



Semi-automated selection of cryo-EM particles in RELION-1.3

Sjors H.W. Scheres

MRC Laboratory of Molecular Biology, Cambridge Biomedical Campus, Francis Crick Avenue, Cambridge CB2 0QH, UK



ARTICLE INFO

Article history:

Received 28 October 2014

Received in revised form 20 November 2014

Accepted 30 November 2014

Available online 6 December 2014

Keywords:

Electron cryo-microscopy

Single-particle analysis

Automated particle picking

ABSTRACT

The selection of particles suitable for high-resolution cryo-EM structure determination from noisy micrographs may represent a tedious and time-consuming step. Here, a semi-automated particle selection procedure is presented that has been implemented within the open-source software RELION. At the heart of the procedure lies a fully CTF-corrected template-based picking algorithm, which is supplemented by a fast sorting algorithm and reference-free 2D class averaging to remove false positives. With only limited user-interaction, the proposed procedure yields results that are comparable to manual particle selection. Together with an improved graphical user interface, these developments further contribute to turning RELION from a stand-alone refinement program into a convenient image processing pipeline for the entire single-particle approach.

© 2014 The Author. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Recent advances in electron cryo-microscopy (cryo-EM) single-particle analysis have made it possible to obtain near-atomic resolution structures for a much wider range of specimens and from much fewer particles than before. Previously, cryo-EM maps with sufficient detail to see amino acid side chains could only be obtained for hundreds of thousands asymmetric units of large icosahedral viruses (Grigorieff and Harrison, 2011). However, last year a ribosome reconstruction with details of around 4 Å was reported from 35 thousand (asymmetric) particles (Bai et al., 2013), and a 20S proteasome structure to 3.3 Å was reported from 1.8 million asymmetric units (Li et al., 2013). More recently, a 3.2 Å map for the yeast mitochondrial large ribosomal subunit was reported from 47 thousand particles (Amunts et al., 2014), a 3.4 Å structure of the F₄₂₀-reducing [NiFe] hydrogenase from 319 thousand asymmetric units (Allegretti et al., 2014), and a 3.4 Å structure of the TRPV1 ion channel from 142 thousand asymmetric units (Cao et al., 2013).

Two developments play an important role in these advances. The first is the development of direct-electron detectors, which are much more efficient at detecting electrons than conventionally used photographic film or charged-coupled devices (CCDs) (McMullan et al., 2009). The higher detection quantum efficiency (DQE) of the new detectors yield images with much improved signal-to-noise ratios (SNRs). This has a “double effect” on the resolution of 3D reconstructions: not only need one average over fewer particles to obtain a given resolution, but one can also align and classify each particle better, so that reconstructions are blurred

to a much smaller extent than before. This then also relates to the second development: that of powerful new image processing algorithms. In particular, unsupervised image classification algorithms may be used to separate projections of distinct 3D structures, so that relatively impure or structurally heterogeneous samples may still lead to high-resolution structure determination, e.g. see (Fernandez et al., 2013; Voorhees et al., 2014). Moreover, as the new detectors are also very fast, one can now record multiple images during irradiation of the sample in the microscope. Since interactions with the incoming electrons cause movement of the sample, movie processing algorithms that correct for these beam-induced movements may further increase resolution (Campbell et al., 2012; Bai et al., 2013; Li et al., 2013). These developments have opened up the possibility to apply high-resolution cryo-EM structure determination to a much wider range of samples than before, which will attract many new researchers to this exciting field.

Together with increased interest in the technique, the call for high-throughput, easy-of-use and automation will also grow. One step in the data processing pipeline of high-resolution structure determination that may take considerable amounts of time and user-input is the selection of particles that are suitable for 3D reconstruction. In the past this process was typically done manually by the researcher, who would sit in front of a computer screen and click on each individual particle. Over the last 15 years, many algorithms to automate this often tedious procedure have been proposed, see (Nicholson and Glaeser, 2001; Zhu et al., 2004) for earlier reviews. More recently, implementations of automated particle picking algorithms were made available in EMAN2 (Tang et al., 2007), SIGNATURE (Chen and Grigorieff, 2007), DOGPICKER

E-mail address: scheres@mrc-lmb.cam.ac.uk

(Voss et al., 2009), XMIPP (Sorzano et al., 2009), and ARACHNID (Langlois et al., 2014) among others. These approaches may broadly be divided into two categories: feature-based and template-based approaches. In the feature-based approaches, different characteristics of the particles are expressed in some numerical manner (features) and features calculated from local areas in the micrographs are compared to a set of expected features. In the template-based approaches, images that express the expectation how the particles look like are correlated against the micrographs, often using fast Fourier-transform (FFT) accelerated algorithms (Roseman, 2003, 2004). The distinction between the two types of approaches is not always clear, as sometimes expected features are calculated from template images themselves. In general, template-based approaches introduce a higher degree of prior information into the picking process than feature-based approaches, which may be both an advantage and a disadvantage. The advantage of using more prior information is that it allows to detect weaker signals. However, the high levels of noise in the micrographs also make the picking task extremely prone to reference bias. Thereby, relying heavier on prior information becomes a disadvantage in cases where this information is incorrect.

This paper describes recent developments in the RELION software (Scheres, 2012a,b) that are centred around a new template-based particle picking algorithm. The choice for a template-based approach was motivated by its larger potential to select particles from noisy data. The workflow proposed is a semi-automated one. The researcher manually picks particles from a low number of micrographs; uses reference-free 2D class averaging inside RELION to calculate average images of these particles; performs template-based automated particle picking with those class averages on all micrographs; and then relies on a new sorting algorithm and further 2D class averaging to remove false positive from the data. To facilitate this process an improved graphical user interface (GUI) was also implemented. Whereas RELION was originally proposed as a stand-alone refinement program, these developments continue the evolution of RELION into a software package that provides a convenient pipeline for most of the single particle analysis tasks. Together with an improved movie-processing approach to correct for beam-induced motion in samples of relatively small particles (Scheres, 2014), the developments presented here represent the main improvements in the latest (1.3) release of RELION.

2. Approach

2.1. Particle picking

The template-based particle picking algorithm proposed employs an additive model with white Gaussian noise in real-space. A micrograph \mathbf{X} that contains N individual particles i at coordinates \vec{t}_i in the micrograph is described as follows (also see Fig. 1A):

$$\mathbf{X}(\vec{r}) = \mu(\vec{r}) + \sigma(\vec{r}) * \left\{ \mathbf{N}(\vec{r}) + \sum_{i=1}^N \mathbf{A}_{k_i}^{\phi_i}(\vec{r} - \vec{t}_i) \right\}, \quad (1)$$

where:

- $\mathbf{X}(\vec{r})$ is the micrograph, i.e. a two-dimensional image that was recorded in the electron microscope, and r describes the two-dimensional position in that image.
- $\mathbf{N}(\vec{r})$ is an image of independent (or white) Gaussian noise with mean zero and standard deviation one.
- $\mu(\vec{r})$ and $\sigma(\vec{r})$ are position-dependent additive and multiplicative normalisation factors that bring the recorded noise levels in the micrograph to mean zero and standard deviation one. Variations in μ and σ with position typically describe experimental variations in ice thickness, electron dose, etc.

