Locating Macromolecular Assemblies in Cells by 2D Template Matching with *cis*TEM

Bronwyn A. Lucas[1#], Benjamin A. Himes[2#], Liang Xue[3,4], Timothy Grant[1†], Julia Mahamid[3], Nikolaus Grigorieff[2*]

[1] Howard Hughes Medical Institute, Janelia Research Campus, Ashburn, VA, United States

[2] Howard Hughes Medical Institute, RNA Therapeutics Institute, The University of Massachusetts Medical School, Worcester, MA, United States

[3] Structural and Computational Biology Unit, European Molecular Biology Laboratory (EMBL), Heidelberg, Germany

[4] Collaboration for joint PhD degree between EMBL and Heidelberg University, Faculty of Biosciences

[#] Equal contribution

[*] For correspondence: niko@grigorieff.org

[†] Present address: John W and Jeanne M Rowe Center for Research in Virology, Morgridge Institute for Research, Madison, WI, USA; Department of Biochemistry, University of Wisconsin-Madison, Madison, WI, USA

Abstract

Over the last decade, single-particle electron cryo-microscopy has become one of the main techniques contributing to the growing library of high-resolution structures of macromolecules and their assemblies. For a full understanding of molecular mechanisms, however, it is important to place them into the broader context of a cell. Traditionally, this context can be visualized in 3D by electron cryo-tomography, and more recently, has also been studied by template matching of 2D images of cells and viruses. A current limitation of the latter approach is the high computational cost that limits the throughput and widespread adoption of this method. We describe here a GPU-accelerated implementation of 2D template matching in the image processing software *cis*TEM that allows for easy scaling and improves the accessibility of this approach. We apply 2D template matching to identify ribosomes in images of frozen-hydrated *Mycoplasma pneumoniae* cells and demonstrate that it can function as a versatile tool for *in situ* visual proteomics and *in situ* structure determination. We compare the results with 3D template matching of tomograms acquired on identical sample locations. We identify strengths and weaknesses of both techniques which offer complementary information about target localization and identity.

Introduction

A major goal in structural biology over the last 70 years has been to understand the molecular mechanisms of biological processes that occur inside cells by studying the underlying proteins and their assemblies, collectively referred to here as complexes, and the many biochemical reactions catalyzed by them. X-ray crystallography and electron microscopy, in particular electron cryo-microscopy (cryo-EM), have generated high-resolution density maps of these complexes that could be interpreted by atomic models, allowing a detailed description of molecular mechanisms (Berman et al., 2002). Of these two structural techniques, cryo-EM has emerged as the more versatile, being applicable to 2D and 3D crystals, helical assemblies, non-crystalline material (single particles), cells and tissues. All cryo-EM techniques have greatly benefited from technical advances over the last decade, including direct electron detectors with improved image quality and speed allowing for movie data collection (Brilot et al., 2012; Campbell et al., 2012; Li et al., 2013), more powerful computers, and new image processing algorithms (Lyumkis et al., 2013; Scheres, 2012), leading to the so-called "resolution revolution" (Kuhlbrandt, 2014). Using single-particle cryo-EM, it has become easier to obtain near-atomic resolution structures of a broad range of complexes, following earlier resolution breakthroughs with 2D crystals, helical assemblies, and highly symmetrical virus particles (Grigorieff & Harrison, 2011; Hasler et al., 1998; Sachse et al., 2008). A particular strength of the single particle technique, as opposed to crystallography, is its ability to sort particles according to their 3D structure using image classification techniques, and to deliver structures of several distinct states of a complex observed under near-native conditions (Abeyrathne et al., 2016). This structural "deconvolution revolution" has been equally important in driving the success of cryo-EM, and often yields a more complete picture of a mechanism than can be obtained from the best resolved structure of a single state.

While the list of high-resolution structures is rapidly growing, a full understanding of molecular mechanisms, and of the role these complexes play within the host organism, requires the broader context of the cell (Alberts, 1998). This context is necessary to understand the functional coupling of different complexes, for example by transient interaction between them, by efficient shuttling of substrates and products, or by a common regulatory mechanism (e.g., transcription-translation coupling (O'Reilly et al., 2020), turnover of nuclear pore complexes (Allegretti et al., 2020) or co-translational protein transport (Braunger et al., 2018)). More native conditions can be maintained in a sample by employing rapid purification of cell lysate (Behrmann et al., 2015). The rapid purification aims to better preserve transient interactions and states compared with traditional purification techniques, but at the expense of purity. The "impure" sample can then be purified *in silico* by modern image classification techniques that are now commonplace in cryo-EM. To achieve close-to native conditions, cryo-EM can also be used to image molecules and complexes directly inside frozen-hydrated cells and tissue at high resolution, making it one of the most promising approaches to add cellular context to structures that have been visualized *in vitro*. To date, the most developed cryo-EM technique to visualize the "molecular sociology" of

cells (Beck & Baumeister, 2016) is electron cryo-tomography (cryo-ET) (Oikonomou & Jensen, 2017).

Similar to single-particle cryo-EM, cryo-ET has seen significant development over the past decade (Wan & Briggs, 2016), including the extension of subtomogram averaging to sub-nm resolution (Schur et al., 2013). By adapting techniques from single-particle cryo-EM to work on small 3D volumes within a tomogram, subtomogram averaging can boost the resolution of reconstructions of complexes visualized *in situ*, thus avoiding the difficulties related to sample purification (A. Bartesaghi et al., 2008; Bohm et al., 2000; Winkler et al., 2009). Subtomogram averaging also benefits from image classification, yielding detailed 3D volumes of cells and sections that show the distribution and functional states of prominent complexes, such as ribosomes and proteasomes, in relation to compartments, membranes, filaments and other cellular structures (Asano et al., 2015; Cai et al., 2018; Mahamid et al., 2016). Subtomogram averaging and the annotation of 3D volumes requires the correct identification of density features in the tomogram, which is still one of the major bottlenecks for the method (Pfeffer & Mahamid, 2018; Zhang, 2019). This can be accomplished by matching with a 3D template density (Frangakis et al., 2002), which may be obtained from a small preliminary set of manually selected subtomograms, or from a different experiment, for example single-particle cryo-EM. 3D template matching (3DTM) has been particularly successful in the study of large cytoplasmic complexes, such as ribosomes, proteasomes and chaperonins (Eibauer et al., 2012; Pfeffer et al., 2018) and even proteasomes inside the nucleus, close to nuclear pore complexes (Albert et al., 2017). However, the nucleus as a whole has proven to be more challenging, due to molecular crowding and the higher density produced by nucleic acids in cryo-EM images (Spahn et al., 2000). This points to a more fundamental problem in cryo-ET: density annotation depends critically on recognizing the overall shape, i.e., the particle envelope, which may be obscured in regions of the cell with high molecular density, or in regions where several complexes connect to form a single continuous density (Grünewald et al., 2002). 3DTM is also susceptible to false-positive detections as the limited resolution of a tomogram (~20 Å (Frank, 2006)) makes correlation coefficients relatively insensitive to internal structure.

