



Validation of the Orthogonal Tilt Reconstruction Method with a Biological Test Sample

Citation

Chandramouli, Preethi, Rogelio A. Hernandez-Lopez, Hong-Wei Wang, and Andres E. Leschziner. 2011. Validation of the orthogonal tilt reconstruction method with a biological test sample. *Journal of Structural Biology* 175(1): 85-96.

Published Version

doi:10.1016/j.jsb.2011.04.012

Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:8893763>

Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Open Access Policy Articles, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#OAP>

Share Your Story

The Harvard community has made this article openly available.
Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)

Validation of the Orthogonal Tilt Reconstruction method with a biological test sample.

Preethi Chandramouli¹, Rogelio Hernandez-Lopez², Hong-Wei Wang³ and Andres E Leschziner^{1#}

¹Department of Molecular and Cellular Biology, Harvard University; ²Department of Chemistry and Chemical Biology, Harvard University; ³Department of Molecular Biophysics and Biochemistry, Yale University.

#Corresponding author

Abstract

Electron microscopy of frozen-hydrated samples (cryo-EM) can yield high resolution structures of macromolecular complexes by accurately determining the orientation of large numbers of experimental views of the sample relative to an existing 3D model. The “initial model problem”, the challenge of obtaining these orientations *ab initio*, remains a major bottleneck in determining the structure of novel macromolecules, chiefly those lacking internal symmetry. We previously proposed a method for the generation of initial models--Orthogonal Tilt Reconstruction (OTR)--that bypasses limitations inherent to the other two existing methods, Random Conical Tilt (RCT) and Angular Reconstitution (AR). Here we present a validation of OTR with a biological test sample whose structure was previously solved by RCT: the complex between the yeast exosome and the subunit Rrp44. We show that, as originally demonstrated with synthetic data, OTR generates initial models that do not exhibit the “missing cone” artifacts associated with RCT and show an isotropic distribution of information when compared with the known structure. This eliminates the need for further user intervention to solve these artifacts and makes OTR ideal for automation and the analysis of heterogeneous samples. With the former in mind, we propose a set of simple quantitative criteria that can be used, in combination, to select from a large set of initial reconstructions a subset that can be used as reliable references for refinement to higher resolution.

Keywords

Electron Microscopy; Random Conical Tilt; Orthogonal Tilt Reconstruction; initial model; missing cone

1. Introduction

Electron microscopy of frozen-hydrated samples (cryo-EM) has emerged as a powerful technique capable of providing structural information on large macromolecular complexes not easily accessible to the more traditional biophysical methods. Cryo-EM reconstructions of “single particles”—macromolecules or assemblies that do not form higher-order arrays—have been increasing in resolution over the last several years and have recently yielded the first few atomic and near-atomic resolution structures (Cong et al., 2010; Liu et al., 2010; Ludtke et al., 2008; Yu et al., 2008; Zhang et al., 2010; Zhang et al., 2008). However, these successes are confined mainly to samples with long histories as benchmarks in the field or to those exhibiting a large degree of internal symmetry (such as icosahedral viruses). In contrast to these spectacular advances, structures of novel macromolecules with low internal symmetry typically have significantly lower resolution. Even worse, a number of examples exist of structures solved independently by different research groups that are in disagreement with each other. These limitations are the result of what is known as the “initial model problem”. High-resolution cryo-EM structures are obtained by determining, with high accuracy, the spatial relationship among the individual molecular images obtained in the microscope. This requires the existence of a reference structure to determine those orientations. This structure is not available when a novel sample is being analyzed and the orientations of the experimental images must be determined *ab initio*.

Traditionally, two approaches have been used to generate an initial reconstruction for a novel single-particle sample: Random Conical Tilt (RCT) (Radermacher et al., 1987) and Angular Reconstitution (AR) (Van Heel, 1987). Both approaches are based on the “Central Section Theorem”. This theorem states that the Fourier Transform of a projection of a volume is equivalent to a central section through the three-dimensional (3D) Fourier Transform of that volume in a direction normal to the projection (Frank, 1996). This means that the images collected experimentally, which are projections of the molecule, sample the 3D Fourier Transform of the structure to be determined. The goal of any reconstruction method is to determine the relative orientations of these projections and fill the 3D Fourier Transform as much as possible. The Random Conical Tilt (RCT) method relies on collecting an image with the sample tilted at a high angle followed by a second image collected with no tilt, resulting in two views of each molecule with a known angular relationship. The untilted molecular images are aligned and sorted into groups (“classes”) representing characteristic views of the molecule. The tilted images, which are physically linked to their untilted counterparts, will be randomly distributed in a cone with its axis perpendicular to the average of the untilted images. These tilted images, which

sample 3D Fourier space, are used to obtain reconstructions (“class volumes”) for each characteristic view. The strengths of the RCT approach are twofold: first, the angular relationship between the tilted and untilted images is known experimentally with relatively high accuracy; second, the untilted images are sorted computationally into separate groups thus allowing for the identification and “purification” (*in silico*) of heterogeneity in the sample. Its main shortcoming lies in the fact that there is a limitation in the extent to which the sample can be tilted in the microscope. This limited angle results in cone-shaped areas in Fourier space that are not sampled, a phenomenon known as the “missing cone”. The artifacts in the reconstruction that result from this incomplete sampling are referred to as the “missing cone problem”. Solving it typically requires merging independent reconstructions that are missing information in complementary parts of Fourier space, a non-trivial process requiring significant user intervention. An automated solution to the problem of merging RCT reconstructions was recently proposed by Sander and colleagues (Sander et al., 2010). In their approach, called “weighted RCT” (wRCT), single-class volumes obtained from frozen-hydrated samples (and thus not suffering from stain-induced flattening) are iteratively aligned and weighted according to their signal-to-noise ratio and cross-correlation coefficient relative to a model updated throughout the process. The key features of the method, in addition to the use of vitrified samples, are the low numbers of images in each single-class volume, which increases their ability to sort out heterogeneity at the classification stage, and the weighting algorithm that optimizes the full sampling of the Fourier transform of each structure (Sander et al., 2010).

Angular Reconstitution (AR) determines the spatial relationships among the images mathematically, rather than geometrically as RCT does, by relying on the fact that any two central sections through a 3D Fourier Transform must share a common line where they intersect. This line can be found either by searching in Fourier space or, as implemented in AR, by comparing one-dimensional projections of the experimental class averages (Van Heel, 1987). An important advantage of this approach is the potential absence of the “missing cone problem”, provided the sample adopt enough orientations on the support. However, while elegant conceptually and very powerful with highly symmetric structures, AR has a major limitation: its main underlying assumption is that all the views whose spatial relationships are being determined are different views of the same object. This assumption fails whenever heterogeneity is present in the sample and no *a priori* knowledge of the structure is available to sort the views into separate groups. This will become a very serious limitation as the complexity—and thus potential conformational and biochemical heterogeneity—of novel samples increases.

A few years ago we proposed a new reconstruction approach based on a modification of the RCT data collection geometry. This method—termed Orthogonal Tilt Reconstruction (OTR)—takes advantage of the robustness in the angular relationship between images obtained by tilting the sample while fully sampling the structure in Fourier space (Leschziner and Nogales, 2006). It thus combines

the strengths of RCT and AR while circumventing their main limitations; the “missing cone problem” and the need for user intervention to solve it are eliminated (Leschziner and Nogales, 2006). OTR has as its only requirement that the sample adopt a large number of orientations relative to the electron beam. Images are collected at two orthogonal tilts (typically $+45^\circ$ and -45°) to obtain the equivalent of a 90° “tilt”, which would be physically impossible in the microscope. Other than in the geometry of data collection, OTR differs very little from RCT. One set of images is aligned and classified into different views, allowing for the sorting out of different species present in the sample; the other is used to reconstruct a volume for each view. Because the images used for reconstruction are orthogonal to those used for alignment and classification the structure is fully sampled in Fourier space and consequently does not suffer from incomplete sampling artifacts.

In our initial presentation of the method, we demonstrated its feasibility and advantages using synthetic data, allowing us to analyze and quantify our results by comparing them with the known structure used to generate the data (Leschziner and Nogales, 2006).

In a subsequent paper, we presented three-dimensional reconstructions of the ATP-dependent chromatin remodeling complex RSC from the yeast *S. cerevisiae* using the OTR method and negatively-stained samples (Leschziner et al., 2007). Although this was the first application of OTR to a biological sample, two other reconstructions of the same complex are available (obtained with the RCT method) and all three disagree with each other (Asturias et al., 2002; Chaban et al., 2008; Skiniotis et al., 2007). Until the discrepancies among these structures are resolved we cannot take our RSC reconstruction as validation of the OTR method. It therefore seemed necessary to us to test OTR on a biological sample of known structure where results could be quantified. This would also allow us to gauge how OTR performs when faced with the artifacts present in real samples that did not exist in the synthetic data and could not be analyzed in the case of RSC where a validated structure did not exist. Ideally, our test molecule would also have been solved by RCT, making comparisons with OTR possible.