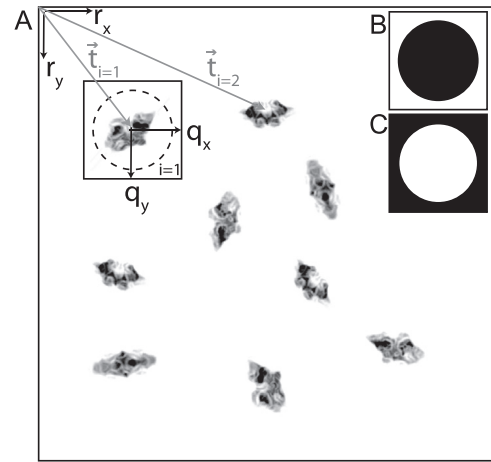


Fig. 1. Schematic representation of the data model. (A) Representation of a micrograph, with coordinate vectors $r = (r_x, r_y)$ inside the micrograph, and coordinate vectors $q = (q_x, q_y)$ inside each particle image. Vectors $t_i = (t_x, t_y)$ place the i th particle inside the micrograph with an unknown in-plane rotation ϕ_i with respect to a common frame of reference. Inset (B), mask \mathbf{M}_o which is used for normalisation of the particle images: average and standard deviation of the background pixels are calculated in the white area of this mask. Inset (C), mask \mathbf{M}_i which is used for the particle sorting algorithm: all statistics on the difference images between each particle and its corresponding template are calculated in the white area of this mask.

- $\mathbf{A}_{k_i}^{\phi_i}$ is one of K known, two-dimensional template images \mathbf{A}_k with internal positions \vec{q} . Typically, the template images are much smaller than the micrograph (in the summation above, the template image is zero outside the defined box size). Therefore, for any given \vec{r} and template \mathbf{A}_{k_i} at position \vec{t}_i , the internal position will be $\vec{q} = \vec{r} - \vec{t}_i$. The K different template images may describe projections in different directions of the same molecule, or they may describe projections of different molecules; k_i describes which of the template images corresponds to the i th particle; and ϕ_i describes the relative in-plane rotation between the particle and that template image.

Given $\mathbf{X}(\vec{r})$ and K template images \mathbf{A}_k , the task at hand is to identify all N combinations of \vec{t}_i , ϕ_i and k_i . Based on positive experiences with maximum-likelihood approaches, e.g. see (Scheres et al., 2007; Scheres, 2012a), the choice was made to implement a probability-based similarity metric for this task. The assumption of Gaussian noise in Eq. (1) naturally leads to a Gaussian similarity measure. Unlike the cross-correlation coefficient, as used for example in the template-based picking program findEM (Roseman, 2003, 2004), the squared difference term inside the Gaussian metric is not invariant to multiplication with or addition of a constant. This means that one needs to account for the varying intensity levels in the recorded micrographs, and one needs to determine the normalisation factors $\mu(\vec{r})$ and $\sigma(\vec{r})$ to bring all particles on the same intensity level.

Upon extraction of individual particles from the micrographs, RELION relies on a normalisation procedure that uses a circle (with a user-defined radius R , see Fig. 1B) to divide each extracted particle image in a background area (outside the circle) and a particle area (inside the same circle) (Sorzano et al., 2004). By subtracting the average value of the pixels in the background area from the entire particle image, and subsequently dividing the entire image by the standard deviation of the pixels in the background area, noise levels with zero-mean and unity-standard deviation are obtained for all particles, independent of variations in ice thickness, exposure or other uncontrolled experimental factors.

The same normalisation procedure is used inside the implemented particle picking procedure. Using similar concepts as in the fast local correlation algorithm inside findEM (Roseman, 2003, 2004), the values of $\mu(\vec{r})$ and $\sigma(\vec{r})$ can be precalculated efficiently for all \vec{r} using fast Fourier transforms (FFTs):

$$\mu(\vec{r}) = \frac{1}{M_o} FT^{-1} \{ FT(\mathbf{X}) FT(\mathbf{M}_o)^* \}, \quad (2)$$

$$\sigma(\vec{r}) = \sqrt{\frac{1}{M_o} FT^{-1} \{ FT(\mathbf{X}^2) FT(\mathbf{M}_o)^* \} - \mu^2(\vec{r})}, \quad (3)$$

where FT and FT^{-1} denote forward and inverse Fourier transform operations, $*$ denotes complex conjugation, and \mathbf{M}_o is a binary mask as depicted in Fig. 1B with M_o white pixels. A second mask \mathbf{M}_i is the inverse of \mathbf{M}_o and has M_i white pixels (Fig. 1C).

Given the data model in Eq. (1), the probability of observing the micrograph with a particle corresponding to template image \mathbf{A}_k in orientation ϕ and position \vec{t} is then given by the multiplication of a Gaussian (with a standard deviation of unity) for each pixel \vec{q} inside mask \mathbf{M}_i :

$$P(\mathbf{X}|\vec{t}, \mathbf{A}_k^\phi) \propto \exp \sum_{\vec{q} \in \mathbf{M}_i} -\frac{1}{2} \left(\frac{\mathbf{X}(\vec{q} + \vec{t}) - \mu(\vec{r})}{\sigma(\vec{r})} - \mathbf{A}_k^\phi(\vec{q}) \right)^2. \quad (4)$$

The range of values of $P(\mathbf{X}|\vec{t}, \mathbf{A}_k^\phi)$ depends on the number of pixels inside mask \mathbf{M}_i and the power of the signal in \mathbf{A}_k^ϕ . In order to define a similarity metric with a pre-defined range, one also calculates the probability of observing the micrograph with only noise at that position, i.e. as for an all-zero template image \mathbf{O} :

$$P(\mathbf{X}|\vec{t}, \mathbf{O}) \propto \exp \sum_{\vec{q} \in \mathbf{M}_i} -\frac{1}{2} \left(\frac{\mathbf{X}(\vec{q} + \vec{t}) - \mu(\vec{r})}{\sigma(\vec{r})} \right)^2. \quad (5)$$