Recently, we described a 2D template matching (2DTM) technique (Rickgauer et al., 2017) that may overcome some of the limitations of 3DTM. It matches projections of 3D templates to features found in single-exposure (2D) images of nominally untilted specimens. Avoiding multiple exposures and high specimen tilt angles helps preserve the high-resolution signal in these 2D images (Brilot et al., 2012), and therefore, 2DTM can utilize this signal to detect complexes with high specificity, as well as high angular and positional in-plane accuracy (x,y coordinates). This added signal comes at the expense of increased structural noise in the images due to overlapping density from other molecules in the cell, and a relatively large error in localizing the depth of the targets within the sample (z coordinate). 2DTM also requires a fine-grained angular search composed of millions of reference projections and correlation maps to be calculated, therefore making the computational workload of a 2D template search relatively high

compared to a more coarse-grained search that is normally done with 3DTM. We describe here an implementation of 2DTM in the software package *cis*TEM (Grant et al., 2018), providing a user-friendly graphical interface and GPU-acceleration to speed up computation. We applied 2DTM to a set of images of *Mycoplasma pneumoniae* using a bacterial 50S large ribosomal subunit (LSU) as a template. *M. pneumoniae* cells are small and electron-transparent and do not require additional sample thinning (Kühner et al., 2009), for example via controlled cell lysis (Fu et al., 2014) or more commonly by using focused ion beam milling (Marko et al., 2007; Rigort et al., 2012; Strunk et al., 2012). We show that 2DTM is a versatile method with potentially broad applications for both *in situ* visual proteomics and *in situ* structure determination, including *de novo* structure determination. We compare the results of 2DTM and 3DTM directly by collecting a subsequent tomogram of the same area analyzed by 2DTM. We show that 2DTM has improved specificity relative to 3DTM and comparable sensitivity given a sufficiently thin sample.

Results

GPU-accelerated 2DTM implemented in *cis*TEM

The large search space required for 2DTM makes this method computationally demanding; a single 1850 x 1850 pixel image required 1000 CPU-hours for a search in the proof-of-principle MATLAB implementation (Rickgauer et al., 2017). To make 2DTM accessible to more users and a broader range of biological questions, we implemented 2DTM in *cis*TEM (Grant et al., 2018) (**Figure 1a**) using C++ and achieved a roughly 23x speed-up compared to the MATLAB implementation. The core 2DTM algorithm is unchanged from its original description as depicted in the flowchart in **Figure 1 - figure supplement 1**. The *cis*TEM implementation may be run as a standalone from the command line interface, or alternatively, using the *cis*TEM graphical user interface (GUI). In addition to a user-friendly interface, the GUI affords several advantages: Firstly, project metadata and image-specific information, such as CTF estimation, are tracked in a database; secondly, the search is easily divided over many CPUs or computers using *cis*TEM's MPI-like dispatch via run-profiles; and thirdly, the results are displayed in an interactive manner allowing for easy interpretation and comparison across multiple search conditions. A screenshot showing the results of a template matching search in the *cis*TEM GUI is shown in **Figure 1a**. The GUI displays the image searched (**Figure 1b**), maximum intensity projection (MIP) (**Figure 1c**) and the plotted results (**Figure 1d**), enabling rapid qualitative interrogation of the template matching results.

Even with these improvements, searching a more typically sized image, like that from a Gatan K3 detector (5760 x 4092 pixels) requires ~7,100 CPU-hours when searching 13 focal planes, which are needed to resolve overlapping densities in a 150 – 200 nm thick sample. Further substantial gains in the CPU-based code are not likely, given that about 85% of the computation for 2DTM is spent on calculating fast Fourier transforms (FFTs) using the already highly

optimized Intel Math-Kernel library (MKL). To circumvent these limitations, we developed a GPU image class in *cis*TEM that has a subset of the same underlying member variables and methods as the corresponding CPU image class. Both implementations use positions on the Euler sphere to divide the search space (**Figure 2a**). The GPU implementation exploits further parallelism via threading, which combined with CUDA streams allows for multiple kernels to execute on the GPU simultaneously (**Figure 2b**). This in turn allows us to create a dynamic load balancing that results in full occupancy of the GPU over a wide range of problem sizes.

To evaluate the performance improvements of our GPU 2DTM implementation relative to the CPU 2DTM implementation, we compared two high-end GPUs (Nvidia GV100) against two high-end 28-core CPUs (Intel Xeon Platinum 8280) installed in the same general-purpose workstation, with all other hardware and variables unchanged. By this metric, the GPU-accelerated implementation of 2DTM achieved an 8.5x speed up relative to the CPU-only implementation (**Figure 2c**). Using IEEE 754 half-precision floating point values (FP16) for the arrays used to track search statistics, namely the pixel-wise sum and sum-of-squares over all orientations, resulted in further acceleration and a reduced memory footprint. The total speed-up was 10.5x (**Figure 2c**). The algorithm scales nearly linearly with the number of GPUs used, as shown for different NVIDIA GPU architectures in **Figure 2d**. The relative computational cost of each step of the GPU-accelerated 2DTM inner loop algorithm is detailed in **Figure 2 - figure supplement 1**.

To avoid cumulative rounding errors at the reduced precision of FP16, we implemented a cascading summation where the sums and sum-of-squares are accumulated over ten search positions. Every tenth search, the results are accumulated in 32-bit single precision into a separate array. Every one-hundredth search, the results from the lower tier are accumulated, and so on, resulting in similar-sized numbers being added. An additional consideration for the sum-of-squares array was needed as the smallest positive number that can be represented by FP16 without being rounded to zero is 2^{-14}. To prevent these subnormal numbers from being flushed to zero, we temporarily multiply the array by a factor of 10^3 until conversion to higher precision farther down the cascade.

Detection of 50S ribosomal subunits by 2DTM

The improved speed and increased throughput of 2DTM enabled us to perform an initial screen of template and search parameters that affect target detection *in situ*. To this end, we collected 2D images of plunge-frozen *M. pneumoniae*. *M. pneumoniae* lacks a cell wall and can be less than 200 nm thick, making it sufficiently thin to allow TEM imaging of whole cells in ice without the need for thinning (Kühner et al., 2009). We performed 2DTM using the *M. pneumoniae* 50S (PDB: in progress; EMDB: 11999) (Tegunov et al., 2021) as a template (**Figure**

3a) and identified 6,558 50S large ribosomal subunits in 220 2D images. This search did not distinguish between isolated subunits and subunits bound to the 30S small ribosomal subunit.

The output from a 2DTM search are SNR values that depend on both the agreement between the template and the target as well as noise in the image (Sigworth, 2004). The noise is predominantly shot noise, as well as background generated by molecules and other cellular material overlapping the target in projection. A target is detected when the SNR value exceeds a threshold at which the average number of false positives per search is set to a user-specified number, usually one (Rickgauer et al., 2017). To improve the match between template and recorded signal, the template can be low-pass filtered to approximate physical blurring of the image due to radiation damage and beam-induced motion, as well as discrepancies between the target in the template, for example due to conformational differences. To low-pass filter the template, we applied a range of B-factors to the initial map generated using *pdb2mrc* ((Tang et al., 2007), see Methods) and found that the average SNR was maximized using a B-factor of about 85 $Å^2$ (**Figure 3 - figure supplement 1a**). Using this value, we found that searching an image composed of exposure-weighted frames (32 $e^-/Å^2$) (Grant & Grigorieff, 2015) increased the average SNR compared to using only the first 8 frames (12.8 $e^-/Å^2$) with or without exposure weighting (**Figure 3 - figure supplement 1b**). We conclude that including additional frames with exposure weighting does not diminish the 2DTM SNR, indicating the constructive contribution of lower-resolution features to the detected signal. Consistent with SNR values being sensitive to defocus errors (Rickgauer et al., 2017, 2020), searching each image with templates sampling a defocus range of 2400 Å in 200 Å steps (13 defocus planes) increased the number of detected targets by ~40% in a 100 nm thick sample and ~400% in a 220 nm thick sample (**Figure 3 - figure supplement 1c**). In addition to increasing the number of detected targets, the defocus search also provides a rough estimate of the out-of-plane position of each 50S. Earlier simulations predicted that protein background would result in reduced target detection in images collected at higher defocus (Rickgauer et al., 2017). In the present experiment, we did not observe a consistent relationship between mean SNR and image defocus over the range of ~500 to ~2200 nm underfocus (**Figure 3 - figure supplement 1d**), suggesting that detection of complexes as large as 50S ribosomal subunits is not prevented by collecting images further from focus.