In this article we present the validation of the OTR method using the yeast exosome bound to the associated subunit Rrp44 (Wang et al., 2007). We chose the exosome—a 398 kDa complex essential for RNA processing in yeast—as our test sample because (1) a Random Conical Tilt reconstruction of the sample, in negative stain, is already available (Wang et al., 2007); (2) crystal structures of the core complex (lacking Rrp44) and the Rrp44 protein have been published (Liu et al., 2006; Lorentzen et al., 2008) and (3) by collecting data from the *exact same grid* used for the RCT reconstruction we could eliminate the effect of sample preparation as a variable in our results. The data presented here confirms the observations we originally made with synthetic data: initial models obtained with OTR are fully sampled in Fourier space (thus lacking artifacts) and can be directly used for refinement without further intervention by the user, allowing for the method to be automated. We also present an approach to select, in a user-independent way, a subset of initial models that are most likely to represent the correct structure.

Finally, an important aspect of the work presented here is that images were obtained using fully automated data collection for OTR geometry as implemented in the Leginon software package (Yoshioka et al., 2007), removing a practical barrier to collecting the relatively large data sets required for OTR.

2. Methods

2.1. Sample preparation

We collected data from the exact same grid Wang and colleagues used for their RCT reconstruction of the exosome (Wang et al., 2007). The grid had been prepared using the “sandwich” method by staining the sample with a 2% uranyl formate solution between two thin layers of carbon on a copper grid (Wang et al., 2007).

2.2. Data acquisition

We collected OTR data at the National Resource for Automated Molecular Microscopy at The Scripps Research Institute. We used a Tecnai F20 microscope operated at 120kV and a magnification of 50,000X. The nominal defocus at the center of the tilted images was 1.50 μ m (underfocus) and the dose was 17.5 e⁻/Å². The images were recorded on a 4k x 4k Tietz SCX CCD camera with a pixel size at the level of the sample of 1.63Å. All the data was acquired using automated OTR as implemented in Leginon (Yoshioka et al., 2007). We collected a total of 130 micrograph pairs from which we extracted 12,692 pairs of particles.

The 0° exosome images we used to refine our initial models for some of the analysis presented here are the same ones used by Wang and colleagues for their RCT reconstruction of the exosome (Wang et al., 2007).

2.3. Data processing: Extraction and preparation of molecular images

We extracted the pairs of tilted images from our micrographs in a semi-automated fashion. First, we obtained the coordinates of the particles on one half of the data set (in this case we used the -45° micrographs) using EMAN’s Boxer program (Tang et al., 2007). We used those coordinates as input for a series of scripts implemented in SPIDER (Frank et al., 1996) that performed the following steps: (1) the micrograph from which coordinates were obtained is aligned to its corresponding tilt mate (the +45° micrograph) by searching over a series of stretches and compressions that mimic deviations from ideal tilt geometry; (2) the alignment parameters (shifts and in-plane rotation) are used to determine the area of overlap between the two micrographs; (3) particles selected in the first micrograph that would not be present in the tilt mate are discarded; (4) the alignment parameters are used to calculate initial estimates for the coordinates of the tilt mate for each particle represented in both micrographs; (5) the

reference particle (-45°) is windowed out of the micrograph within a relatively large box and the coordinates calculated in the previous step are used to box out the corresponding area in the tilt mate ($+45^\circ$); (6) the two boxed out areas are aligned to each other and the resulting shift is used to adjust the initial estimates for the coordinates of the particle in the $+45^\circ$ micrograph; (7) the original coordinates for the -45° particles and the refined coordinates for the $+45^\circ$ particles are used to window out the entire data set.

In order to do CTF correction at the individual particle level, which is necessary due to the tilted nature of our images, we used CTFTILT (Mindell and Grigorieff, 2003) to obtain both the defocus at the center and the tilt geometry parameters of each micrograph. We discarded particles whose defocus was greater than $1.8\ \mu\text{m}$ and/or whose astigmatism was greater than $500\ \text{nm}$; this resulted in a data set containing 11,811 pairs of particles.

After extraction, we performed the following processing steps in SPIDER: (1) CTF correction by phase flipping; (2) two-fold decimation (final pixel size: 3.26\AA); (3) ramp correction; and (4) contrast enhancement by histogram fitting.

At this point, we imported the data into IMAGIC (van Heel et al., 1996) where we applied both low-pass (20\AA) and high-pass (250\AA) filters.

The final combined (-45° and $+45^\circ$) data set contained a total of 23,622 particles.

2.4. Data processing: Alignment and classification

We performed cycles of alignment and classification using the software package IMAGIC (van Heel et al., 1996). We used the entire data set (-45° and $+45^\circ$) in this step without separating the two tilts. The rationale behind merging the two halves is that a $+45^\circ$ particle is the $+90^\circ$ tilt mate of a -45° image while a -45° particle is the -90° tilt mate of a $+45^\circ$ image.

After the alignment and classification of the entire data set had converged we split the particles into 4 groups representing major common views and continued the processing for a few additional cycles; this helped improve the details of the class averages.

We incorporated higher frequency information during the last few cycles of alignment and classification by changing the low-pass filtration from 20\AA to 15\AA .

Our final set consisted of 101 classes containing an average of approximately 230 particles/class.

2.5. Data processing: Analysis of distribution of in-plane rotation angles from alignment and classification.

We performed this analysis as follows: (1) We extracted the cumulative in-plane rotation angles from cycles of multi-reference alignment (MRA) and classification for each class to be analyzed; (2) We divided the full range of possible rotation angles (0° to 360°) into 18 bins of 20° each; (3) Within each

class we used a binomial distribution to determine whether the number of particles whose angles are found in a given bin are within the expected range given the total number of particles in that class; (4) Bins containing too few or too many particles ($p < 0.001$) were assigned a value of “0” and those containing a number of particles within the statistically expected range were given a value of “1”; (5) We assigned each class a final score corresponding to the sum of the scores given to the 18 bins, where a class with a score of 18 would have no bins with a number of particles that deviates from what is expected statistically. Figure S3 shows this approach with a specific example.

2.6. Data processing: Reconstruction of volumes

We built the Euler angular file for the reconstruction of initial volumes as previously described (Leschziner, 2010). As was the case for alignment and classification, we used both -45° and $+45^\circ$ particles for reconstruction. We generated volumes using the command BP 32F in SPIDER (Frank et al., 1996), followed by six cycles of refining the translational parameters of the particles used for the reconstruction. We generated reconstructions for each of the 101 classes obtained (see 2.4).

2.7. Data processing: Refinement of volumes

We refined the OTR initial models by performing 14 cycles of projection-matching in SPIDER (Frank et al., 1996) against the 0° data collected by Wang et al. (Wang et al., 2007). In order to perform the refinements we had to interpolate our initial models, which had a pixel size of 3.26\AA , to the pixel size of the data used by Wang and colleagues (5.18\AA).

2.8. Data analysis: The exosome reference

We wanted to generate an exosome reference structure for our analysis that was not heavily biased by the RCT initial model used to generate the published structure (Wang et al., 2007). To do this, we low-pass filtered the 19\AA exosome reconstruction using a Butterworth filter with pass-band and stop-band frequencies of $1/130\text{\AA}$ and $1/110\text{\AA}$, respectively, and used the resulting ellipsoid as the starting reference for 14 cycles of projection-matching refinement in SPIDER (Frank et al., 1996) against the 0° data collected by Wang et al. We obtained a structure with a resolution of 24\AA (by the 0.5 FSC criterion) that was very similar to the published one (see Figure S2). This structure was used for all the comparisons in this work, except for Figure 2C where the published structure is shown.

2.9. Data analysis: The “Average Fourier Ring Correlation Resolution”

The “Average Fourier Ring Correlation Resolutions” are averages of multiple resolution measurements (in \AA) obtained from Fourier Ring Correlations (FRCs) (Saxton and Baumeister, 1982) calculated between pairs of images. We have used three types of Average FRC Resolution in this work: (1) An Average FRC Resolution (*self*) that measures how well a given initial model matches its

specific experimental data (*i.e.* the images used to generate it); (2) An Average FRC Resolution (*all*) that measures how well a given initial (or refined) model matches the experimental data in general (class averages) (Figure 3B,C) and (3) An Average FRC Resolution (Θ) that measures how well a given initial model matches the reference exosome structure as a function of the tilt angle (Θ) used to generate the projections.