Subsequently, one calculates the ratio of $P(\mathbf{X}|\vec{t}, \mathbf{A}_k^\phi)$ and $P(\mathbf{X}|\vec{t}, \mathbf{O})$, which will be denoted as $R_{\phi,k}(\vec{t})$, using:

$$\begin{aligned} R_{\phi,k}(\vec{t}) &= \frac{P(\mathbf{X}|\vec{t}, \mathbf{A}_k^\phi)}{P(\mathbf{X}|\vec{t}, \mathbf{O})} \\ &= \exp \sum_{\vec{q} \in \mathbf{M}_i} \frac{\mathbf{X}(\vec{q} + \vec{t}) \mathbf{A}_k^\phi(\vec{q})}{\sigma(\vec{r})} - \frac{\mu(\vec{r}) \mathbf{A}_k^\phi(\vec{q})}{\sigma(\vec{r})} - \frac{1}{2} \left(\mathbf{A}_k^\phi(\vec{q}) \right)^2. \end{aligned} \quad (6)$$

If $R_{\phi,k}(\vec{t}) > 1$, the position \vec{t} is more likely to correspond to a particle \mathbf{A}_k^ϕ than to solvent. The expected value for $R_{\phi,k}(\vec{t})$ for an image according to the data model in Eq. (1) is calculated as:

$$E\langle R_k \rangle = \exp \frac{1}{2M_i} \sum_{\vec{q} \in \mathbf{M}_i} (\mathbf{A}_k)^2. \quad (7)$$

Therefore, one can define a similarity metric that adopts values within a meaningful range by expressing the fraction:

$$S_{\phi,k}(\vec{t}) = \frac{R_{\phi,k}(\vec{t}) - 1}{E\langle R_k \rangle - 1}. \quad (8)$$

The value $R_{\phi,k}(\vec{t}) - 1$ expresses how much more likely the position \vec{t} is to correspond to a particle \mathbf{A}_k^ϕ than to solvent. For perfect signal and white Gaussian noise, this value will be close to $E\langle R_k \rangle - 1$. In practice, the templates are not perfect and the noise is not white, which results in typical values of $R_{\phi,k}(\vec{t}) - 1$ being smaller than $E\langle R_k \rangle - 1$. Therefore, useful threshold values for peak searching in $S_{\phi,k}(\vec{t})$ often lie within the range (0, 1].

The calculation of $S_{\phi,k}(\vec{t})$ for all ϕ and k can be done efficiently as follows. The calculation of $\sum_{\vec{q} \in \mathbf{M}_i} \mathbf{A}_k^\phi$ and $\sum_{\vec{q} \in \mathbf{M}_i} (\mathbf{A}_k^\phi)^2$ is invariant to \vec{t} and ϕ , and thus need to be calculated only once for each template. Calculation of the remaining unknown in Eq. (6) for all \vec{t}, k and ϕ may again be calculated using FFT-accelerated cross-correlation:

$$\sum_{\vec{q} \in \mathbf{M}_i} \mathbf{X}(\vec{q} + \vec{t}) \mathbf{A}_k^\phi(\vec{q}) = FT^{-1} \{ FT(\mathbf{X}) FT(\mathbf{A}_k^\phi)^* \}, \quad (9)$$

provided that $\mathbf{A}_k^\phi(\vec{q}) = 0$ for $\vec{q} \notin \mathbf{M}_i$. For each template image \mathbf{A}_k and each discretely sampled in-plane rotation $\phi = 1, \dots, \Phi$, one calculates $S_{\phi,k}(\vec{t})$ for all \vec{t} . This involves $K \times \Phi$ evaluations of Eq. (9), whereas Eqs. (2) and (3) only need to be evaluated once for each micrograph. The user controls Φ through the definition of an angular sampling rate (typically 5 degrees). The FFT libraries used only deal with squared images. Rectangular micrographs are padded with white Gaussian noise to obtain squared images. This is done internally, so that the interaction with the user does not change.

Peak searching is first done independently for all $S_{\phi,k}(\vec{t})$. Potential particle positions are only selected for local maxima (where the four neighbouring pixels are smaller than the peak value), and if the peak value is higher than a given user-defined threshold. For each image $S_{\phi,k}(\vec{t})$, the peaks are pruned based on a user-defined minimum inter-particle distance. All peaks within this distance from each other are clustered together. Within each cluster, the peak with the highest value of $S_{\phi,k}(\vec{t})$ is kept, and remaining peaks within the minimum inter-particle distance of the kept peak are discarded. This is done recursively, such that more than one peak from each cluster may be kept, but all pruned peaks will be at least the minimum inter-particle distance from each other. The pruned peaks for all templates k and all in-plane rotations ϕ are then combined, and the combined list of peaks is pruned using the same algorithm. The final result is a list of N particle coordinates \vec{t}_i , each with a value for the corresponding template k_i and in-plane rotation ϕ_i .

Although not explicitly written as such in Eq. (1), the implementation in RELION-1.3 is done in such a way that CTF-corrected template images are provided, and that internally \mathbf{A}_k is calculated by applying a given Contrast Transfer Function (CTF) of the micrograph to the template images. In this way, a fully CTF-corrected picking algorithm is obtained.

2.2. Particle sorting

The new sorting algorithm provides a fast way to identify incorrectly picked particles from the data. For every extracted particle image \mathbf{P}_i , the sorting program subtracts an associated template image \mathbf{A}_k in its given in-plane orientation ϕ . The resulting difference image is used to calculate an arbitrary number of statistical features. If \mathbf{P}_i corresponded to a true particle, the difference image should contain only background noise. If \mathbf{P}_i was incorrectly picked as a particle, then the difference image will contain features that cannot be described by background noise alone. As the sorting depends on the availability of a template image, it may be performed at any stage of the image processing when a 2D-template image has been assigned to each particle: i.e. directly after the auto-picking, or after any 2D or 3D classification or refinement. After 2D classification and auto-picking, \mathbf{A}_k will correspond to one of the K 2D templates used; after 3D classification or refinement, \mathbf{A}_k will correspond to projections of 3D template(s) in a given direction.

The sorting algorithm itself is similar to the one described in Scheres (2010), where one calculates a Z-score for an arbitrary number of features in each particle, and then sorts all particles based on the average of all these Z-scores. The features in the sorting algorithm in RELION are all based on the difference image between each particle and its aligned template. In particular, they comprise the mean, standard deviation, skewness and kurtosis of the difference image; as well as the standard deviation between the standard deviations calculated in four quadrants of the difference image. All these values are calculated within mask \mathbf{M}_i (Fig. 1C).

The resulting average Z-scores may then be used for displaying the particles in a sorted manner. Particles with a high average Z-score are often high-contrast false positives such as ice, protein

the new interface, as well as a schematic of the proposed semi-automated particle selection procedure that links together the auto-picking, sorting and reference-free 2D class averaging algorithms (Scheres, 2012b).