De novo structure determination using 2DTM

In our current 2DTM implementation, each detected target is assigned an x,y location, three Euler angles and a defocus value (z coordinate) (**Figure 3a**). These parameters can be used to calculate a 3D reconstruction using standard single-particle methods. It is well known that a reconstruction calculated from particles that were identified using a template may suffer from significant template bias (Henderson, 2013), reproducing only features of the template and not

the targets. However, as shown by Rickgauer et al. (2017) and discussed below, applying an absolute (rather than relative) threshold based on a known noise distribution, also limits template bias and new features not present in the template may be visible in a reconstruction derived from detected targets. To test this, we calculated a reconstruction using the results from searching 220 images of *M. pneumoniae* cells (see above), selecting targets only from the best images that had more than nine detected targets and at least one target with an SNR value above 9. Using the 5,080 targets that met these criteria, out of the total of 6,558 detected targets, we calculated a 3D reconstruction (**Figure 3b**). The 20 Å-filtered reconstruction reproduces the 50S template as expected, but also shows clear, albeit weaker density for the 30S subunit that was not present in the template, and thus cannot be due to template bias (**Figure 3b,c**).

The difference map calculated using *diffmap* (Grigorieff, 2021b), between the *M. pneumoniae* 50S template and the 3D reconstruction shows density in regions of the 50S that have been shown to be flexible from *in vitro* reconstructions of the ribosome, specifically, around the L1 stalk and the L7/L12 stalk (**Figure 3c**). The *Escherichia coli* L1 stalk moves ~45 – 60Å relative to 50S during translocation (reviewed in (Ling & Ermolenko, 2016)), and the C-terminal domain of L10 and N-terminal domain of L12 are known to be flexible (Diaconu et al., 2005). Moreover, we observed density in each of the three tRNA binding sites on the small subunit, which were not present in the template (**Figure 3d**). The density is consistent with tRNAs representing multiple sates and may also reflect the binding of other factors, such as translational GTPases in the A-site. It is therefore likely that the calculated reconstruction represents an average of different conformations adopted by the ribosome *in vivo*, as expected in actively growing cells. This average differs therefore from the single conformation used for the template, giving rise to the features observed in the difference map. Smaller features in the difference map may also correspond to noise in the reconstruction, as well as features that result from inaccuracies in the resolution scaling of the template against the reconstruction before subtraction. More accurate scaling may be achieved by scaling according to local resolution estimates, which were not obtained here, as well as masking to exclude parts in both maps that do not overlap.

2DTM reveals species-specific structural features

To test the specificity of 2DTM and to evaluate whether species-specific structural differences would preclude detection by 2DTM, we searched the same 220 images of *M. pneumoniae* with a 50S template derived from *Bacillus subtilis* (PDB: 3J9W). *B. subtilis* 50S is structurally closely related to *M. pneumoniae* 50S but displays differences in rRNA sequence and protein composition (Grosjean et al., 2014; Sohmen et al., 2015) (**Figure 4a**). Comparing the two models by Fourier shell correlation (FSC) and taking 0.5 as a similarity threshold (**Figure 4 - figure supplement 1a**) suggests that the effective resolution of the *B. subtilis* template is limited to ~4.7Å. Using *B. subtilis* 50S as a template, we identified 2,874 50S locations, less than half

that were identified using *M. pneumoniae* template (**Figure 4b**), and with significantly lower 2DTM SNR values (P <0.0001, Mann-Whitney U test, **Figure 4c**). Again, to limit targets to the best images, we included only images that had more than two detected targets and at least one target with an SNR value above 9. Using the 1,172 targets from these images, we calculated a 3D reconstruction (**Figure 4d**). As before, the 20 Å-filtered reconstruction reproduces the 50S template with clear density for the 30S subunit and shows features consistent with translating ribosomes (**Figure 4d-e**). However, unlike before, the 3D reconstruction also shows extensive differences in the 50S relative to the *B. subtilis* 50S template (**Figure 4e**). Since this reconstruction was generated using ~5 fold fewer particles, some of the additional density likely reflects an expected higher level of noise. Despite this, further examination of the 3D reconstruction revealed that some of the features visible in the 50S density that deviate from the *B. subtilis* template likely result from *M. pneumoniae* specific features (O'Reilly et al., 2020) (**Figure 4e, Figure 4 - figure supplement 1c**). Specifically, the 3D reconstruction showed density consistent with *M. pneumoniae* specific C-terminal extensions of protein L29, and L22 and protein L9 (**Figure 4f, Figure 4 - figure supplement 1c**) (O'Reilly et al., 2020). The latter was absent from the *B. subtilis* model despite being encoded in the *B. subtilis* genome (Sohmen et al., 2015). The 3D reconstruction also lacked density for protein uL30, which is present in the *B. subtilis* template, but absent from the *M. pneumoniae* genome (Grosjean et al., 2014) (**Figure 4g**). Moreover, the 3D reconstruction shared additional unattributed features with the previously determined reconstruction (**Figure 3b**) that do not derive from either template (**Figure 4 - figure supplement 1c**). We conclude that a *B. subtilis* 50S template can be used as a homology model to identify *M. pneumoniae* ribosomes and distinguish species-specific features, and that including high-resolution signal in 2DTM does not overly bias the 3D reconstruction when a high SNR threshold is used. These results further demonstrate the reliability of the 2DTM and its potential to directly resolve complex structure *in situ*.

Detection of 50S ribosomal subunits by 3DTM

Cryo-ET combined with 3DTM is currently one of the most commonly used approaches for locating molecules within cells using available structural information. Rather than using high-resolution templates to search 2D images, as is done in 2DTM, 3DTM locates molecules in tomograms using templates filtered to 30 Å or lower (Himes & Zhang, 2018). To compare the detection of 50S ribosomal subunits by 2DTM and 3DTM, we collected 19 2D images of *M. pneumoniae* cells followed by tomograms of an overlapping area. These tomograms were collected after the 32 $e^-/Å^2$ exposure for images used with 2DTM, using standard protocols (**Figure 5a**). The pre-exposure of 32 $e^-/Å^2$ affects the signal in the tomograms at higher resolution but is expected to have only a small effect in the resolution range relevant for 3DTM, i.e., 20 Å and lower (Grant & Grigorieff, 2015). To identify 50S subunits with 3DTM, we followed established protocols using PyTOM and a 30 Å low-pass filtered 50S template

generated in a previous study (O'Reilly et al., 2020). The found targets were ranked according to their cross-correlation score and the top 600 hits were selected for each tomogram, followed by alignment and 3D classification with *RELION 3.0.8* (Zivanov et al., 2018). The selection of the top 600 hits ensured that more than 90% of potential targets in one tomogram, including free 50S and 50S in assembled 70S, were included. Combining *RELION* classification results with cross-correlation ranking (**Figure 5 - figure supplement 1**) shows that the hits with the lowest scores contained fewer than 10% of 50S and 70S targets.

We aligned the coordinates of the 652 50S subunits identified by 2DTM with 983 50S subunits identified by 3DTM (see Methods), both containing 50S targets corresponding to single 50S subunits and 70S ribosomes (**Figure 5b**). Within the area common to the 2D images and tomograms, we identified 576 2DTM targets that were within a 100-Å distance in the x,y plane, and within 20° in each Euler angle of a 3DTM target. These limits included ~95% of all paired targets (**Figure 5c-d**). The paired targets represent ~90% of all 2DTM targets and ~60% of all 3DTM targets in this area (**Figure 5c**, upper). We found that the proportion of 2DTM targets with a corresponding 3DTM target was consistent across the images examined (**Figure 5d**). In contrast, the proportion of the 3DTM targets detected in the 2DTM search was variable and showed a negative correlation with sample thickness (**Figure 5e**). This is consistent with our and prior observations that 2DTM is particularly sensitive to sample thickness (**Figure 5 - figure supplement 2**) (Rickgauer et al., 2017, 2020), which likely contributes to the high false negative rate of 2DTM relative to 3DTM.