The Average FRC Resolution (*self*) was obtained by calculating FRCs between each experimental image used to reconstruct a given class volume and the corresponding projection of that volume. We do this calculation after the final round of refinement of the translational parameters of the input images (see 2.6). We converted each 0.5 FRC point to a resolution (in Å); the average of all these resolutions for a given initial model is the Average FRC Resolution (*self*).

We calculated the Average FRC Resolution (*all*) shown in Figures 3B,C by performing a multi-reference alignment between 195 evenly spaced projections of a given initial (or refined) model and a set of 229 experimental class averages. The projections of the volume had an angular distance of 10° and the experimental class averages and an average of approximately 100 particles/class. The output of this alignment are the best-matching pairs of projection and class average aligned to each other. We calculated a FRC for each pair and converted the frequency corresponding to the 0.5 FRC point to a resolution (in Å). The global average among all the resolutions obtained for a given initial model is the Average FRC Resolution (*all*) (see Figure S4).

In order to calculate the Average FRC Resolution (Θ) shown in Figure 5 we performed 14 cycles of projection-matching refinement of the initial volumes in SPIDER (Frank et al., 1996) using the 0° data collected by Wang and colleagues (Wang et al., 2007). We aligned the refined volumes to the exosome reference (see 2.8) using OR 3Q in SPIDER and applied the alignment parameters to the initial (non-refined) models. Once the OTR reconstructions had been aligned we calculated projections from them and from the exosome reference for parallel rings that were 10° apart along the tilt direction (Θ) and with an angular distance between projections of 10° within each ring. We calculated the FRC between corresponding projections from an experimental volume and the exosome reference and converted the frequency corresponding to the 0.5 FRC point to a resolution (in Å). The average among all the resolutions obtained for each Θ ring is the Average FRC Resolution (Θ) (see Figure S5).

As it was the case for the refinement of the initial models (2.7), calculating the Average FRC Resolution (Θ) required an interpolation because of the different pixel sizes of the OTR initial models and the refined exosome reference (2.8). In this case we interpolated the exosome reference (5.18Å/pixel) to the dimensions of the OTR data (3.26Å/pixel).

We padded all the projections to 512 x 512 pixels before calculating the FRCs to reduce the noise in the function.

We measured the similarity between initial (or refined) models and the exosome reference (Figure 3) by calculating a Fourier Shell Correlation (Harauz and van Heel, 1986) between each aligned model and the exosome structure. The graphs in Figure 3 show the inverse of the frequency (*i.e.* resolution) corresponding to the 0.5 FSC point.

2.10. Data analysis: Missing cone-based masking of projections for Average FRC Resolution (Θ) calculation

We used SPIDER to generate a binary volume representing the missing cone geometry corresponding to a RCT reconstruction obtained with data collected at 55° (see Figure S6A). We generated each binary mask as follows: (1) we projected the missing cone volume using the Euler angles of the projections to be masked; (2) we thresholded the projections of the missing cone volume to make them binary; (3) we filtered them to smoothen the edges and made them binary again by thresholding. We made the projections larger than the images to be masked and windowed their centers out before use; this avoids having the normal circular contour of projections, which would filter out the corners of the Fourier transforms of the images (see Figure S6B-D).

We imported the projections of the reconstructions into MATLAB (version 7.9.0, R2009b) using the M-file collection of functions as implemented by Bill Baxter (version 1.0, Feb 2009, B. Baxter Copyright (C) 2009 Health Research Inc.) and calculated the Fourier transform in MATLAB. We also imported the binary masks using MATLAB to create the corresponding masking matrix. We used the `fftshift` MATLAB function to shift the components of the matrix to match the corresponding locations in the two dimensional Fourier transform from each projection. We then multiplied the Fourier transform of the projection by the masking matrix and performed an inverse Fourier transform to generate the masked projection, which was finally exported back to SPIDER for FRC calculation.

2.11. The RCT data

In order to make all initial models directly comparable we used the images and angular data from Wang and colleagues (Wang et al., 2007) to generate again all the RCT initial models using the same scripts we used for our OTR data.

3. Results

3.1. The exosome adopts enough orientations and appears amenable to OTR

The main requirement for a sample to be reconstructed using OTR is that it must adopt a large enough number of orientations on the grid (Leschziner, 2010). Strictly speaking, the minimum requirement is that the macromolecule show orientations representing a 45° precession about an axis

perpendicular to the support. This is sufficient to obtain one view of the molecule and a fully-sampled 3D reconstruction (Leschziner, 2010). In practice, this scenario is rather unlikely and one would want as many orientations as possible.

There is no way of determining, without *a priori* knowledge about the structure, whether a sample adopts enough orientations on the support. The heuristic approach we previously proposed (Leschziner et al., 2007) is based on the idea that a sample that adopts truly random orientations would give rise to the same set of views (class averages) regardless of the tilt used during data collection. In more realistic cases, we would expect that samples adopting multiple orientations would result in a number of common views present in both tilted and untilted data sets. Since $\pm 45^\circ$ images are available from OTR data collection and untilted data is often collected for an initial characterization of any new sample, this comparison is easy to implement.

We already knew that the exosome adopted a sufficient number of orientations from the distribution of Euler angles reported for the refinement of the original RCT reconstruction (Wang et al., 2007). However, we wanted to validate our approach with this test sample. We generated 55 class averages from 4,726 untilted images and 65 class averages from 9,792 $\pm 45^\circ$ images. We subjected the two sets, independently, to reference-free alignment and classification and then performed a multi-reference alignment to find the best-matching pairs of class averages. Figure 1 shows that we could obtain a number of very similar views of the exosome at both 0° and $\pm 45^\circ$.

3.2. OTR generates initial reconstructions that appear unaffected by flattening

After alignment and classification, we generated 101 classes with an average of approximately 230 particles/class as well as their corresponding reconstructions. Figure 2A shows a selection of five single-class OTR reconstructions that we judged to be “good” due to their overall similarity to the known exosome structure (Fig.2C). Projection-matching refinement of these volumes against the 0° data used by Wang and colleagues for their exosome work (Wang et al., 2007) yielded structures very similar to the published one in every case (compare Fig.2D with 2C).

Interestingly, despite the fact that the OTR data was collected from the same grid used for the original RCT reconstruction of the exosome, flattening is not apparent in the OTR initial models (Fig.2A), which have dimensions reminiscent of those of the final exosome structure (Fig.2C). This is in contrast to the RCT reconstructions where flattening is pronounced (Fig.2B).

While we do not fully understand the source of this apparent absence of flattening (discussed further in 4.3), this is a phenomenon we had already observed when we applied OTR to the reconstruction of the chromatin remodeling complex RSC (Leschziner et al., 2007).

Figure 2 also shows that RCT reconstructions typically have better defined features relative to OTR volumes when viewed along the direction of the beam (second row in Fig.2A and B). We

observed the same phenomenon when we initially tested OTR with synthetic data (Leschziner and Nogales, 2006). Possible reasons for this are discussed in 4.2.

3.3. How do we select the best initial models from a large set of reconstructions?

The previous section presented a few OTR initial models that we judged to be “good” by their similarity to the known exosome structure. However, many of the reconstructions in our set are of much lower quality. Even worse, in most realistic situations one would not have a “correct” reference structure to be used in the selection and/or validation of initial models. Therefore, we wanted to find some metric, or combination thereof, that would allow us to rank a set of initial reconstructions. The goal was to be able to select a subset that would be most likely to represent correct structures and perform well during refinement. Ideally, this ranking would be performed in an automated fashion and would not require any visual inspection. Given that we would always generate a relatively large set of initial models, we are not aiming to find every good initial model but rather a subset that is likely to represent structures present in the sample.

