paper (GE Healthcare), and plunged into liquid ethane.

Three tomograms of synapses were used in this study (datasets CCD 1, CCD 2 and DDD). Imaging and reconstruction for dataset DDD was similar to that already presented for datasets CCD 1 and 2 (Fernández-Busnadiego et al., 2010). Specifically, tilt series were collected under a low dose acquisition scheme (Koster et al., 1997) using microscopes CM300 [Philips], TF30 Polara [FEI] (datasets CCD 1 and 2) and Titan Krios [FEI] (dataset DDD). All microscopes were equipped with a field emission gun operated at 300 kV, with a post-GIF energy filter (Gatan) operated in the zero-loss mode and with a computerized cryostage designed to maintain the specimen temperature <−150 °C. Images from the first two microscopes were recorded on a 2 k × 2 k charge-coupled device camera (Gatan) and from the third one on direct electron detector device (K2 Summit operated in the counting mode). Tilt series were typically recorded from −60° to 60° with a 2° angular increment. Pixel sizes were 0.68 nm (CM300), 0.66 nm (Polara) and 0.34 nm (Titan) at the specimen level. The defocus was set to −9μm for datasets CCD 1 and 2, while for the dataset DDD Volta phase-plate with nominal defocus of −1μm (Danev et al., 2014) was used. The total dose was kept <100$e^-/Å^2$. Tilt series were aligned using gold beads as fiducial markers, and 3D reconstructions were obtained by weighted back projection (WBP) using Imod (Kremer et al., 1996) and TOM toolbox (Nickell et al., 2005). During reconstruction, the projections were binned twice (final voxel size of 2.73, 2.64 and 1.37 nm) and low pass filtered at the post-binning Nyquist frequency.

Tomograms CCD 1 and 2 were subsequently denoised by anisotropic nonlinear diffusion (Frangakis and Hegerl, 2001) as implemented by (Fernandez and Li, 2003). The following parameters wee used in all cases: CED/EED balance 0.5, proportion of CED along 2nd eigenvector 0.75, proportion of smoothing 0, initial sigma 0.25, sigma for averaging 2, time interval 0.04, diffusion mode standard hybrid (0). Parameters K (constant for edge enhansing diffusion) and C (constant for coherence enhancing diffusion) were chosen as specified in the text.

2.3. Determination of boundaries

Boundaries required for the segmentation procedure comprised a portion of the presynaptic plasma membrane directly apposed to the postsynaptic side (the active zone membrane) and the synaptic vesicle membranes (Fig. 1 A2, B2 and C2). The active zone membrane and maximum-diameter profile (along the z axis) of each vesicle were manually traced using Amira (TGS) segmentation tools, so that the complete membranes, but no outside material were included. For each vesicle, the center was determined as the center of mas of the corresponding profile, the radius was calculated from the distance between the center and all the pixels on the profile edge, and the vesicles were then created by substituting each profile with a sphere having the previously determined center and radius (all steps implemented in Pyto). The resulting boundaries on individual tomographic slices are shown on the middle panels of Fig. 1 A3, B3 and C3. Alternatively, boundaries can be detected by automated tools (Fernandez, 2012). The comparison with vesicles segmented by manual tracing of all profiles confirmed that spheres provided a good approximation for vesicles as described previously (Harris and Sultan, 1995). In addition, the region where the segmentation is to be performed was outlined in Amira (TGS). This region does not need to be outlined precisely, but it should not include boundaries nor any other cellular structures such as organelles, larger vesicles or cytoskeletal filaments.

2.4. Pyto software design and implementation

2.4.1. Segments and contacts

Every segment created by the segmentation procedures presented here consists of mutually interconnected voxels. That is, each two voxels of a segment are connected via a sequence of neighboring (directly contacting) voxels belonging to the same segment, and no voxel from one segment can be connected to a voxel of any other segment. The choice of structuring element determines what constitutes a neighborhood (direct contact) of a voxel. Here, two (3D) voxels are considered to be in direct contact (neighborhood) if they share a face, that is a voxel (regarded as a

Please cite this article in press as: Lučić, V., et al. Hierarchical detection and analysis of macromolecular complexes in cryo-electron tomograms using Pyto software. J. Struct. Biol. (2016), http://dx.doi.org/10.1016/j.jsb.2016.10.004