The number of targets detected in 2DTM depends on the SNR threshold, which is determined by the desired false-positive rate. Assuming a Gaussian noise distribution, we calculated a threshold SNR of 7.61 to allow for a false positive rate of one per image (Rickgauer et al., 2017). To test if a lower SNR threshold improves the agreement between 2DTM and 3DTM, we determined the proportion of detected 2DTM 50S targets with a matched 3DTM target at different SNR thresholds (**Figure 5f**). Decreasing the SNR threshold to 7.00 increases the number of detected targets >2-fold (from 652 to 1442), but the number of 2DTM targets with a corresponding 3DTM target increased only ~1.2-fold (from 576 to 714) (**Figure 5c**, lower). The plot of matching targets in **Figure 5f** shows that at SNR thresholds >7.6, the proportion of 2DTM / 3DTM paired targets was ~90-100%, while lowering the SNR threshold below 7.6 resulted in a sharp decrease, indicating that below this threshold, more non-matching targets were found than matching targets. The value of 7.6 agrees closely with the 7.61 threshold at which one false positive per image is expected (**Figure 5f**, dashed line). We conclude that while lowering the SNR threshold below this value increases the number of true targets identified by 2DTM, based on validation by matching 3DTM targets (**Figure 5c**, lower), the lower threshold also introduces more non-matching targets, many of them likely false positives. This experimentally validates our use of a Gaussian noise model and the Neyman-Pearson threshold criterion and shows that the SNR threshold can be used to set a desired false positive rate suitable for a particular experimental design.

At a threshold of 7.61, the number of 2DTM targets not detected in the 3DTM search was 76, >4 times higher than the expected 19 false positives for the 19 searched images and representing ~12% of the 2DTM peaks (**Figure 5g**). This is consistent with the estimated 3DTM false negative rate of ~10% (**Figure 5 - figure supplement 1**), and suggests that these may represent false negatives in the 3DTM search. We also noted that in samples of ~100 nm we detected ~90% of the 50S identified with 3DTM (**Figure 5e**). We therefore estimate that given optimum sample parameters, 2DTM can detect 50S subunits with comparable sensitivity to 3DTM.

2DTM can validate targets identified by 3DTM

While 2DTM tends to miss targets close to the chosen SNR threshold, one of the key limitations of 3DTM is a high false discovery rate and the difficulty in discriminating between true and false positives. This shortcoming can be partially mitigated by manually curating the detected 3DTM targets and subsequent classification of the subtomograms. The manual curation requires experience, is time-consuming and generally not reproducible. Progress has been made to improve the discrimination between true and false positives using deep learning algorithms as demonstrated with synthetic data (Gubins et al., 2019) although on real data they currently perform only as well as 3DTM (Moebel et al., 2020).

We investigated the use of 2DTM as an independent screen to validate 3DTM targets, by comparing the list of detected 3DTM targets before manual curation and 3D classification with the list of high-confidence 2DTM targets. We found that 2DTM detects ~60% of the 3DTM targets that are classified as 50S or 70S, while only detecting <5% of the 3DTM targets that likely represent false positives (**Figure 6a**). Thus 2DTM can clearly discriminate targets that were excluded by classification of 3DTM targets with *RELION* (Bharat & Scheres, 2016). This shows that 2DTM provides complementary evidence that can be used to reduce the rate of false positives of 3DTM.

To evaluate whether we could locate 30S subunits associated with 50S by template matching, and to compare the results to the *RELION* classification of the 3DTM results, we used the locations and orientations of the 50S targets detected with the *M. pneumoniae* template to perform a local search for the 30S subunit (see Methods), sampling a range of +/- 12.5 degrees, i.e., larger than the rotational range of the 30S relative to the 50S. Of the 700 50S locations searched, 314 (~45%) had detectable 30S (**Figure 6b**), fewer than the ~70% of the 3DTM targets that were classified as 70S using *RELION*. We speculate that this is due to structural discrepancies that likely exist between the 30S template and specific instances of the small ribosomal subunit in our images. 30S is known to undergo significant conformational changes during the functional cycle of the ribosome in addition to inter-subunit rotation (see Discussion).

We also compared the x,y locations (**Figure 6c**), z coordinates (**Figure 6e**) and orientations (**Figure 6d**) assigned by 2DTM to those assigned by 3DTM. Comparing the 2DTM results with the refined 3DTM results shows that ~95% (576/603) of the targets aligned in x and y had an angular distance of 20° or smaller, with a median of 7° (using Eq. (1)) (**Figure 6d**). The median in-plane distance of 12 Å is likely due to deformation of the sample under the electron beam, which is more pronounced with the higher electron dose used in tilt series, and to a lesser extent due to the lower resolution of 3DTM. In contrast to the x,y locations, which were calculated in steps of one pixel (1.27 Å), the defocus values of 2DTM were sampled in 200 Å steps and, correspondingly the median out-of-plane distance between detected 2DTM and 3DTM targets was much higher at +/- 84 Å (**Figure 6e**). We used the program *refine_template* (see Methods) to perform a local defocus refinement, which reduced the median out-of-plane distance to +/- 59 Å (**Figure 6e**), less than 1/3 the diameter of the 50S. The 5-fold greater error in z relative to the x,y plane is consistent with the previously reported weaker dependence of the 2DTM SNR values on defocus (Rickgauer et al., 2017).

We note that some of the 50S subunits detected by 2DTM overlap in the image and were found in comparable locations by 3DTM (e.g., **Figure 6f-i**). This confirms that 2DTM, as previously proposed (Rickgauer et al., 2017) can effectively detect overlapping particles without sample tilt.

Discussion

In this study, we provide strategies to overcome some of the major limitations of molecular localization in cells using 2DTM. Firstly, we present a new implementation of 2DTM in *cis*TEM, which offers a user-friendly interface and substantial acceleration by running the most demanding computations on GPUs. Secondly, we show that averaging of complexes identified with 2DTM can reveal differences between the template and the target and allow identification of interacting complexes that may be difficult to detect alone. Thirdly, we demonstrate that 2DTM detects *bona-fide* ribosomes in *M. pneumoniae* cells with high specificity and comparable sensitivity relative to 3DTM. In the following, we highlight some of the technical advances of our implementation, discuss applications of, and possible improvements to 2DTM and how 2DTM and 3DTM can benefit from each other.

GPU acceleration increases the throughput of 2DTM

To make 2DTM useful in practice, we have increased its speed by ~10-fold relative to our C++ CPU implementation in *cis*TEM via GPU acceleration, making it possible to search multiple images in hours instead of days. Flexible load balancing via CUDA streams, accessible to the user by adding threads in their *cis*TEM run profile, ensure a simple mechanism to take full

411 advantage of computers containing variable Nvidia GPU hardware architectures and adapt to
412 different problem sizes.

413 GPUs have so many processing units (cores) that efficient algorithms, like the FFT, are often
414 limited by memory transfers. We have ameliorated this problem by taking advantage of the FP16
415 format, storing two variables in the low and high 16 bits of a 32 bit __half2 vector data type,
416 thereby roughly doubling the memory bandwidth. Due to the low SNR intrinsic to cryo-EM data,
417 the loss of numeric precision associated with FP16 does not substantially alter the effective SNR
418 of the data stored in FP16 format. We expect that this could be useful for computation in other
419 modalities of cryo-EM.