Initially, we wondered whether some direct measure of resolution might help us identify the best initial reconstructions within our data set. We tested both the Fourier Shell Correlation (FSC) (Harauz and van Heel, 1986) and Rmeasure (Sousa and Grigorieff, 2007). Unfortunately, while a few of the “good” volumes shown in Figure 2A had relatively good resolutions by the 0.5 FSC criterion, others ranked near the bottom of our 101 initial volumes. In order to determine more systematically whether the FSC was indeed a poor predictor of an initial model’s quality we plotted the resolution of each initial model (as the 0.5 FSC point) against an FSC calculated between that model and the exosome. The latter measurement would be an indication of a model’s similarity to the final structure. To calculate this similarity, we used a projection-matching refined version of each initial model to align it to our exosome model (see 2.8) and then applied the alignment parameters to the original initial models before calculating the FSC. In order to address the need for an unbiased exosome structure for all the comparisons performed in this work—one that minimized the bias introduced by the RCT initial models—we took advantage of the fact that the 0° data is robust enough to refine even when the starting model is relatively featureless (HW, unpublished data). We generated an ellipsoid of the right dimensions by low-pass filtering the 19Å exosome reconstruction to 120Å and used this as the starting reference for projection-matching refinement against the 0° data. The resulting structure had a resolution of 24Å by the 0.5 FSC criterion and looked very similar to the published one (Figure S2).

The plot of the resolution of the initial models against the FSC between them and the refined exosome structure shows a weak correlation with a Pearson coefficient of 0.4 (Figure 3A, white squares). The resolution of the initial models is an even poorer predictor of their ability to refine to the correct structure; when we plotted the resolution of the initial models against the similarity between their refined versions and the exosome structure (as the 0.5 FSC expressed as resolution) we obtained a

correlation coefficient of 0.26 (Figure 3A, black circles).

Rmeasure was even less reliable as it tends to fail with low resolution structures (Sousa and Grigorieff, 2007). We saw a sharp break in the distribution of resolutions we obtained for the initial models with Rmeasure, with about half of our reconstructions in the 35Å to 47Å range and the rest all giving a resolution of 90Å (data not shown).

Given these results we decided to switch to a strategy where we would select the best initial models by a process of elimination, using two different criteria to eliminate reconstructions from the set.

For the first round of selection we inspected the distribution of in-plane rotation angles from the cycles of alignment and classification, a strategy we introduced previously for the reconstruction of the chromatin remodeling complex RSC (Leschziner et al., 2007) and that we discussed more extensively recently (Leschziner, 2010). The goal is to avoid generating any reconstruction where the particles in a class show a non-random distribution of in-plane rotation angles. There are two reasons for these distributions to arise when aligning and classifying tilted data: (1) the particles adopt some preferred orientations on the support and/or (2) the sample has been flattened to some degree by the staining process (Leschziner, 2010). Regardless of the origin of the non-random distribution, the gaps in information in Fourier space that result from it would create artifacts in the corresponding reconstructions. Our strategy for detecting non-random distributions of alignment angles is described in detail in 2.5 and outlined in Figure S3. Briefly, we divided the full range of in-plane rotation angles (360°) into bins of 20° each and determined, for each bin, whether the number of particles with rotation angles in that range is lower or higher than would be expected by chance ($p < 0.001$). Bins that show a statistically significant deviation from the expected value are flagged and each class is assigned a score that reflects the total number of flagged bins. We do not yet have an objective criterion to determine where the cutoff should be in terms of tolerance of flagged bins in a given class; for the work presented here we arbitrarily decided to discard any class with 7 or more flagged bins, *i.e.* with fewer than 2/3 of the bins containing the expected number of particles. This resulted in our discarding 29 out of 101 classes.

For our second criterion we reasoned that good initial models would be those that best account for the experimental data. We decided to gauge this by measuring how well projections from a given initial model matched the experimental class averages. Specifically, we generated a large number of evenly-spaced projections (195) from each of the 72 initial models that had not been discarded in the previous step and performed a multi-reference alignment between them and a set of experimental class averages (229, containing an average of approximately 100 particles each). Once a best-matching class average was identified for each projection of a model and aligned with it, we calculated a Fourier Ring Correlation (FRC) between them and extracted the resolution (in Å) corresponding to the 0.5 FRC. Finally, we calculated an average from the 195 resolution values and called this the “Average FRC Resolution (*all*)” for that particular initial model (Figure S4). To test whether this metric was a better

predictor of an initial model's quality, we plotted the Average FRC Resolution (*all*) for each initial model against the similarity (0.5 FSC) between either the model or its refined version and the final exosome structure (Figure 3B). These plots are equivalent to those discussed above for the resolution (FSC) of the initial models (see Figure 3A). The Average FRC Resolution (*all*) appeared to be a much better indicator of a model's quality: the correlation coefficient for the plot of Average FRC Resolution (*all*) vs. the similarity between initial models and the final exosome structure was 0.65 (compared with 0.4 for the FSC in Figure 3A) (Figure 3B, white squares). The Average FRC Resolution (*all*) is also a good predictor of a model's ability to refine to the correct structure: the correlation coefficient for the plot of Average FRC Resolution (*all*) vs. the similarity between refined models and the final exosome structure was 0.63 (compared with 0.26 for the FSC in Figure 3A) (Figure 3B, black circles).

Finally, we wanted to see whether the Average FRC Resolution (*all*) was also a good indicator of a *refined* model's similarity to the correct structure. This might allow us to detect, after refinement, those reconstructions closest to the correct structure. We calculated Average FRC Resolution (*all*) for the refined versions of the 72 initial models that had passed the first selection and plotted these values against the similarity between each refined model and the final exosome structure (Figure 3C). The correlation we observed between the two measures, 0.63, suggests that the Average FRC Resolution (*all*) could also be used to further select reconstructions after a refinement has been performed on the initial models.

When calculating Average FRC Resolutions we could either (1) average the frequencies corresponding to the 0.5 FRC points and convert the average frequency into a resolution or (2) convert each 0.5 FRC frequency into a resolution and average those resolutions. We chose the latter because it gives greater weight to low resolutions, making it more sensitive to "bad" matches between a reconstruction and the data and therefore more likely to discriminate against the worse initial models.

Our ranking of the initial models appears to match well our visual assessment. Of the five initial models shown in Figure 2, which were selected independently of the ranking, three were among the top four volumes (out of 72 that had passed the angular distribution criterion): the first volume in Figure 2 ranked at #1, the third at #2 and the fourth at #4. The other two were ranked as #9 (fifth volume in Fig.2) and #17 (second volume in Fig.2).

Based on the Average FRC Resolution (*all*)'s performance we wondered whether a measure of how well a given reconstruction accounts for the set of images used to generate it (rather than the overall data) would be a good predictor of a volume's quality. We used an approach analogous to that of the Average FRC Resolution (*all*) to obtain an average of the resolutions (0.5 FRC) calculated between each image used to reconstruct a given class volume and the corresponding projection of that volume. We called this measure the Average FRC Resolution (*self*). This parameter, however, seemed to be a poor predictor of an initial model's similarity to the actual structure; the correlation coefficient between the Average FRC Resolution (*self*) and the similarity between the initial model and the

exosome structure was 0.44 (data not shown).

Finally, we wondered how much the ranking was influenced by the number of particles included in a given initial model. We plotted the number of particles per volume against that volume's Average FRC Resolution (*all*) (the parameter used for the final ranking); the correlation coefficient between these two variables was -0.26, suggesting that the number of images was not a major factor in determining an initial model's quality (data not shown).

3.4. A quantitative comparison between RCT and OTR initial models

The better an initial model, the more it will resemble the final structure. Since we did have a final structure in this case, we decided to compare how good our initial OTR reconstructions and Wang *et al.*'s initial RCT reconstructions were by quantifying their resemblance to the exosome. As mentioned above (3.3), we wanted to avoid using the published 19Å exosome structure (Wang *et al.*, 2007) as our reference as it would be biased given that it was obtained from one of the RCT initial models we wanted to compare. Therefore, we used as our reference exosome the structure we obtained by performing a projection-matching refinement using Wang *et al.*'s 0° data and a low-pass filtered (120Å) version of their published structure as the initial reference (Figure S2). As outlined above, we used refined versions of the initial models to calculate alignments to the exosome structure but then applied the alignment parameters to the original initial models. These aligned models are the ones we compared to the exosome structure.

We decided to use Fourier Ring Correlations (FRCs) for our similarity measure, as they would allow us to detect anisotropic distribution of information in the initial models. We had shown in the past, using synthetic data, that projections from RCT reconstructions in the direction where the missing cone effects are most severe (perpendicular to the beam axis) fared less well than those from OTR initial models in terms of their similarity to the corresponding projections from the actual structure (Leschziner and Nogales, 2006). We wanted to make this comparison with a biological sample and extend our original analysis, which we had restricted to projections in the 0° and 90° directions, to the full range of tilt (Θ) angles (0° - 90°). In order to make our comparison more significant statistically we used again our "Average FRC Resolution" parameter but this time calculated it for each set of projections sharing the same Θ value ("Average FRC Resolution (Θ)") (Figure S5). Our expectation, based on our previous work, was that for reconstructions with comparable number of images, projections from RCT volumes would perform better than those from OTR volumes at low tilt angles but deteriorate beyond the experimental tilt angle. On the other hand, we expected OTR volumes to give projections of similar resolution irrespective of their direction and that these would perform better than projections from RCT volumes at tilt angles beyond the experimental angle for RCT.