3. Experimental procedures

The particle selection procedure outlined above was tested on two previously published data sets. Firstly, it was tested on the keyhole limpet hemocyanin (KLH) data set that was used as a benchmark for testing particle picking algorithms in the so-called particle selection “bakeoff” (Zhu et al., 2004). This data set consists of 82 defocus pairs of micrographs at 2.2 Å/pixel that were acquired on a Philips CM200 microscope at 120 kV using a $2k \times 2k$ Tietz CCD camera and the Leginon system (Suloway et al., 2005). The first micrograph of each pair was recorded near to focus (NTF: 1 μm); the second one was recorded far from focus (FFF: 3 μm). The accumulated dose on each micrograph was approximately 10 electrons per Å². In the bakeoff, particles were selected either manually or (semi-) automatically using different computer programs, and all results were compared to each other. In the results described below, the particles selected by the algorithms in RELION are compared to the manually selected coordinates by Mouche, who only picked side views without any additional overlapping densities.

All calculations were performed on Dell M620 computing nodes of twelve 2.8 GHz Xeon cores and 48 Gb of RAM each.



The recommended semi-automated particle selection procedure in RELION-1.3 consists of at least five steps (Fig. 2B), each of which requires intervention by the user.

In the first two steps, the user manually selects particles from a subset of the micrographs and uses these particles to calculate reference-free 2D class averages. The number of particles necessary to calculate suitable templates ultimately depends on the SNR of the data, but 50–100 cryo-EM particles per template appears to be a useful guideline. In this case, from the first ten FFF micrographs, 264 particles were selected manually in step one. At this point, both top and side views were included. In the second step, reference-free 2D class averaging with 10 classes yielded two classes that were much larger than the others, representing a side view and a top view (Fig. 3A). These two class averages were selected to be used as templates for the particle picking algorithm. To prevent model bias, or “Einstein-from-noise” artefacts (also see Section 5), these templates were low-pass filtered to (strictly) 20 Å.

Fig. 2. Improved GUI and workflow. (A) Screenshot of the new GUI in RELION-1.3. (B) Proposed workflow for semi-automated particle selection in RELION-1.3. After CTFs have been estimated for all micrographs, the particle selection procedure consists of five steps (numbered 1–5), as explained in more detail in [Section 4](#).

In the third step, one performs the actual automated particle picking. At this point, there are two parameters to be optimised: the picking threshold, with higher values resulting in smaller, cleaner data sets; and the minimum inter-particle distance. The user-controlled angular sampling rate was kept fixed at 5 degrees for all calculations in this paper. Depending on the shape of the templates and the SNR in the micrographs, useful values of the threshold may vary from one data set to another. Fig. 3B shows how the recall, precision and false discovery rate of the auto-picking algorithm vary with the picking threshold. If TP is the number of particles that are both selected by RELION and Mouche, FP is the number of particles that is selected by RELION but not by Mouche, and FN is the number of particles that is selected by Mouche but not by RELION, then $\text{recall} = TP / (FN + TP)$, $\text{precision} = TP / (FP + TP)$, and false discovery rate (FDR) = $FP / (FP + TP)$, also see (Langlois and Frank, 2011). The minimum inter-particle distance often requires less optimisation, as values of around 60–90% of the particle diameter have been found to be useful in many cases. To accelerate the testing of different values for the picking threshold and the minimum inter-particle distance, the auto-picking program allows one to write out intermediate images with the optimal values (over all ϕ) of $S_{\phi,k}(\vec{r})$ for each template k . These so-called FOM maps are micrograph-sized images and writing many of them to disk may quickly become a bottle neck. Therefore, the parallel version of this program is disabled when writing out FOM maps, and it is recommended to write FOM maps only for a few representative micrographs of the data set. Once the FOM maps have been written to disk, reading them back in again and

picking peaks with different values of the picking threshold and the minimum inter-particle distance may be done in seconds. This allows one to optimise these values for the chosen representative micrographs. The optimised values may then be used to pick particles in the entire data set. In this run one no longer writes out FOM maps, and the program may be run in parallel to speed up the calculations. Two micrographs were selected for the parameter optimisation, which resulted in a picking threshold of 0.3 and a minimum inter-particle distance of 300 Å (which is approximately two-thirds of the diameter of mask M_i). Auto-picking of all 82 micrographs was done in 5 min using 41 cores in parallel, i.e. taking approximately 2.5 min per micrograph on each core. The RAM requirements for the auto-picking were approximately 400 Mb. Since the manually selected data set only contained side views, at this point all particles that were picked as a top view were discarded.

In the fourth step, one sorts the autopicked particles using the algorithm described in Section 2.2, one displays the particles sorted on the calculated average Z-scores, and one manually discards bad particles with the highest average Z-scores. Using a single core, the average Z-score calculation took less than a minute for the 1195 selected KLH side views. Manual inspection of the approximately 80 particles with the highest average Z-scores led to the removal of 44 particles, and was done in less than 2 min. Fig. 3C shows the 15 particles with the highest average Z-score as an example of what type of particles get discarded at this stage.

Finally, in the fifth step, the remaining particles are subjected to reference-free 2D class averaging, and manual inspection of the resulting classes is used to discard those particles that do not