420 With a few precautions, we can retain sufficient precision, i.e., greater than what is used for our
421 detection threshold, to obtain the same net results as with the 64-bit double precision
422 accumulators used in the CPU code. We have also taken advantage of increased memory
423 bandwidth of FP16 by converting the FFT of the input image, which is only transformed once at
424 the outset of the algorithm, into FP16 format. When this array is read into the GPU's streaming
425 multi-processors for cross-correlation, it is converted to 32-bit single precision as part of our
426 cuFFT callback routine prior to the conjugate multiplication and inverse FFT. Due to the limited
427 range of FP16, we did not investigate whether this reduced precision could be used directly in
428 the FFT in 2DTM. In principle, this could yield another ~1.6x speedup by doubling the inverse
429 FFT computation speed. Additionally, it may be possible to exploit the fact that each template is
430 padded with zeros to the full image size prior to taking the FFT, resulting in many redundant
431 zero-valued 1D FFTs in the common row-column transform approach for 2D FFTs.

432 Support for the new Bfloat16, which has a larger range than FP16, and hardware acceleration for
433 asynchronous memory movement on chip in the new Nvidia Ampere architecture, may also
434 provide additional opportunities to accelerate the current algorithm, which is predominantly
435 memory-bandwidth limited. Finally, further acceleration may be gained by improving the
436 algorithm itself, perhaps via a hierarchical search as is done in single-particle cryo-EM. For
437 example, a coarse search followed by a local refinement may improve both the speed and overall
438 accuracy of 2DTM.

439

440 **Template matching and noise overfitting**

441 Template matching is a well-established method to pick particles in single-particle cryo-EM
442 images for further processing and 3D reconstruction (Scheres, 2015; Sigworth et al., 2010). The
443 advantage of using a template that matches the structure of the particle to be reconstructed is
444 more discriminatory picking, and the possibility to exclude contaminants and other features in an
445 image that do not represent valid particles. Using a template for picking can also lead to template
446 bias, i.e., a 3D reconstruction that reproduces the template rather than the true particle structure

(Henderson, 2013; Subramaniam, 2013; Van Heel, 2013). It is therefore important to validate features in a reconstruction derived from particles identified by template matching. This is commonly done by observing features in the reconstruction that were not present in the template, and that are known to be true. For example, templates can be low-pass filtered to 20 Å or lower (Scheres & Chen, 2012), and the emergence of high-resolution features such as secondary structure or amino acid side chains will then validate the reconstruction. In the current release of *cis*TEM (Grant et al., 2018), a Gaussian blob is used for particle picking, and recognizable low and high-resolution features visible in the reconstruction serve as validation. Since 2DTM uses high-resolution features to identify targets, reconstructions have to be validated differently, for example by identifying additional *a priori*-known density features that were not present in the template. In the case of using the 50S templates for 2DTM, the reconstructions showed clear density for the 30S subunit, as well as at tRNA binding sites, both of which validate the reconstructions (**Figure 3c-d**). However, despite this validation, there may still be a template bias in the reconstruction (Stewart & Grigorieff, 2004). While it is not possible to quantify this bias without knowledge of the unbiased reconstruction, the strength of the bias will depend on the degrees of freedom accessible during template matching, i.e., the number of search locations. By increasing the SNR threshold to a value where only one false positive per search is expected on average, almost all search locations that do not contain matching signal are excluded. The high SNR threshold used in our 50S template search, therefore, limits the template bias in the reconstruction. Indeed, apart from the large additional density corresponding to the 30S subunit, there are also several smaller differences between the template and the reconstruction within the region of the 50S subunit that can be related to conformational changes occurring during translation (**Figure 3c-d**, see Results). This confirms that template bias in this reconstruction must be small relative to the unbiased signal represented by the reconstruction.

Template optimization

To realize an approach that combines 2DTM and 3DTM, a better understanding of how resolution affects 2D and 3D template searches will be needed. In this study, we found that applying a B-factor of 85 $Å^2$ to the 50S template maximized the mean detection SNR in this dataset (**Figure 3 - figure supplement 1a**). This value matches closely the B-factor of 86 $Å^2$ used to correct the *in situ* reconstruction of the *M. pneumoniae* 70S from subtomograms (Tegunov et al., 2021). Since this B-factor affects all atoms in the template equally, we expect that a more accurate method of generating templates that uses variable B-factors to account for local differences in mobility will further improve detection. We propose that given a sufficiently accurate method to calculate template densities, 2DTM could be used to investigate the relative impact of these and other, not yet considered, factors. For example, we could use biochemical restraints to control conformational/compositional heterogeneity and thereby investigate the process of radiation damage which likely varies based on a complex's chemical composition and

485 environment. Alternatively, if we can correctly model sample motion and radiation damage, we
486 could use 2DTM to probe for more detailed aspects of an individual target's structural identity
487 and composition, which may vary based on location in the cell or lifecycle of the organism to be
488 investigated. It will therefore be important in the further development of 2DTM to also develop
489 more reliable and accurate methods to generate templates from atomic coordinates, and to
490 comprehensively model cryo-EM images.

491 Templates in the present study used density maps generated by *pdb2mrc* (Tang et al., 2007), and
492 projections were calculated using the simple linear interpolation algorithms implemented in
493 *cis*TEM (Grant et al., 2018). These simplifications contributed to the acceleration of our 2DTM
494 implementation, compared to the more accurate density modeling and projection calculation used
495 previously (Rickgauer et al., 2017 & 2020). The simplifications, as well as deformation of the *M.*
496 *pneumoniae* cells under the electron beam (Tegunov et al., 2021), leading to noticeable blurring
497 in some of our images, likely affect the SNR values obtained in our template searches. Indeed,
498 the SNR values we observe are lower on average than expected based on the molecular mass of
499 our template (about 1.2 MDa), as well as previously observed values obtained with different
500 samples and template structures (Rickgauer et al., 2017 & 2020). Furthermore, the *M.*
501 *pneumoniae* 50S atomic model built into a 3.5-Å map (Tegunov et al., 2021) may contain atomic
502 coordinate errors that are larger than those in the 60S and 40S models used by Rickgauer et al.
503 (2020), which were built into 2.9-Å maps. Coordinate errors will further decrease the SNR
504 values obtained in a template search. Accurate atomic models and methods to simulate realistic
505 cryo-EM images (Himes & Grigorieff, 2021) will therefore be important to maximize the
506 detectability of molecules by 2DTM.

507

508 **2DTM with multiple templates**

509 2DTM employs the matched filter, which is the statistically optimal detector for a deterministic
510 signal in wide-sense stationary noise (McDonough, 1995). With zero-mean Gaussian noise, the
511 output of the matched filter is an SNR determined by the ratio of the cross-correlation coefficient
512 to the standard deviation of the noise in an image. Given that the noise is roughly constant for a
513 given image, the SNR measured in 2DTM is ultimately limited by the template's molecular
514 mass, making smaller complexes more difficult to detect. In our earlier work, we limited the
515 SNR threshold to a value allowing, on average, one false positive per search. This threshold is
516 clearly not optimal for every experiment and should be adjusted based on the experimental
517 rationale and design.

518 A match in 2DTM indicates that the template is sufficiently similar to the target to yield an SNR
519 above the significance threshold, but does not indicate that the target conforms to the exact
520 structure of the template. Indeed, significant peaks were detected using the *B. subtilis* 50S as a
521 template which has substantial differences to the *M. pneumoniae* 50S (discussed below). Since

target detection depends on pre-existing structures, detection by 2DTM is necessarily biased towards targets that are sufficiently similar to the state represented by the template, and other states may be missed. As a consequence, targets detected with 2DTM likely reflect a subset of the total complement of *in vivo* states. To reduce bias, variable regions could be removed from the template, thereby also reducing the potentially detectable signal in a template search. To detect targets that are in different states without loss of signal, multiple templates representing different states might be used. The problem of mismatched states is particularly evident when performing a local search for bound 30S subunits using the coordinates of detected 50S subunits. Despite the 30S subunits being well within the size range expected to be detectable by 2DTM in images of samples with a thickness of 150 – 200 nm (Rickgauer et al., 2017), we found only ~45% of the identified 50S bound with 30S subunits, just over half of the expected 70% based on the analysis of tomograms from *M. pneumonia* ((O'Reilly et al., 2020), **Figure 6b**). It is unlikely that this low detection rate is due to a bias of 2DTM towards detecting isolated 50S subunits because (i) a 3D reconstruction based on detected 50S subunits shows clear density for the 30S subunit (**Figure 3d**) and (ii) there was no significant difference in the detection of 50S and 70S targets, as validated by 3DTM (**Figure 6 - figure supplement 1**). Since the 30S subunit is highly dynamic and can adopt multiple conformations, it is likely that our false negatives arose from conformations that are not sufficiently similar to the 30S template that we used in our search.