Figure 4 shows a small subset of the projections we generated to measure the similarity

between each initial model and the final exosome structure. For these comparisons, we used the top three OTR initial models according to the Average FRC Resolution (*all*) ranking and the three RCT models shown in Figure 2. The projections in Figure 4 already show the expected behavior for RCT and OTR volumes. The RCT projections match those from the exosome structure better at low tilt angles ($\Theta = 0^\circ$ and 30°) but progressively deteriorate at higher angles ($\Theta = 60^\circ$ and 90°). The figure also illustrates how the flattening of the sample is apparent in the RCT reconstructions but absent from the OTR initial models (compare the projections from the RCT and OTR models at $\Theta = 60^\circ$ and 90° with the projections from the exosome). Finally, the streaks of density commonly associated with the missing cone can be observed in the projections from the RCT volumes at $\Theta = 90^\circ$ but are absent from any of the projections from the OTR initial models.

These observations were confirmed, quantitatively, by the plot of Average FRC Resolution (Θ) values as a function of tilt angle (Θ) (Figure 5A): the OTR initial models show an isotropic distribution of information, matching the exosome structure to the same extent regardless of the Θ angle used to generate the projections. The RCT volumes, as we had seen visually in Figure 4 and previously with synthetic data (Leschziner and Nogales, 2006), matched the exosome structure better at low Θ values and gradually deteriorated as Θ went beyond the experimental tilt angle, which was 55° for the RCT reconstructions (Wang et al., 2007).

When we saw the results shown in Figure 5A we became concerned about the effect the missing cone could have on the Fourier Ring Correlations. Given that FRCs are calculated by multiplying the Fourier transforms of the reference projection and the corresponding projection from the experimental volume, they would be influenced by any missing data in Fourier space. As the tilt angle used to generate projections from the RCT initial volumes is increased, the area in their Fourier transforms containing data decreases, effectively being masked by the missing cone. When Θ reaches 90° , the Fourier transform of a projection from an RCT initial volume arising from data collected at 55° will not have information in approximately 40% of its area relative to a projection from a volume with no missing cone. It was therefore possible that the differences we observed between RCT and OTR volumes in Figure 5A simply reflected this lack of information as the effect of the missing cone becomes more severe in RCT reconstructions while not affecting those from OTR. Of course, this loss of information is real and a problem that affects RCT reconstructions, but we wanted to make sure we were not overestimating the difference in information content between the OTR and RCT initial models. In order to address this concern, we generated a binary 3D volume representing the missing cone geometry for RCT (Figure S6A). The Z axis of the missing cone was aligned with the Z axis of the reference exosome structure, which had been in turn aligned to the Z axis of the published structure. Then, whenever we generated a projection from an RCT or OTR initial model, we used the same set of

Euler angles to generate a mask from the binary missing cone volume (see 2.10 and Figure S6B-D). We applied this binary mask to the Fourier transforms of both the RCT and OTR projections prior to calculating the FRCs (Figure S6E-H). The areas being compared in the RCT and OTR projections would now be the same and the FRC should report only on the quality of the information present in those areas not covered by the mask. As we expected, reducing the amount of information present in the Fourier transforms resulted in a deterioration of the FRCs beyond the experimental tilt angle (55°) for most volumes (Figure 5B). However, this deterioration was more dramatic for the RCT than the OTR initial models (Figure 5B), showing that the anisotropy seen for the RCT reconstructions in Figure 5A was not an artifact of our measurement.

4. Discussion

4.1. OTR can generate robust initial models

We have presented here a validation of OTR's ability to generate single-class initial reconstructions that are fully sampled in Fourier space and can be used as references for refinement without further intervention by the user.

We chose as our test case a molecular complex whose structure had been independently solved by RCT (Wang et al., 2007) yet the actual EM grid used for data collection was the exact same one in both cases. We can therefore rule out any contribution from sample preparation to the differences we have observed. The only other differences between the two data sets, besides the geometry of data collection, were the electron source, detector and size of the data set. The data used by Wang and colleagues was collected using a LaB₆ filament (in a Tecnai T12 microscope operated at 120kV) on film while our data was collected using a Field Emission Gun (in a Tecnai F20 microscope operated at 120kV) on a CCD camera. After digitization, the pixel size was smaller for the OTR data set (1.63Å vs 2.59Å before decimation). The size of the OTR data set was larger than the RCT one: 12,692 vs. 5,000 pairs of particles. We were helped further by the fact that both tilts (-45° and $+45^\circ$) can be pooled together throughout the data processing, bringing the effective size of the data set to 25,384 particles. While OTR's requirement for multiple orientations of the sample on the support necessitates larger data sets, these are no longer limiting given automated data collection, even for these specific geometries (Yoshioka et al., 2007).

The comparisons presented here between OTR and RCT initial models and a refined exosome structure recapitulate the observations we had made when we introduced OTR using synthetic data (Leschziner and Nogales, 2006). In particular, we observed again that projections from RCT reconstructions are of higher quality when generated at 0° (the direction of the class average) while OTR reconstructions performed better at 90° , the direction where the missing cone effects are most

severe and become evident in the projections from RCT volumes. We extended our analysis here and compared projections from RCT and OTR reconstructions generated at 10° intervals from 0° to 90°. The OTR volumes, which are fully sampled in Fourier space, show an isotropic distribution of resolutions when their projections are compared with equivalent ones from the refined exosome structure (Figures 4 and 5A). The RCT reconstructions show higher resolution at low tilt angles (up to $\Theta = 30^\circ$), at which point they begin to deteriorate and become worse than projections from the OTR initial models once Θ reaches the experimental tilt angle (55°) (Figures 4 and 5A). We have showed here that this difference between OTR and RCT reconstructions at higher Θ angles is not simply a result of the total amount of information present in the projections because of the missing cone; even after we applied a binary mask in Fourier space to make all projections equivalent in terms of their contents in Fourier space, the projections from the RCT initial models deteriorated more than those from the OTR reconstructions (Figure 5B).

The isotropic distribution of information in single-class OTR volumes is one of the method's main strengths. This even sampling of Fourier space makes the references robust and removes the need for any additional user-driven data processing to fill missing data as is typically the case in RCT. In fact, when we combined two or three single-class volumes, the resulting models did not perform any better than the single-class components in the Average FRC Resolution (Θ) measurement (data not shown). Weighted RCT (wRCT), a method recently proposed by Sander and colleagues (Sander et al., 2010), eliminates the need for user intervention during the merging of RCT reconstructions in order to fill the missing data. The method uses a weighting algorithm to account for the amount of overlap of information in Fourier space between two RCT reconstructions; this avoids the bias that favors alignments that lead to volumes having their missing cones in the same orientation (and thus not filling the missing data). While wRCT should make the process of generating fully-sampled initial models from RCT a more objective and robust process, the advantage provided by OTR is that it completely eliminates the need to combine single-class reconstructions. This makes the approach ideally suited for the generation of initial models from heterogeneous samples where decisions regarding the identity of the different molecular species would be better postponed until after an initial refinement has been performed.

Given that the exosome, at approximately 400 kDa, is a relatively small macromolecular complex, we would expect even better performances as the molecular weight increases.

It should be emphasized that the ability to obtain refined volumes that resemble the published exosome structure using OTR initial models is not simply a consequence of the robustness of the exosome data. Even though we were able to obtain a correct exosome structure when we used an initial model consisting of an ellipsoid with the correct dimensions (see 2.8 and Figure S2), OTR initial models that ranked low according to their Average FRC Resolution (*all*) yielded refined structures that

diverged more significantly from the published exosome structure (Figure S7). Those volumes that most resemble the published structure are typically associated with the better Average FRC Resolution (*all*) values (Figure S7), indicating that the data we presented in Figure 2 reflect the actual quality of the OTR initial models.

4.2. Why do RCT reconstructions perform better at low angles?

We observed that RCT reconstructions appear “better” when viewed along the direction of the beam (Figures 2 and 4) and result in higher resolutions when compared with a reference structure at low tilts, *i.e.* using projections generated with Θ angles that are smaller than the tilt angle used for RCT data collection (Figure 5). We had made similar observations using synthetic data when we introduced OTR (Leschziner and Nogales, 2006). These observations are likely due to two key differences between untilted and tilted data that result in the better performance of the former in alignment and classification, as can be seen by comparing class averages obtained from exosome particles collected at 0° or $\pm 45^\circ$ (see Figure 1). First, tilted data contains, by definition, particles spanning a relatively large range of defocus values. Although we correct the CTF of particles individually (see 2.3), large differences in defocus values will affect alignment and classification. The second, and possibly stronger effect is the more severe manifestation of the stain-induced flattening in tilted images. Even though we take advantage of this to some extent later during data processing (see 4.3) this flattening would be expected to have a strong impact in the quality of the class averages. We discuss possible approaches to address these limitations in 4.5.