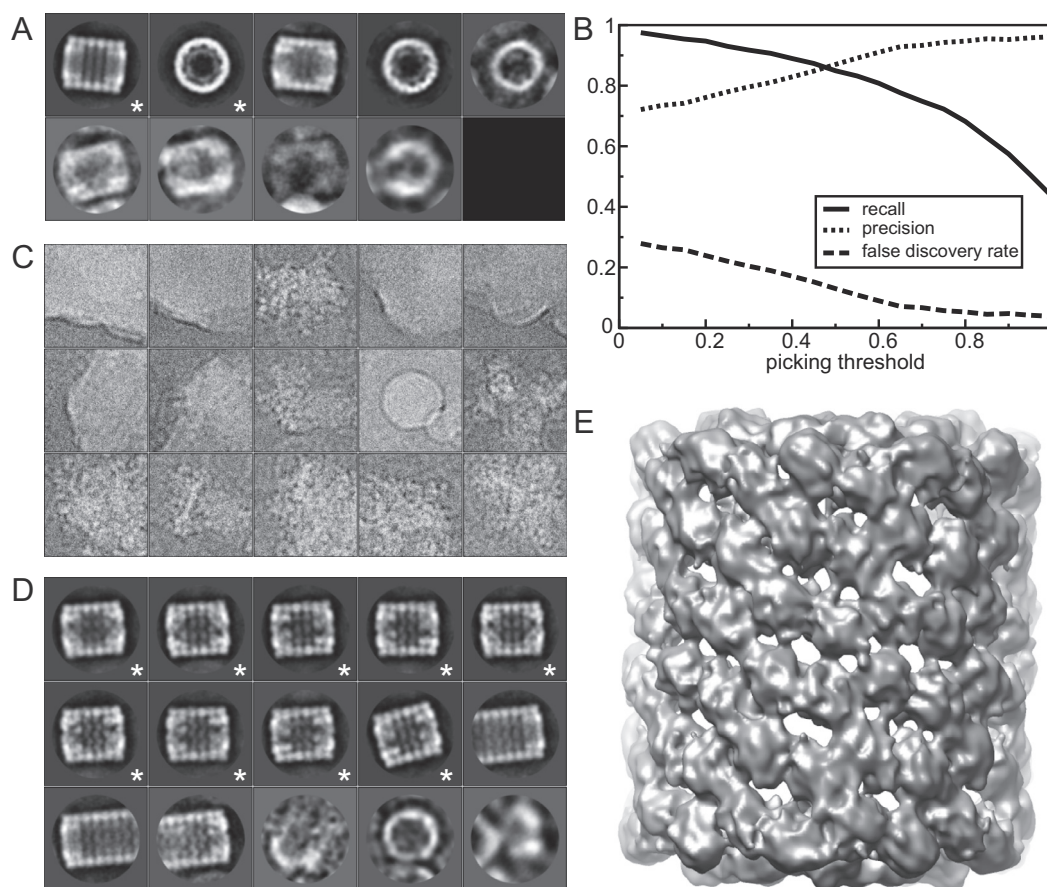


Fig. 3. Particle selection for the KLH data. (A) The ten reference-free class averages (ordered from larger to smaller classes) that were calculated from the manually selected particles. The two classes indicated with an asterisk were selected as templates for the auto-picking. (B) Curves of precision, recall and false discovery rate against the pick threshold. A picking threshold of 0.3 was chosen. (C) The 15 particles with the highest average Z-scores after sorting. (D) The 15 largest classes (ordered from larger to smaller) after 2D class averaging of the auto-picked particles. Particles assigned to the classes indicated with an asterisk were selected for subsequent 3D refinement. (E) 3D map after refinement of the semi-automatically selected particles from the combined near-to-focus (NTF) and far-from-focus (FFF) KLH data sets.

Table 1

Number of selected particles for each particle set, and the recall and false discovery rate (FDR) at the different steps in the RELION procedure for the far-from focus (FFF) and near-to-focus (NTF) data sets.

Data set	Nr. selected	Recall	FDR
Mouche	1042	–	–
FFF step 3	1195	0.92	0.20
FFF step 4	1151	0.91	0.18
FFF step 5	1048	0.92	0.10
NTF step 3	1261	0.91	0.25
NTF step 4	1173	0.91	0.19
NTF step 5	1064	0.90	0.12

average into good classes. In this case, the 2D class averaging with 25 classes was performed in 7 min using 36 cores in parallel, and 1048 particles were selected from 9 good classes (Fig. 3D). Depending on the sample, one may repeat the 2D class averaging step several times, and/or also perform 3D classification to further enrich the data set. Also, after any of these steps, one may recalculate the average Z-scores with the improved templates. For this paper, such additional classifications or sortings were not performed.

To further demonstrate the potential of the particle selection procedure for closer to focus images, it was also applied to the 82 NTF micrographs. The comparison of the selected particles after each step for both the FFF and the NTF micrographs with the Mouche coordinates is shown in Table 1. These results suggest that in some cases collecting defocus pairs may actually not be necessary when semi-automatically selecting particles in RELION. The combined 2112 particles from both the FFF and the NTF micrographs were used directly in a 3D refinement, using the preliminary 3D reconstruction that is distributed with these data as an initial model. This refinement yielded a reconstruction with a resolution of 11 Å according to the gold-standard FSC = 0.143 criterion (Scheres and Chen, 2012) (Fig. 3E).

4.2. Particle selection for the β -galactosidase data set

The same five-step particle selection procedure was also applied to the β -galactosidase data set. Manual picking in 5 micrographs yielded 2555 particles, which were used for a first 2D classification into 25 classes, and 10 of the resulting class averages were used as templates in the autopicking procedure (Fig. 4A). The picking threshold and the minimum inter-particle distance were set to 0.4 and 130 Å, respectively. Fig. 4B shows the performance of the picking algorithm for different thresholds. Auto-picking all 84 micrographs took approximately 1 h on 43 cores (i.e. taking approximately half an hour per micrograph on a single core), and took approximately 1.5 Gb of RAM.

In this case, the auto-picking algorithm was observed to give obviously false positives for some micrographs that showed high-variance artifacts like edges of the carbon holes or dust particles. Most probably this is caused by low values for $P(\mathbf{X}|\vec{r}, \mathbf{O})$ in the denominator of $R_{\phi,k}(\vec{r})$. Therefore, in experimental applications it may be beneficial to manually supervise the auto-picking results by deleting obviously false positives in the micrographs. The new display program makes this task relatively straightforward. However, in what follows such intervention was not performed in order to better reflect the recall and false discovery rate for the proposed algorithms with minimal user-interaction.

Sorting of the 52,495 particles that were picked automatically took 2 min on 8 cores. After visual inspection, the 4185 particles with the highest average Z-scores were discarded. Many of these were corresponding to the obviously false positives in the high-variance regions of the micrographs. The remaining particles were subjected to 2D classification with 200 classes, which took approximately 16 h on 64 cores. Good classes showed white particles with protein-like details on a black background, whereas many bad classes showed low-resolution blobs or images with many features in the background. From a total of 39 selected classes a final number of 42,755 particles were selected.