Below an SNR of 7.2, the number of detected 2DTM targets without a 3DTM match is less than statistically expected (**Figure 5g**). This suggests that the statistics, which are based on a Gaussian noise model, might suffer from bias, for example because the correlation coefficients calculated for each search location are not completely independent from each other. While further investigation will be needed to model the noise in correlation maps more accurately, the current bias due to residual correlations between search locations results in a slight overestimation of the number of false positives at a given SNR. Incorporating more accurate noise models, and properly accounting for uncertainty in the model (reference structure) will require the replacement of the SNR values currently used to evaluate 2DTM results by a more general probabilistic framework. The use of likelihood ratios, for example, will make 2DTM more robust, as well as open up new avenues to explore molecular heterogeneity.

2DTM as a tool to investigate diverse species

One of the limitations of 2DTM is a reliance on pre-existing high-resolution structures. Outside of a few model organisms, high-resolution structures are not available for the vast majority of species. We show that, despite differences in their structures, a *B. subtilis* 50S template can detect 50S in *M. pneumoniae* cells (**Figure 4**). Thus, it is possible to use structures from related organisms to identify the locations and orientations of complexes with 2DTM. Mismatches between the template and cellular target resulting from species-specific structures reduce the

2DTM SNR (**Figure 4c**). By comparing the 2DTM SNRs of a series of structures from different species, it may be possible to infer evolutionary relationships in a manner analogous to DNA sequence comparison. Structural comparison with 2DTM would present the additional advantage of defining structural conservation which may not be evident by sequence comparison alone without the need to build a detailed molecular model of homologs in each species.

We have shown that it is possible to generate *in situ* 3D reconstructions of ribosomes from 2D images of cells, without the need to collect a tilt series (**Figures 3&4**). Template bias (discussed above) becomes especially pertinent when using a template from a different species when there is no existing structure to validate the obtained map. In a 3D reconstruction from 50S located by 2DTM with a *B. subtilis* 50S template we identified density corresponding to *M. pneumoniae* specific protein structures, and failed to detect density corresponding to a *B. subtilis* specific protein (**Figure 4f-g**). This demonstrates that by using a sufficiently high threshold to prevent a preponderance of false positives, targets identified using 2DTM templates from different species can be averaged to generate *in situ* 3D reconstructions with minimal template bias. Many biologically important organisms are difficult to grow at scales necessary for protein-purification and *in vitro* structure determination. Cryo-ET is also often limited by computationally demanding image processing, thus hindering routine *in situ* structure determination. 2DTM could be used as an alternative tool to accelerate *in situ* structure determination in non-model organisms.

2DTM and 3DTM complement each other

A significant challenge in using cryo-EM to study complexes *in situ* arises from the thick and crowded nature of the sample. The structural noise from neighboring particles is only correlated at low-resolution in projections recorded at different sample tilts. 3DTM is able to benefit from this indirectly by searching reconstructions which essentially "disentangle" the targets from this structural noise. While overlapping density cannot be removed in single-tilt images used for 2DTM, overlapping targets can still be separated based on their different x,y coordinates and defocus values (z coordinate). For example, using 2DTM we were able to discern overlapping 50S ribosomes, as validated by comparison with the 3DTM results (**Figure 6f-i**). Thus, separating overlapping particles does not require collecting a tilt series. Furthermore, because the structural noise is only correlated at low-resolution, a fine-grained search as in current versions of 2DTM, carried out with two or three tilted images may provide a better detection strategy to be implemented in the future.

While the results of 3DTM do not seem to be correlated with sample thickness within the tested range (**Figure 5 - figure supplement 2**), it is difficult to tell if this is due to a better overall performance of 3DTM in thicker samples, for example due to the separation of overlapping densities, or if 3DTM is simply less sensitive to the signal degradation associated with thicker

596 samples, for example by multiple and inelastic scattering. The latter seems more likely since
597 3DTM depends primarily on low-resolution signal, which in cryo-EM data is usually
598 substantially stronger than high-resolution signal and therefore remains detectable in
599 thicker samples despite signal degradation. While the stronger reliance on low-resolution
600 features makes 3DTM less sensitive to sample thickness and image-degrading factors, its high
601 false positive rate requires extensive expert curation, often exceeding the computational run time
602 of the original search. Even where time is not a concern, manual curation is difficult if not
603 impossible in dense regions of a cell, like the nucleus, or for particles smaller than a ribosome
604 that are not easily visually discernible. Particle classification approaches such as those
605 implemented in *RELION* can be useful for removing false positives. However, their performance
606 on noisy subtomograms of relatively small complexes remains problematic.

607 Using the full electron dose in a single exposure allows for the inclusion of high-resolution
608 information in a 2DTM search, which improves its precision and enables detection in dense
609 molecular environments. However, as we show, this comes at the expense of lower recall in thick
610 samples, and the lack of multiple sample tilts lowers the positional accuracy perpendicular to the
611 image plane. In principle, 3DTM could also take advantage of high-resolution information, but in
612 practice this is currently not achieved, possibly due to increased specimen motion on multiple
613 exposures, errors in the CTF determination from very low-dose exposures, difficulty in aligning
614 tilt series, and increasing sample thickness as a function of tilt (Voortman et al., 2014). New
615 approaches that take into account sample deformation (Himes & Zhang, 2018; Tegunov et al.,
616 2021) may help boost the high-resolution signal in tomograms.

617 3DTM may benefit from machine learning algorithms that can be trained to distinguish signal
618 from noise and thereby reduce false positives (Moebel et al., 2020). Although, to our knowledge,
619 this has not yet been tested, machine learning may also improve the performance of 2DTM for
620 similar reasons. Better overall performance of template matching may additionally be achieved
621 by combining the high precision of 2DTM with the high recall of 3DTM, making this approach
622 an effective and informative strategy for visual proteomics. In such a strategy, a zero-tilt image
623 with higher dose is collected before collecting a tilt series. The molecules of interest are then
624 identified in both 2D image and 3D tomogram, the 2D search is used to validate hits in the
625 tomogram, and the 3D coordinates provide context in 3D space and higher accuracy in the z
626 coordinates (e.g., **Figure 5a-b**). Similar approaches that leverage the different information
627 available from 2D images and 3D tomograms have been suggested previously (Alberto
628 Bartesaghi et al., 2012; Sanchez et al., 2020). Moreover, since the image modalities (2D vs 3D)
629 are distinct in the two methods, the noise and background in the data are only partially related: (i)
630 Layers of the sample overlapping in the 2D image can be separated in the tomogram. (ii) The
631 imaging parameters (effective image contrast/defocus, electron dose) as well as the random noise
632 accompanying all cryo-EM data differ between the 2D and 3D data. (iii) Template matching in
633 2D and 3D relies on different resolution ranges – 3DTM depends strongly on shape information
634 (~20 Å) while 2DTM depends on 3 – 5 Å resolution. Overlapping hits can therefore be

considered to be true positives with high confidence. The requirement of dual detection may allow the detection threshold to be lowered in each search, increasing the overall number of true positives. Such an approach could speed up workflows by avoiding the need for labor-intensive classification and refinement of 3DTM datasets, while incorporating the additional structural context provided by a tomogram.