4.3. Why are OTR reconstructions not affected by flattening?

As mentioned in section 3.2, the OTR single-class reconstructions have relative dimensions reminiscent of those of the final exosome structure and do not display the flattening that can be seen in the RCT initial model(s) used to generate that structure (Figure 2). This is despite the fact that all data (both RCT and OTR) was collected from the same grid, ruling out any variability due to sample preparation.

We do not fully understand all the sources of this effect at this point but we believe that two factors may be at play. First, we do implement a selection step where we discard classes that show a biased distribution of in-plane rotation angles from the cycles of alignment and classification. As discussed above (2.5 and 3.3), one possible source for a biased distribution is the presence of flattening in the data. Therefore, we may be discarding those classes that would give rise to reconstructions most affected by flattening. Second, it is possible that OTR’s geometry leads to an “averaging out” of flattening. In the case of RCT, all particles within a class (and their tilt mates) have the same orientation on the support and are therefore affected by flattening along the same

direction. When a reconstruction is generated, the flattening is also reconstructed and becomes apparent. In the case of OTR, however, every particle in a class (and its tilt mate) has a different orientation on the support and is therefore affected by flattening in a different way. One might expect then that the effect of any flattening remaining after our initial selection would be somewhat diluted as a large number of particles are combined in a reconstruction. This could explain to some extent the fact that surface representations of the RCT initial models tend to show features more reminiscent of those present in the final exosome structure. Since these effects are very difficult to test in a meaningful way using synthetic data we may not be able to fully explain the absence of flattening in OTR reconstructions.

It should be emphasized that OTR's geometry makes the method, unlike RCT, inherently incapable of generating structures that show evidence of flattening. In the ideal scenario of a very large, noise-free data set where different degrees of flattening can be fully sorted out, it would not be possible to generate a single 3D reconstruction exhibiting flattening; every class representing flattening would consist of images with the exact same in-plane orientation and therefore in a set of tilt mates that sample only one central section in Fourier space. In this scenario, the only classes that could yield fully sampled 3D reconstructions are those containing images arising from particles not affected by flattening. In more realistic cases, classes containing images reflecting relatively small amounts of flattening could give rise to full 3D reconstructions where the flattening is averaged out because particles with different orientations on the support are affected by flattening in different ways. We believe it is the combination of this phenomenon with the selection of classes based on their angular distribution (see 2.5 and 3.3) that may be responsible for the absence of flattening in the initial models presented here.

4.4. User-free selection of the best references

One of our main goals in this work was to find some parameter(s) that would allow us to identify, from a large number of initial models, a subset that would be most likely to represent the correct structure(s). Importantly, this identification should not require any intervention by the user so it can be automated. OTR initial models are particularly well suited to this type of approach because their full sampling of Fourier space means they can be used directly as references in refinement without the merging of single-class volumes that is usually performed in RCT reconstructions in order to fill the missing data.

The first step in our selection of initial models is actually performed at the level of the classes. This is where we eliminate a subset of classes that show a strong bias in the distribution of in-plane rotation angles from alignment and classification (see 2.5 and 3.3 and Figure S3). Although the ranking of classes according to the bias they exhibit can be easily automated, we have not yet identified a criterion that would allow us to set a threshold for their elimination. In this work we arbitrarily chose a

value that removed approximately 25% of the classes but this value was chosen mainly based on how many classes we wanted to exclude and on the visual inspection of some initial models. It would be useful to find some relationship between the bias found in a class and some other parameter that would allow us to set the exclusion threshold automatically for any new data set.

Once we generated the initial reconstructions we determined their resolution (by the 0.5 FSC criterion) to test whether it was a good indicator of their similarity to the actual structure. Our data showed that this was not the case (Figure 3A). On the other hand, our “Average FRC Resolution (*all*)” parameter, designed to measure how well an initial model accounts for the experimental data (class averages), performed much better (Figure 3B). Clearly, the Average FRC Resolution (*all*) is not a perfect indicator of an initial model’s quality as the distribution we see in the plot is still fairly broad at the intermediate resolutions. However, our goal was not to identify every single good initial model but rather to be able to find, without visual inspection, a small subset that is likely to perform well in refinement. The usefulness of the Average FRC Resolution (*all*) criterion is supported by two independent observations: first, three of the five OTR initial models we selected visually (by their similarity to the exosome structure) ranked among the top four volumes according to their Average FRC Resolution (*all*) (Figure 2 and section 3.3) and second, the top three ranked volumes, selected *without* any visual inspection, performed similarly well when their projections were compared to those of the exosome (Figures 4 and 5).

It should be noted that our comparisons between OTR and RCT initial models and the final exosome structure (the y-axis FSCs in Figure 3 and the y-axis Average FRC Resolution (Θ) in Figure 5) combine, to some extent, two different effects. All these comparisons rely on our ability to align an initial model to the exosome structure. As described in 3.3, our approach consisted in aligning refined versions of the initial models and applying the alignment parameters to the original reconstructions. Since better initial models will yield better-refined structures they will also be better aligned to the reference exosome structure. Therefore, initial models showing a poorer performance (particularly the FSCs in Figure 3) most likely combine contributions from the true quality of the initial model (the one that could be assessed if alignment were perfect) and an additional penalty resulting from its poorer alignment to the exosome reference. Although we have no way of disentangling these contributions, in the end we are only interested in those initial models that perform best, as those will be the ones to be selected for further processing. These initial models refine well and are thus well aligned to the reference structure.

Fully automated reconstruction using OTR geometry is already available as part of the “allA” toolbox for initial model generation (Voss et al., 2009). The additional selection tools we have introduced here could easily be incorporated into that platform to make the entire process user-independent.

4.5. The future

Our immediate goal is to find ways of improving the performance of OTR data in alignment and classification and narrow the gap in resolution we observed between OTR and RCT reconstructions at low theta values (see 4.2 above). We would also like to bypass altogether the artifacts arising from negative staining. Our current strategy to accomplish this is to move to frozen-hydrated samples and image these using spot-scanning (Downing, 1991) with dynamic focusing (Downing, 1992). Using frozen-hydrated samples will both remove staining artifacts and should also increase the number of orientations adopted by a sample, making more of them amenable to OTR. Spot-scanning, where each image is collected as a raster of independent “spots” instead of as a single flood-beam exposure, allows for each individual spot to be focused separately (dynamic focusing) thus removing the defocus gradient currently present in our OTR data.

Acknowledgments

We would like to thank Craig Yoshioka and Neil Voss for their help during data collection at NRAMM. We would also like to thank members of the Leschziner group for useful discussions. RHL is supported by CONACYT (Mexico) and the Fundación Mexico en Harvard. AL is supported by a fellowship from the Alfred P. Sloan Foundation. HW is supported by the Smith Family Foundation.

Figure Legends

Figure 1. Similar class averages can be obtained from 0° and +/-45° data. We collected images from a negatively stained exosome sample at -45° and +45°. We performed alignment and classification on 9,792 particles obtained from this data set as well as on 4,726 particles from the 0° data collected by Wang and colleagues (Wang et al., 2007) from the same sample grid. We then aligned the two sets of class averages, containing on average 150 particles/class for the +/-45° data and 86 particles/class for the 0° data, to each other to find the best-matching ones. A subset of them is shown here to illustrate that several distinct views can be found in both sets.

Figure 2. Single-class OTR initial reconstructions. (A) We selected five single-class OTR initial reconstructions based on their overall similarity to the published exosome structure; three views are shown of each. **(B)** This panel shows three RCT initial models: the first one (“merged volume A”), consists of 6 merged single-class volumes and is the one used by Wang *et al.* to generate their final exosome structure (Wang et al., 2007); the second (“merged volume B”) is an additional volume generated by merging 5 single-class volumes and the third one is the best of the 6 single-class volumes

merged into “merged volume A”. These volumes are shown in the same orientations used for the OTR volumes in (A). **(C)** The published exosome reconstruction is shown in the same three views used in (A) and (B). **(D)** We refined, by projection-matching, each of the initial reconstructions shown in (A) against the 0° data used by Wang *et al.* to generate their final exosome structure. A single view, corresponding to the middle one in panels (A-C) is shown in this case. The number of images that were used to generate the OTR and RCT reconstructions is indicated in parentheses above the volumes themselves. We oriented the RCT volumes (B) in such a way that the view in the second row is equivalent to looking at the sample along the beam axis; this is also the direction of stain-induced flattening. The views in the first and third row are perpendicular to the one in the second row and would therefore show the effects of flattening most clearly.