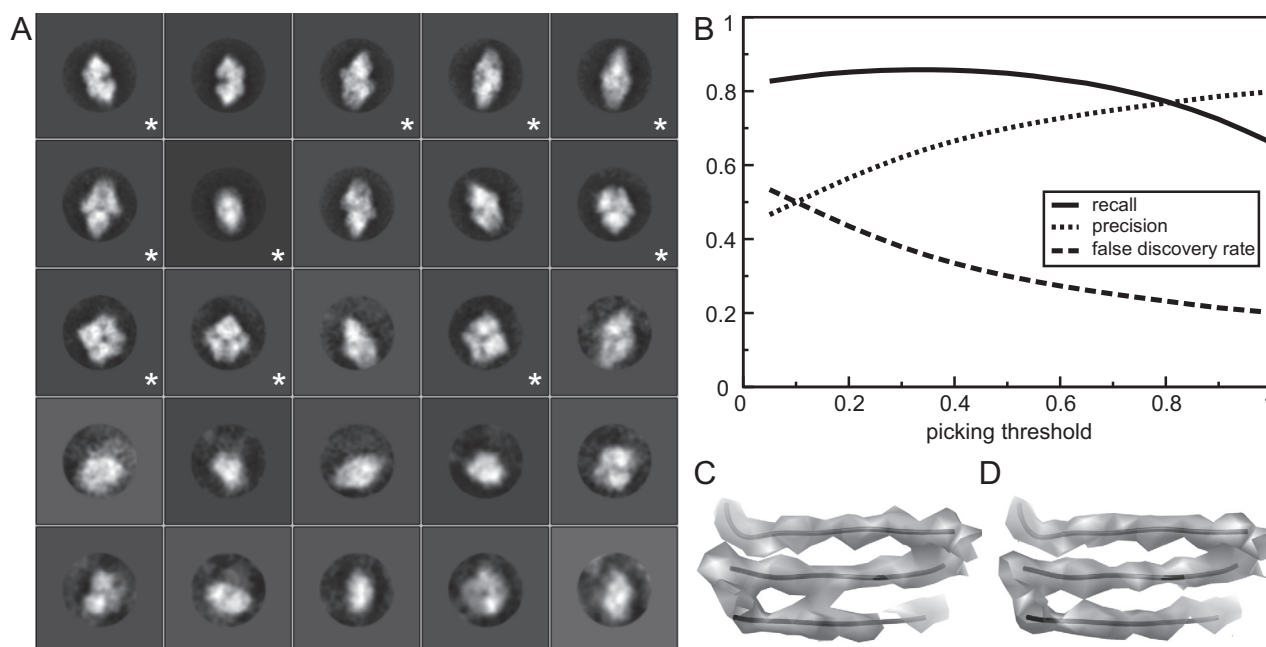


Fig. 4. Particle selection for the β -galactosidase data. (A) The 25 class averages that were calculated from the manually selected particles (ordered from larger to smaller class). The 10 class averages indicated with an asterisk were selected as templates for the auto-picking. (B) Curves of precision, recall and false discovery rate against the pick threshold. A picking threshold of 0.4 was chosen. (C) Map obtained after 3D refinement with the manually picked particles. (D) Map obtained after 3D refinement with the semi-automatically selected particles.

Table 2 shows the comparison between the particle sets at each of these steps and a manually selected data set of 40,863 particles by Richard Henderson. To further compare the two particle sets, a 3D refinement with the semi-automatically selected particle set was compared with an identical 3D refinement of the manually selected particles. In both cases a 60 Å low-pass filtered crystal structure of β -galactosidase (PDB PDB3I3I3E) (Dugdale et al., 2010) was used as an initial model. These refinements led to a resolution of 4.2 Å for both data sets, and the density map of the semi-automatically selected particles appears to be at least as good as the manual one (cf Fig. 4C and D).

5. Discussion

The new semi-automated particle picking and sorting algorithms, in combination with the selection of good classes after reference-free 2D class averaging lead to relatively high recalls and low false discovery rates when compared to alternative approaches in the original bakeoff study (Zhu et al., 2004), even when using the near-to-focus KLH micrographs. However, in the bake-off only particle picking approaches were compared, whereas in the approach described here lower thresholds may be used to avoid false negatives at the picking stage, and the sorting and 2D-classification algorithms may be used to remove false positives. For the β -galactosidase data set, refinement of a data set that was selected semi-automatically led to a map that was as least as good as a map obtained from manually picked particles. Therefore, the procedures proposed here may be an attractive alternative to the tedious process of manual particle selection. To facilitate future comparison with other automated picking approaches, the β -galactosidase micrographs and the manually selected coordinates from Richard Henderson were uploaded to the EMPIAR data base at the EMDB (entry EMPIAR-10017).

However, template-based particle picking does come with a potentially dangerous pitfall. As was pointed out recently in a series of comments on a controversial cryo-EM structure of the HIV-1 envelope glycoprotein trimer (Henderson, 2013; van Heel, 2013), using templates to select particles from noisy micrographs may be subject to strong template bias. This was termed “Einstein-from-noise”, in reference to the classical experiment where pure-noise images are aligned to a picture of Einstein in order to reproduce the Einstein image from averaging over noise only, see also (Shatsky et al., 2009). The template-based picking algorithm in RELION does not form an exception to this general problem. To illustrate this, the auto-picking algorithm was re-run on all β -galactosidase micrographs, but this time with a much lower threshold of 0.1. This led to 70,942 particles being picked, of which 62,230 were selected after sorting. Reference-free 2D class averaging with these particles revealed several artificial, “Einstein-from-noise” classes (indicated with an asterisk in Fig. 5A). Whereas good classes show high-resolution protein-like features, the artificial

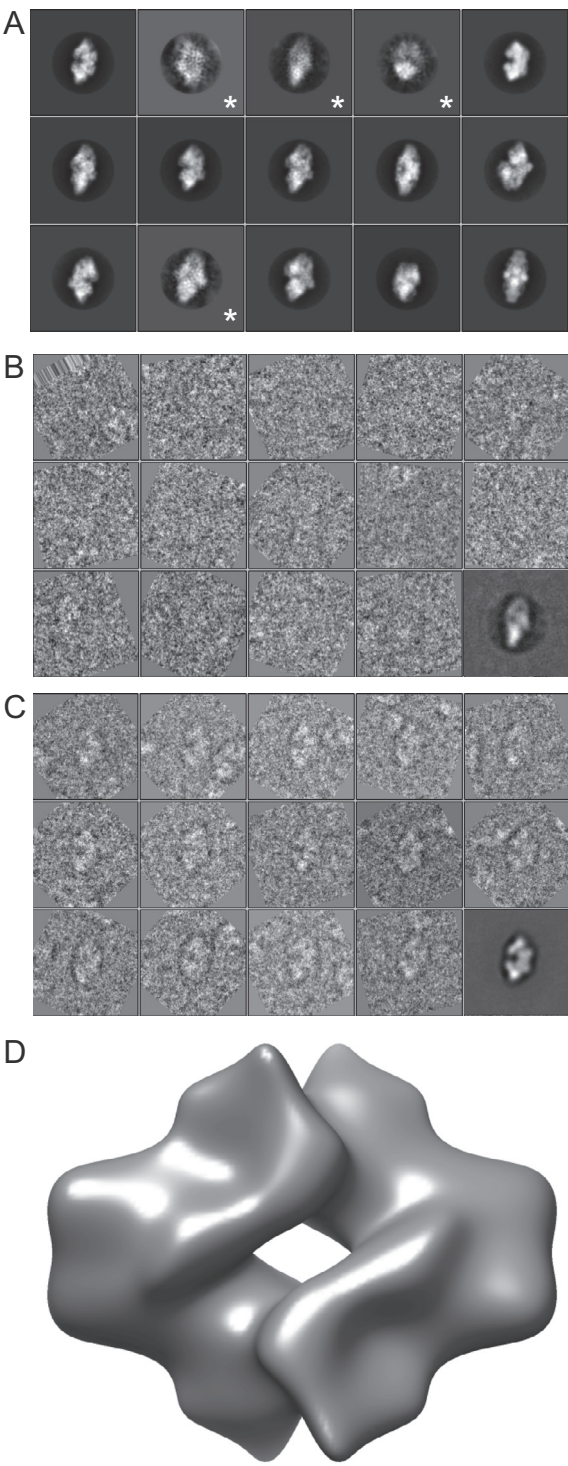


Fig.5. The “Einstein-from-noise” pitfall. (A) Class averages for the 15 largest classes (ordered from larger to smaller) after sorting and 2D class averaging of the auto-picked particles that were picked with a threshold of 0.1. Class averages indicated with an asteriks were identified as artificial classes caused by template bias (see Section 5). (B) Examples of particle images assigned to one of the artificial classes: the third class in A. No clear particles are visible. The lower-right image shows the average of all assigned particles in this class without any CTF correction. (C) Examples of particle images assigned to a good class: the first class in A. Particles are clearly visible. The lower-right image shows the average of all assigned particles in this class without any CTF correction. (D) 3D map obtained from 17,082 particles that were assigned to artificial classes.