Future of template matching and conclusions

We describe here the implementation of a GPU-accelerated 2DTM method into the graphical user interface of the open-source software *cis*TEM. 2DTM effectively detects ribosomes in 2D images of frozen *M. pneumoniae* cells at higher specificity, but lower sensitivity than 3DTM in samples thicker than 100 nm. We propose that an effective, high-confidence strategy for *in situ* visual proteomics would combine 2DTM and 3DTM and would not be limited to species for which high-resolution structures are available. We also propose that the search space of 2DTM could be further expanded to include a more detailed analysis of the signal present in an image, allowing for interrogation of conformational or compositional variability, by searching with a multi-template library. We demonstrated that 3D reconstructions from targets detected by 2DTM reveal new features not present in the template. This technique, therefore, has the potential to deliver *in situ* structures at high resolution, similar to subtomogram averaging but requiring less experience and time. While it requires prior knowledge of the target structure to be reconstructed, i.e., a template, the 2DTM workflow (including data collection) is significantly faster than tomography, and may therefore be more suitable for high-throughput studies. Finally, the need for averaging may limit detailed interrogation of molecular structures to more abundant complexes. The spatial organization, structures, composition and functional states of rare complexes can still be studied in cells by 2DTM, provided they are detectable in 2D images. With further improvements, we expect 2DTM to reveal new insights into the molecular mechanisms of biological processes in their native cellular context.

Methods

Cell culture, grid preparation and cryo-EM imaging

Mycoplasma pneumoniae sample preparation and cryo-EM imaging were carried out as described previously (O'Reilly et al., 2020; Tichelaar et al., 2020). In brief, *M. pneumoniae* cells were grown on Quantifoil gold grids in modified Hayflick medium and the grid was quickly washed with PBS buffer with 10 nm gold fiducial beads (Aurion, Germany) before plunge-freezing. The grids were imaged in a 300 keV Krios TEM (ThermoFisher) equipped with a direct

detector (Gatan) and a quantum post-column energy filter (Gatan). Data collection was done using SerialEM (Mastronarde, 2005).

For the 2D images paired with tilt-series data, 19 cells were first imaged following the SPA standard, with the magnification of 215,000 (pixel size 0.65 Å) and exposure time of 2 seconds (20 frames, total exposure dose 32 $e^-/Å^2$). The target defocus of 0.2 µm to 0.5 µm. Tilt-series were collected using the dose-symmetric scheme (Hagen et al., 2017) with the following settings: magnification 81,000 (pixel size 1.7 Å), tilt range -60° to 60° with 3° increment, exposure time per tilt 1 second, total exposure dose ~129 $e^-/Å^2$. The target defocus remained the same within tilt-series, and ranged from 1.5 µm to 3.5 µm between tilt-series.

A dataset of 220 2D images of *M. pneumoniae* were collected on a K3 camera (Gatan) and at a magnification of 81,000 (pixel size 1.503 Å). The K3 camera was run in non-CDS mode with the following settings: exposure time 1.678 second, 24 frames, total exposure dose ~31 $e^-/Å^2$.

2D template generation

To generate templates for 2DTM, we used the computer program *pdb2mrc* (Tang et al., 2007) to convert PDB-formatted atomic coordinates into 3D density. To reduce potential aliasing, we initially generated an over-sampled density map at half the pixel size of the image to be searched. The sampling rate was then halved to the final value pixel size by Fourier cropping using the program *resample* from the *cis*TEM software package (Grant et al., 2018). The resulting 3D density was low-pass filtered with the standalone program *bfactor* (Grigorieff, 2021a) using a B-factor of 85 $Å^2$ and placed into a 256 x 256 x 256 voxel box, about twice the size of the 50S ribosomal subunit it contained. The final pixel size for searching the 19-image dataset of *M. pneumoniae* that was compared with tomograms was of 1.27 Å. For computational efficiency, we used a final pixel size of 1.5 Å to search the 220-image dataset. As template models, we used the *M. pneumoniae* ribosome (PDB: in progress; EMDB 11999) and the *B. subtilis* ribosome (PDB: 3J9W). We separated the small and large ribosomal subunits and, for the *B. subtilis* ribosome, removed the mRNA, tRNAs and MifM protein from the coordinate file using custom Perl scripts. The resulting coordinates were aligned with *USCF Chimera* (Pettersen & Goddard, 2004) to the 50S template used for 3DTM to place them in a common coordinate system. This ensured that the 3D positions and Euler angles found by 2DTM and 3DTM referred to the same coordinate system and could be directly compared. For local 30S searches, the atomic coordinates corresponding to the body of the *M. pneumoniae* 30S were used to generate a template as described above. For this we combined the results of searches with two templates, one with the proteins and rRNA sequence corresponding to the 30S head and one without, to account for the greater conformational variability of the small subunit relative to the large subunit.

2D template matching in *cis*TEM

Movie frames were aligned with *unblur* using the "optimal exposure filter" (Grant & Grigorieff, 2015) and including all frames in the final sum, unless otherwise indicated. Defocus was determined by *CTFFIND4* (Rohou & Grigorieff, 2015) from within the *cis*TEM GUI (Grant et al., 2018). Templates for 2DTM were imported into *cis*TEM as 3D volumes and 2DTM was performed on all images, including a defocus search over a 2400 Å range centered on the average defocus determined by *CTFFIND4* with a 200 Å step, using a 2.5° out of plane angular search step and a 1.5° in plane angular search step, assuming C1 symmetry and a minimum target - target distance (peak radius) of 10 pixels (~13 Å). As previously described, the SNR values resulting from 2DTM were further normalized by subtracting the mean of the SNR values for all orientations at each location, and dividing by their standard deviation (Rickgauer et al., 2017, 2020). The best peak radius to use in a search depends on the B-factor affecting the signal in the images and may, therefore, vary for different experiments.

Detection of 30S by local search using 50S 2DTM coordinates

To detect 30S subunits bound to 50S subunits, we wrote *refine_template*, a computer program that is accessible through *cis*TEM's GUI, or as a command line tool. *refine_template* reads the 3D template and the output files of a template search, including the MIP and alignment parameters for detected targets. Using these data, it calculates template projections and performs a local refinement of the projection parameters. This refinement can be used, for example, to find the defocus values that maximize the cross-correlation coefficient between projection and image. It can also be used to detect molecules and complexes bound to already detected targets. In this case, *refine_template* is used with the search results obtained for the detected targets, but with the bound complex to be detected as the input template. The locally refined alignment parameters found for the bound complex will be close to those found for the original template, provided both the original template used in the search and the bound complex in the refinement share the same coordinate system.

To ensure the coordinate systems of the *M. pneumoniae* 30S and 50S subunits were aligned, the coordinates for the 30S and 50S subunits were separated from the 70S into two coordinate files and used to generate 3D templates for 2DTM for each of them (see above). For this last step, it was important not to use the *-center* flag of *pdb2mrc* to make sure the 3D densities remained in their correct locations relative to the other subunit. 50S targets were then detected using the 50S template, followed by local refinement with *refine_template* and the 30S template. For the refinement, we searched within an angular range of +/- 12.5° using a 2.5° step. A 30S was deemed to be present when the refined x,y coordinates were within 20 Å of the original coordinates found for the 50S template.