Figure 3. Measurement of the ability of different parameters to identify reliable initial models.

In order to obtain the measurements shown in this figure we aligned every OTR initial model to our exosome reference (see 2.8 and 2.9). **(A)** The resolution of each initial model (taken as the 0.5 Fourier Shell Correlation) is plotted against the similarity (as the resolution corresponding to the 0.5 FSC) between the initial (white squares) or refined (black circles) model and the exosome reference. **(B)** The “Average FRC Resolution (*all*)” (see 2.9) of each initial model is plotted against the similarity (as the resolution corresponding to the 0.5 FSC) between the initial (white squares) or refined (black circles) model and the exosome reference. **(C)** The “Average FRC Resolution (*all*)” of the refined models is plotted against the similarity (as the resolution corresponding to the 0.5 FSC) between them and the exosome reference.

Figure 4. Comparison among projections of the OTR and RCT initial models and the final exosome structure as a function of tilt angle (Θ).

We aligned the three top-ranked OTR initial models (as determined by their “Average FRC Resolution (*all*)”) and the three RCT initial models shown in Figure 2 to the final exosome structure (see 3.4). We generated evenly spaced projections for all these volumes for three different Θ values (30° , 60° and 90°) as well as the single possible projection for 0° . The figure shows equivalent projections along each column. The volumes that gave rise to the projections are shown (in color) on the left of the projections panel as a cross-reference to Figure 2.

Figure 5. Quantification of the similarity between OTR or RCT initial models and the final exosome structure as a function of tilt angle (Θ).

(A) Evenly spaced projections (with an angular distance of 10°) were calculated for the same volumes shown in Figure 4 for $0^\circ \leq \Phi < 360^\circ$ and $0^\circ < \Theta < 90^\circ$, every 10° . A Fourier Ring Correlation was calculated between each projection of an initial model and the corresponding projection of the final exosome structure and the resolution (in Å) corresponding

to the 0.5 Fourier Ring Correlation was extracted. An “Average FRC Resolution (Θ)” was calculated from all the projections having a common Θ value (*i.e.* corresponding to the same tilt angle). This Average FRC Resolution (Θ), as well as its standard error, is plotted for the three top-ranked OTR initial models as well as the three RCT initial models shown in Figures 2 and 4 against the tilt angle (Θ). **(B)** This panel is equivalent to that shown in (A) except that a mask was applied, in Fourier space, to the projections of both RCT and OTR volumes to restrict the information used in the calculation of the Fourier Ring Correlations to the area not affected by the missing cone in RCT (see 3.4 for a detailed explanation).

Figure S1. Example of an OTR “tilt pair”. The figure shows a typical pair of micrographs collected using standard OTR geometry as implemented in Legimon (Yoshioka et al., 2007), with one micrograph collected with the sample tilted to -45° **(A)** and the second with the sample tilted to $+45^\circ$ **(B)**. The insets show blown up versions of the areas highlighted by the dotted squares and correspond to areas containing tilt mates. The data was collected at the National Resource for Automated Molecular Microscopy (The Scripps Research Institute).

Figure S2. Comparison between the published exosome structure and the reconstruction we generated for our analysis in this work. The figure shows three views (the same ones used in Figure 2) of both the published exosome structure (light grey) (Wang et al., 2007) and the reconstruction we generated (dark grey) by doing a projection-matching refinement starting with a low-pass filtered (120Å) version of the published structure. The structure on the right is the ones we used as the reference for all the comparisons reported in this work.

Figure S3. Quantification of bias in the distribution of in-plane rotation angles for images within a class. This figure shows an example of our strategy for determining whether any bias exists in the distribution of in-plane rotation angles resulting from the alignment of images within a class and quantifying the extent of that bias. **(A)** A visual representation of the distribution of in-plane rotation angles in a class. The plot shown at the bottom of this panel is equivalent to looking at the edges of the images used in a reconstruction from the top (symbolized by the eye). These plots can be used to detect strongly biased distributions visually. **(B)** In order to quantify any bias seen within a class, we divided the full range of in-plane rotation angles (0° - 360°) into 18 “bins” of 20° each. For each class, we know the number of particles it contains (N) as well as the in-plane rotation angle for each image in the class; from this angle we determined how many particles fall within a given bin (X). We used a binomial distribution to determine, given N particles in the class, whether a bin has too few or too many particles ($p < 0.001$). **(C)** We then flagged each bin as “good” (within the expected values) or “bad” (statistically

too many or too few particles). The total number of “good” bins is assigned to the class as the indicator of its distribution of in-plane rotation angles. A score of 18 would correspond to a class with no bias according to the criteria we used here.

Figure S4. Calculation of the “Average Fourier Ring Correlation Resolution”. (A) We generated evenly spaced projections from a given initial model. We used each one of these projections to search the best-matching experimental class average in the data set (B). We aligned the projection and the class average to each other (C) and calculated a Fourier Ring Correlation between them (D). We extracted the frequency corresponding to the 0.5 Fourier Ring Correlation and converted that to a resolution (in Å) and calculated a global average (“Average FRC Resolution (*all*)”) from all the projections (E).

Figure S5. Calculation of the “Average Fourier Ring Correlation Resolution” between an initial experimental volume and a reference as a function of the projection angle Θ . We aligned the experimental initial volume (shown in red in A) to a reference volume (shown in grey in B) using a refined version of the experimental volume for the alignment and applying the alignment parameters to the original one. We generated evenly-spaced projections (10° apart) from both the experimental (A) and reference volumes (B) with $0^\circ \leq \Phi < 360^\circ$ and $0^\circ < \Theta < 90^\circ$. (The drawings in (A) and (B) only show a few of the projections generated for $\Theta = 50^\circ$ and $\Theta = 90^\circ$.) This approach results in a total of 195 projections, with 1 projection having $\Theta = 0^\circ$ (the direction of the class average) and the rest ranging from 5 projections for $\Theta = 10^\circ$ to 34 projections for $\Theta = 80^\circ$ (there are only 18 projections for $\Theta = 90^\circ$ because the remaining ones are mirrors of the first half). We used each pair of projections having the same set of Euler angles (*prj_i* in A and B) to calculate a Fourier Ring Correlation between them (C). We extracted the frequency corresponding to the 0.5 Fourier Ring Correlation and converted that to a resolution (in Å). We calculated an “Average FRC Resolution (Θ)” from all the FRCs between projections sharing **the same Θ value**, that is, lying along the same horizontal ring (as shown by the $\Theta = 50^\circ$ and $\Theta = 90^\circ$ rings in the figure) (D). Projections from RCT reconstructions that share a common Θ value should be affected equally by the Missing Cone. We repeated the process for $0^\circ < \Theta < 90^\circ$ every 10° and plotted these Average FRC Resolution (Θ) values against Θ for each initial volume we analyzed (E).

Figure S6. Strategy for weighting Fourier Ring Correlations according to the distribution of information due to the Missing Cone. The goal of this approach is to make sure that differences we observe in Fourier Ring Correlations for OTR and RCT initial volumes are not due to the different

distribution of information in Fourier space due to the missing cone, present in RCT but absent in OTR. In this approach, we generate a binary mask representing the missing cone and use it to restrict the calculation of Fourier Ring Correlations to those areas in Fourier space where both RCT and OTR contain data. **(A)** A 3D binary volume showing the part of Fourier space that contains information in the case of a RCT reconstruction obtained from data collected at a tilt angle of 45° . This volume is tilted towards the viewer; the beam axis is coincident with the axis of the cone seen in the volume. **(B)** A side view of the distal half of the same volume, this time with the Missing Cone axis aligned along the vertical. The solid areas are those containing information. We can create images from this volume equivalent to a central section for any Θ angle used to generate projections to calculate Fourier Ring Correlations. The 2D binary mask shown in **(D)** corresponds directly to the central section of the volume shown in **(B)** and is the mask that we would use whenever $\Theta = 90^\circ$. This is the tilt angle where information content is minimal in Fourier space for a RCT reconstruction. **(C)** A similar 2D binary mask, this time for a projection generated using $\Theta = 60^\circ$. As we move away from $\Theta = 90^\circ$ the amount of information present in a projection increases. The implementation of this Missing Cone corrections is as follows: **(E)** We align a given initial experimental volume and a reference to each other (see 3.4); **(F)** We calculate projections of these volumes using a range of Φ and Θ angles, as shown in Figure S4; **(G)** Standard Fourier Ring Correlations are calculated by using the Fourier transforms of these projections; **(H)** To correct for distribution of information in Fourier space, we apply a binary mask (obtained from the 3D binary mask as outlined above) directly to the Fourier transform and use these masked Fourier transforms to obtain the “corrected” Average FRC Resolution (Θ).