Table 2
Number of selected particles for each particle set, and the recall and false discovery rate (FDR) at the different steps in the RELION procedure for the β -galactosidase data set.

Data set	Nr. selected	Recall	FDR
Henderson	40,863	–	–
Step 3	52,495	0.85	0.34
Step 4	48,310	0.81	0.31
Step 5	42,755	0.79	0.25

classes show merely low-resolution ghosts of the templates with superimposed high-resolution noise. Another noticeable difference between these classes is the angular accuracy that RELION estimates (Scheres, 2012b): for true classes this accuracy is often better than for artificial classes. Analysis of the individual particles that were assigned to the artificial classes shows that they are mostly empty particles (Fig. 5B). Moreover, averaging of these particles (without CTF-correction or masking) shows a black circle around the ghost image of the template. This black circle is the ghost image of the circular mask around the template image, which had slightly negative, i.e. black, values in the background. Particles assigned to good classes are clearly visible in individual images, and averaging over these does not show the black circle (Fig. 5C), which should not be mistaken for the typical black “aura” around an average that has not been CTF-corrected.

The extent of bias in template-based particle picking is not to be under-estimated. For the reference-free class averages shown in Fig. 5A, the second, third and fourth most populated 2D classes were identified as artificial classes. In order to be able to distinguish these false classes from the true ones, it is highly recommended to low-pass filter the templates used for auto-picking. For both the KLH and the β -galactosidase data sets, the auto-picking templates were filtered using a filter that strictly drops to zero beyond 20 Å. As selection of individual particles from noisy micrographs is mainly driven by relatively low frequencies, low-pass filtering of the templates is not expected to have a large impact on the results. But, the artificial class averages will not contain any features beyond that resolution, and true classes may then easily be distinguished from false ones if they contain useful features to higher resolution. Even if artificial classes were to be incorrectly included at this point, a subsequent gold-standard 3D refinement with only such particles would not be expected to reach resolutions beyond the low-pass filter used. To illustrate this, 17,082 particles that were assigned to artificial classes were subjected to 3D refinement, which yielded a featureless reconstruction at a reported resolution of 29 Å (Fig. 5D). It is important to note that the opposite is certainly not true. If the template images used in the auto-picking were not low-pass filtered, then model bias in the picked particles might still lead to spuriously high FSC values, regardless whether the FSC was calculated using gold-standard procedures, or whether some sort of reference-free 2D class averaging had been performed in between (van Heel, 2013). Therefore, one should not trust any reconstruction with a resolution that does not extend beyond the resolution of the templates that were used for the particle-picking.

Taken together, the new algorithms presented here in combination with an improved graphical-user interface provide a streamlined processing workflow for single-particle analysis in RELION. With the notable exception of initial model generation, all steps of a single-particle structure determination may now be performed from the RELION-1.3 interface: starting with initial micrograph inspection and ending at the generation of a final map that is suitable for atomic model building and refinement. As in previous releases, RELION-1.3 relies on CTFFIND3 (Mindell and Grigorieff, 2003) for CTF determination, and the new release employs a wrapper to RESMAP (Kucukelbir et al., 2014) to calculate local resolution variations in the final reconstruction. RELION-1.3 is open-source software and may be downloaded from <http://www2.mrc-lmb.cam.ac.uk/relion>. The semi-automated particle selection workflow presented here has already been useful in our own research on a 3.2 Å structure of the *Plasmodium falciparum* ribosome (Wong et al., 2014) and a 4.5 Å structure of the human γ -secretase complex (Lu et al., 2014). Hopefully, it will also contribute positively to the research of others who wish to use cryo-EM structure determination.

Acknowledgments

I am grateful to Shaoxia Chen, Greg McMullan and Richard Henderson for providing the β -galactosidase data, and to Richard Henderson for manually selecting particles from the β -galactosidase micrographs. I am also thankful to many members of the LMB, in particular Xiaochen Bai, Rafael Fernandez-Leiro and Chris Russo, for providing feedback on the RELION implementation, and to Jake Grimm and Toby Darling for assistance with computing. This work was funded by the UK Medical Research Council through Grant MC_UP_A025_1013.