Figure S7. Visual assessment of the quality of the OTR initial models as a function of their “Average Fourier Ring Correlation Resolution”. In order to visually assess how OTR initial models performed as a function of their Average FRC Resolution we took every 5th initial volume according to its ranking (*i.e.* #1, #6, #11... #71) and compared its refined version to the published exosome structure. We only used volumes from the 72 that had passed the angular distribution criterion. **(A)** This plot is taken directly from that shown in Figure 3B (see Figure 3B for details); it shows the data for the subset of 15 volumes we analyzed for this figure. The numbers adjacent to the white squares in the plot refer to the ranking of each initial model among the set of 72 initial volumes that passed the angular distribution criterion; these numbers correspond to the “ranking” numbers in (B). **(B)** Each OTR initial volume was refined against the same 0° data used by Wang and colleagues to obtain their exosome structure (Wang et al., 2007) (see 2.7 for details). In addition to showing the initial (top row) and refined (bottom row) models, this panel also lists the Average FRC Resolution (*all*) obtained for each initial model (in Å) as well as the resolution (according to the 0.5 FSC criterion) and the similarity to the exosome (also measured as the 0.5 FSC point) measured for each refined model

(see 2.8 for details). The Average FRC Resolution of the initial models and the similarity of the refined structures to the exosome are the values plotted in the X and Y axis, respectively, in (A). The published exosome structure is shown in grey in the inset at the bottom right of the panel; we aligned all the other volumes visually to match this view of the complex. We filtered all the volumes to 30Å for the purposes of this figure.

Bibliography

- Asturias, F.J., Chung, W.H., Kornberg, R.D., and Lorch, Y. (2002). Structural analysis of the RSC chromatin-remodeling complex. *Proc Natl Acad Sci U S A* **99**, 13477-13480.
- Chaban, Y., Ezeokonkwo, C., Chung, W.H., Zhang, F., Kornberg, R.D., *et al.* (2008). Structure of a RSC-nucleosome complex and insights into chromatin remodeling. *Nat Struct Mol Biol* **15**, 1272-1277.
- Cong, Y., Baker, M.L., Jakana, J., Woolford, D., Miller, E.J., *et al.* (2010). 4.0-Å resolution cryo-EM structure of the mammalian chaperonin TRiC/CCT reveals its unique subunit arrangement. *Proc Natl Acad Sci U S A* **107**, 4967-4972.
- Downing, K.H. (1991). Spot-scan imaging in transmission electron microscopy. *Science* **251**, 53-59.
- Downing, K.H. (1992). Automatic focus correction for spot-scan imaging of tilted specimens. *Ultramicroscopy* **46**, 199-206.
- Frank, J. (1996). *Three-Dimensional Electron Microscopy of Macromolecular Assemblies* (San Diego, California, Academic Press, Inc.).
- Frank, J., Radermacher, M., Penczek, P., Zhu, J., Li, Y., *et al.* (1996). SPIDER and WEB: processing and visualization of images in 3D electron microscopy and related fields. *J Struct Biol* **116**, 190-199.
- Harauz, G., and van Heel, M. (1986). *Direct 3D reconstruction from projections with initially unknown angles* (Amsterdam, Elsevier).
- Leschziner, A. (2010). The orthogonal tilt reconstruction method. *Methods Enzymol* **482**, 237-262.
- Leschziner, A.E., and Nogales, E. (2006). The orthogonal tilt reconstruction method: an approach to generating single-class volumes with no missing cone for ab initio reconstruction of asymmetric particles. *J Struct Biol* **153**, 284-299.
- Leschziner, A.E., Saha, A., Wittmeyer, J., Zhang, Y., Bustamante, C., *et al.* (2007). Conformational flexibility in the chromatin remodeler RSC observed by electron microscopy and the orthogonal tilt reconstruction method. *Proc Natl Acad Sci U S A* **104**, 4913-4918.
- Liu, H., Jin, L., Koh, S.B., Atanasov, I., Schein, S., *et al.* (2010). Atomic structure of human adenovirus by cryo-EM reveals interactions among protein networks. *Science* **329**, 1038-1043.
- Liu, Q., Greimann, J.C., and Lima, C.D. (2006). Reconstitution, activities, and structure of the eukaryotic RNA exosome. *Cell* **127**, 1223-1237.
- Lorentzen, E., Basquin, J., Tomecki, R., Dziembowski, A., and Conti, E. (2008). Structure of the active subunit of the yeast exosome core, Rps44: diverse modes of substrate recruitment in the RNase II nuclease family. *Mol Cell* **29**, 717-728.
- Ludtke, S.J., Baker, M.L., Chen, D.H., Song, J.L., Chuang, D.T., *et al.* (2008). De novo backbone trace of GroEL from single particle electron cryomicroscopy. *Structure* **16**, 441-448.
- Mindell, J.A., and Grigorieff, N. (2003). Accurate determination of local defocus and specimen tilt in electron microscopy. *J Struct Biol* **142**, 334-347.
- Radermacher, M., Wagenknecht, T., Verschoor, A., and Frank, J. (1987). Three-dimensional reconstruction from a single-exposure, random conical tilt series applied to the 50S ribosomal

- subunit of Escherichia coli. *J Microsc* *146* (Pt 2), 113-136.
- Sander, B., Golas, M.M., Luhrmann, R., and Stark, H. (2010). An approach for de novo structure determination of dynamic molecular assemblies by electron cryomicroscopy. *Structure* *18*, 667-676.
- Saxton, W.O., and Baumeister, W. (1982). The correlation averaging of a regularly arranged bacterial cell envelope protein. *J Microsc* *127* (Pt 2), 127-138.
- Skiniotis, G., Moazed, D., and Walz, T. (2007). Acetylated histone tail peptides induce structural rearrangements in the RSC chromatin remodeling complex. *J Biol Chem* *282*, 20804-20808.
- Sousa, D., and Grigorieff, N. (2007). Ab initio resolution measurement for single particle structures. *J Struct Biol* *157*, 201-210.
- Tang, G., Peng, L., Baldwin, P.R., Mann, D.S., Jiang, W., *et al.* (2007). EMAN2: an extensible image processing suite for electron microscopy. *J Struct Biol* *157*, 38-46.
- Van Heel, M. (1987). Angular reconstitution: a posteriori assignment of projection directions for 3D reconstruction. *Ultramicroscopy* *21*, 111-123.
- van Heel, M., Harauz, G., Orlova, E.V., Schmidt, R., and Schatz, M. (1996). A new generation of the IMAGIC image processing system. *J Struct Biol* *116*, 17-24.
- Voss, N.R., Lyumkis, D., Cheng, A., Lau, P.W., Mulder, A., *et al.* (2009). allA Toolbox for Ab Initio 3-D Reconstructions in Single-particle Electron Microscopy. *J Struct Biol*.
- Wang, H.W., Wang, J., Ding, F., Callahan, K., Bratkowski, M.A., *et al.* (2007). Architecture of the yeast Rrp44 exosome complex suggests routes of RNA recruitment for 3' end processing. *Proc Natl Acad Sci U S A* *104*, 16844-16849.
- Yoshioka, C., Pulokas, J., Fellmann, D., Potter, C.S., Milligan, R.A., *et al.* (2007). Automation of random conical tilt and orthogonal tilt data collection using feature-based correlation. *J Struct Biol* *159*, 335-346.
- Yu, X., Jin, L., and Zhou, Z.H. (2008). 3.88 Å structure of cytoplasmic polyhedrosis virus by cryo-electron microscopy. *Nature* *453*, 415-419.
- Zhang, J., Baker, M.L., Schroder, G.F., Douglas, N.R., Reissmann, S., *et al.* (2010). Mechanism of folding chamber closure in a group II chaperonin. *Nature* *463*, 379-383.
- Zhang, X., Settembre, E., Xu, C., Dormitzer, P.R., Bellamy, R., *et al.* (2008). Near-atomic resolution using electron cryomicroscopy and single-particle reconstruction. *Proc Natl Acad Sci U S A* *105*, 1867-1872.