References

- Allegretti, M., Mills, D.J., McMullan, G., Khlbrandt, W., Vonck, J., 2014. Atomic model of the f420-reducing [NiFe] hydrogenase by electron cryo-microscopy using a direct electron detector. *eLife* 3.
- Amunts, A., Brown, A., Bai, X.-C., Licer, J.L., Hussain, T., Emsley, P., Long, F., Murshudov, G., Scheres, S.H.W., Ramakrishnan, V., 2014. Structure of the yeast mitochondrial large ribosomal subunit. *Science* 343 (6178), 1485–1489.
- Bai, X.-C., Fernandez, I.S., McMullan, G., Scheres, S.H., 2013. Ribosome structures to near-atomic resolution from thirty thousand cryo-EM particles. *eLife* 2, e00461.
- Campbell, M., Cheng, A., Brilot, A., Moeller, A., Lyumkis, D., Veesler, D., Pan, J., Harrison, S., Potter, C., Carragher, B., Grigorieff, N., 2012. Movies of ice-embedded particles enhance resolution in electron cryo-microscopy. *Structure* 20 (11), 1823–1828.
- Cao, E., Liao, M., Cheng, Y., Julius, D., 2013. TRPV1 structures in distinct conformations reveal activation mechanisms. *Nature* 504 (7478), 113–118.
- Chen, J.Z., Grigorieff, N., 2007. SIGNATURE: a single-particle selection system for molecular electron microscopy. *J. Struct. Biol.* 157 (1), 168–173.
- Chen, S., McMullan, G., Faruqi, A.R., Murshudov, G.N., Short, J.M., Scheres, S.H.W., Henderson, R., 2013. High-resolution noise substitution to measure overfitting and validate resolution in 3d structure determination by single particle electron cryomicroscopy. *Ultramicroscopy* 135, 24–35.
- Dugdale, M.L., Dymianiw, D.L., Minhas, B.K., D'Angelo, I., Huber, R.E., 2010. Role of met-542 as a guide for the conformational changes of phe-601 that occur during the reaction of β -galactosidase (*Escherichia coli*). *Biochem. Cell Biol.* 88 (5), 861–869.
- Fernandez, I.S., Bai, X.-C., Hussain, T., Kelley, A.C., Lorsch, J.R., Ramakrishnan, V., Scheres, S.H.W., 2013. Molecular architecture of a eukaryotic translational initiation complex. *Science (New York, N.Y.)* 342 (6160), 1240585.
- Grigorieff, N., Harrison, S.C., 2011. Near-atomic resolution reconstructions of icosahedral viruses from electron cryo-microscopy. *Curr. Opin. Struct. Biol.* 21 (2), 265–273.
- Hall, S.R., 1991. The STAR file: a new format for electronic data transfer and archiving. *J. Chem. Inf. Comput. Sci.* 31 (2), 326–333.
- Henderson, R., 2013. Avoiding the pitfalls of single particle cryo-electron microscopy: Einstein from noise. *Proc. Natl. Acad. Sci. USA* 110 (45), 18037–18041.
- Kucukelbir, A., Sigworth, F.J., Tagare, H.D., 2014. Quantifying the local resolution of cryo-EM density maps. *Nat. Methods* 11 (1), 63–65.
- Langlois, R., Frank, J., 2011. A clarification of the terms used in comparing semi-automated particle selection algorithms in cryo-EM. *J. Struct. Biol.* 175 (3), 348–352.
- Langlois, R., Pallesen, J., Ash, J.T., Nam Ho, D., Rubinstein, J.L., Frank, J., 2014. Automated particle picking for low-contrast macromolecules in cryo-electron microscopy. *J. Struct. Biol.* 186 (1), 1–7.
- Li, X., Mooney, P., Zheng, S., Booth, C.R., Braumfeld, M.B., Gubbens, S., Agard, D.A., Cheng, Y., 2013. Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM. *Nat. Methods* 10 (6), 584–590.
- Lu, P., Bai, X.-C., Ma, D., Xie, T., Yan, C., Sun, L., Yang, G., Zhao, Y., Zhou, R., Scheres, S.H.W., Shi, Y., 2014. Three-dimensional structure of human-secretase. *Nature* 512 (7513), 166–170.
- McMullan, G., Chen, S., Henderson, R., Faruqi, A., 2009. Detective quantum efficiency of electron area detectors in electron microscopy. *Ultramicroscopy* 109 (9), 1126–1143.
- Mindell, J.A., Grigorieff, N., 2003. Accurate determination of local defocus and specimen tilt in electron microscopy. *J. Struct. Biol.* 142 (3), 334–347.
- Nicholson, W.V., Glaeser, R.M., 2001. Review: automatic particle detection in electron microscopy. *J. Struct. Biol.* 133 (23), 90–101.
- Roseman, A.M., 2003. Particle finding in electron micrographs using a fast local correlation algorithm. *Ultramicroscopy* 94 (3–4), 225–236.
- Roseman, A.M., 2004. FindEM – a fast, efficient program for automatic selection of particles from electron micrographs. *J. Struct. Biol.* 145 (1–2), 91–99.
- Scheres, S.H., 2014. Beam-induced motion correction for sub-megadalton cryo-EM particles. *eLife* 3, e03665.
- Scheres, S.H.W., 2010. Classification of structural heterogeneity by maximum-likelihood methods. In: *Cryo-EM, Part B: 3-D Reconstruction. Methods in Enzymology*, vol. 482. Academic Press, pp. 295–320.

- Scheres, S.H.W., 2012a. A bayesian view on cryo-EM structure determination. *J. Mol. Biol.* 415 (2), 406–418.
- Scheres, S.H.W., 2012b. RELION: implementation of a bayesian approach to cryo-EM structure determination. *J. Struct. Biol.* 180 (3), 519–530.
- Scheres, S.H.W., Chen, S., 2012. Prevention of overfitting in cryo-EM structure determination. *Nat. Methods* 9 (9), 853–854.
- Scheres, S.H.W., Gao, H., Valle, M., Herman, G.T., Eggermont, P.P.B., Frank, J., Carazo, J.M., 2007. Disentangling conformational states of macromolecules in 3d-EM through likelihood optimization. *Nat. Methods* 4 (1), 27–29.
- Shatsky, M., Hall, R.J., Brenner, S.E., Glaeser, R.M., 2009. A method for the alignment of heterogeneous macromolecules from electron microscopy. *J. Struct. Biol.* 166 (1), 67–78.
- Sorzano, C.O.S., de la Fraga, L.G., Clackdoyle, R., Carazo, J.M., 2004. Normalizing projection images: a study of image normalizing procedures for single particle three-dimensional electron microscopy. *Ultramicroscopy* 101 (2–4), 129–138.
- Sorzano, C.O.S., Recarte, E., Alcorlo, M., Bilbao-Castro, J.R., San-Martín, C., Marabini, R., Carazo, J.M., 2009. Automatic particle selection from electron micrographs using machine learning techniques. *J. Struct. Biol.* 167 (3), 252–260.
- Suloway, C., Pulokas, J., Fellmann, D., Cheng, A., Guerra, F., Quispe, J., Stagg, S., Potter, C.S., Carragher, B., 2005. Automated molecular microscopy: the new legion system. *J. Struct. Biol.* 151 (1), 41–60.
- Tang, G., Peng, L., Baldwin, P.R., Mann, D.S., Jiang, W., Rees, I., Ludtke, S.J., 2007. EMAN2: an extensible image processing suite for electron microscopy. *J. Struct. Biol.* 157 (1), 38–46.
- van Heel, M., 2013. Finding trimeric HIV-1 envelope glycoproteins in random noise. *Proc. Natl. Acad. Sci. USA* 110 (45), E4175–E4177.
- Vinothkumar, K.R., McMullan, G., Henderson, R., 2014. Molecular mechanism of antibody-mediated activation of galactosidase. *Structure* 22 (4), 621–627.
- Voorhees, R.M., Fernandez, I.S., Scheres, S.H.W., Hegde, R.S., 2014. Structure of the mammalian ribosome-Sec61 complex to 3.4 resolution. *Cell* 157 (7), 1632–1643.
- Voss, N.R., Yoshioka, C.K., Radermacher, M., Potter, C.S., Carragher, B., 2009. DoG picker and TiltPicker: software tools to facilitate particle selection in single particle electron microscopy. *J. Struct. Biol.* 166 (2), 205–213.
- Wong, W., Bai, X.-C., Brown, A., Fernandez, I.S., Hanssen, E., Condrón, M., Tan, Y.H., Baum, J., Scheres, S.H., 2014. Cryo-EM structure of the *Plasmodium falciparum* 80S ribosome bound to the anti-protozoan drug emetine. *eLife*, e03080.
- Zhu, Y., Carragher, B., Glaeser, R.M., Fellmann, D., Bajaj, C., Bern, M., Mouche, F., Haas, F.d., Hall, R.J., Kriegman, D.J., Ludtke, S.J., Mallick, S.P., Penczek, P.A., Roseman, A.M., Sigworth, F.J., Volkman, N., Potter, C.S., 2004. Automatic particle selection: results of a comparative study. *J. Struct. Biol.* 145 (1–2), 3–14.