3D reconstruction using 50S 2DTM coordinates

We wrote a new computer program called *make_particle_stack*, as part of the *cis*TEM image processing package. It reads the results of a template search, including the MIP with the peaks indicating detected targets, as well as the Euler angles of the detected targets, the original image that was searched, and other imaging parameters (defocus, amplitude contrast, lens aberrations, beam energy and pixel size). On output, it will generate a stack of boxed-out targets and a list of alignment parameters that can be used for 3D reconstruction with *cis*TEM. Using *make_particle_stack* and *cis*TEM, we generated a reconstructions from 5,080 and 1,172 50S targets detected in 220 images of *M. pneumoniae* cells, using the Euler angles and x,y locations assigned by the *M. pneumoniae* and *B. subtilis* 50S template search, respectively. The reconstructions had nominal resolutions of 4.3 Å and 5.5 Å, respectively (FSC threshold of 0.143, (Rosenthal & Henderson, 2003), **Figure 3 - figure supplement 1e**, **Figure 4 - figure supplement 1b**). However, visible density of known ribosome features (not shown) suggested that this resolution estimate was unrealistically high, as expected due to the well-known effect of template bias (Stewart & Grigorieff, 2004). We therefore low-pass filtered the reconstructions to 20 Å resolution. The low-pass filtered reconstructions reproduced the 50S density, except for a few details in peripheral features (see Results). They also revealed additional clear density for the 30S subunit, albeit weaker than the 50S density, and density that corresponds to tRNAs in conformations consistent with translating ribosomes (see Results).

3D template matching

For tilt-series data, movie frames were aligned on-the-fly using the SerialEM plug-in *alignframes*. Tilt-series alignment was carried out in IMOD (Mastronarde & Held, 2017). Warp was used to estimate CTF, import tilt-series alignments from IMOD, reconstruct tomograms and subtomograms (Tegunov et al., 2021). Cell thickness was estimated based on central Y-Z sections of the tomograms.

We used PyTom to perform 3DTM of tomograms as described previously (Hrabe et al., 2012; O'Reilly et al., 2020). The 50S ribosome template was generated from a previous 50S class average from *M. pneumoniae* (O'Reilly et al., 2020). The density was low-pass filtered to 30 Å resolution prior to template matching in PyTom. 3DTM hits were extracted from the scores map generated by PyTom and were ranked by cross-correlation scores for each tomogram. Subtomograms for the top 600 hits in each tomogram were extracted using Warp. 3D classification and refinement were done in *RELION* 3.0.8 (Zivanov et al., 2018). In total, 3062 70S and 6336 50S subtomograms were classified, after removing the 1827 false positives, i.e., particles that could not be aligned (**Figure 5 - figure supplement 1**).

Comparison of 2DTM and 3DTM

To compare targets found by 2DTM and 3DTM, we wrote a new computer program, *align_coordinates*, as part of *cis*TEM (Grant et al., 2018), to identify and align target coordinates. *align_coordinates* assumes that the z-axes of the coordinate systems of the 2DTM and 3DTM searches are aligned, i.e., they refer to the same sample tilt. Therefore, to find the angular alignment and x,y offset between a set of reference coordinates and a second set of sample coordinates, only the x,y coordinates need to be considered. In a first step, *align_coordinates* calculates all possible vectors between any two x,y coordinates in a coordinate set. A rotation matrix is applied to the sample coordinate vectors, and the vector with the smallest vector difference within a given tolerance is found with respect to each reference vector. The algorithm performs a systematic search of the rotation angle in 0.5° steps to find the rotation that produces the largest number of matching vectors while giving a higher weight to better matching vectors. The result of this systematic search is a rough rotation angle and a list of corresponding coordinates that define a coordinate transform. The search also identifies coordinates in each set that do not have corresponding coordinates in the other set. The coordinate transform is then refined, and the list of corresponding coordinates is updated using a finer local search with angular step of 0.01° and 1-Å steps along the x and y axes. The best coordinate transform is selected based on a least-squares fit between the corresponding coordinates.

In the comparison of 2DTM and 3DTM coordinates, the latter were used as a reference set, and the distance threshold for correspondence was set to 110 Å. The final list of corresponding targets was further reduced by limiting distances to 100 Å and below. By providing the dimensions of the field of view of the images used for 2DTM, it was also possible to identify the corresponding area within the larger field of view of tomograms analyzed by 3DTM. For the calculation of percentages of corresponding targets identified by each search, only targets in the overlapping areas were counted.

To compare the angular orientations for a given coordinate pair, the 2DTM template was first aligned relative to the 3DTM template (see above), and the Euler angles of the detected targets were recorded. The total angular distance between each orientation, represented by the rotation matrices R_2 and R_3, given by the Euler angles, was calculated using the equation (Huynh, 2009)

$$\Delta\vartheta_{2,3} = \cos^{-1}\left[\frac{tr\left(R_2 R_3^T\right)-1}{2}\right]. \qquad \text{Eq. (1)}$$

Acknowledgements

We are grateful to Wim Hagen for advice on cryo-EM data acquisition, and Peter Rickgauer for providing numerous fruitful discussions and advice on implementing template matching in *cis*TEM. JM acknowledges support from the EMBL and the European Research Council (760067).

Author contributions

BAL, Designed experiments, analyzed data, discussed and interpreted results, and wrote and revised the manuscript; BAH, Designed and wrote GPU-enabled software for 2D template matching, discussed and interpreted results, and wrote and revised the manuscript; LX, Designed experiments, collected and analyzed data, discussed and interpreted results, and helped edit the manuscript; TG, Implemented the CPU version of 2D template matching in *cis*TEM and designed and implemented the graphical user interface and data structure used with it, and helped edit the manuscript; JM, Designed experiments, discussed and interpreted results, and helped edit the manuscript; NG, Conceived the project, designed experiments, analyzed data, discussed and interpreted results, and wrote and revised the manuscript.

Competing Interests

The authors declare no competing interest.

Software Availability:

All the code used for the 2D template matching has an open source license and is freely available from the *cis*TEM github repository, https://github.com/timothygrant80/cisTEM. For the purpose of reproducibility, the code base at the time of submission is publicly archived (Himes et al., 2021).

Figure legends

Figure 1: *cis*TEM GUI implementation of 2DTM. (a) Screenshot showing the results of a 2DTM search in the *cis*TEM GUI (located in the "Experimental" tab). The panel on the left shows all images searched. Images may be searched individually (column #1) or as batch jobs (column #2). The Results tab shows the locations, orientations and SNR values of each detected target in a list, as well as the original image (b) (membrane highlighted in yellow), the maximum intensity projection (MIP, c) and the plotted result (d), which shows the best-matching orientation of the template at each detected location. The survival histogram (subpanel in (a)) shows the SNR values for all search locations (blue line) and compares this with the survival histogram of Gaussian noise (red line). This is used to establish the threshold at which a single false positive is expected per image. Scale bar in (b) = 500Å.

Figure 2: GPU acceleration of 2DTM in *cis*TEM. (a) The angular search space is distributed among any number of processors using the home-grown MPI-like socket communication in the *cis*TEM GUI (Grant et al., 2018). Unlike MPI, if fewer processors are available than requested (shaded box) processing may still proceed. (b) To expose further parallelism, additional host threads may be requested to subdivide each angular group to maximize occupancy on the GPU. Each host thread queues up a series of GPU kernels into its respective stream, and then returns to calculate the next projection and initiates its transfer to the GPU (green box). This way, close to 100% of the CPU and GPU is used during computation. (c) GPU acceleration relative to optimized CPU-based calculation, of which 85% is spent on MKL-based FFTs. Kernel-fusion using cuFFT callbacks and custom data structures combined with flexible kernel launch parameters ensure the GPUs stay saturated, enabling an 8x speedup. A total of 10.5x speedup is achieved by optimizing data throughput using the vectorized FP16 format for storing results. (d) The code scales nearly linearly with the number of GPUs and tracks with the total memory bandwidth of a given model. All timing was obtained using a padded K3 image with 4096 x 5832 pixels, and searching one defocus plane with 2.5°/1.5° angular steps.

Figure 3: 2DTM detects ribosomes in *Mycoplasma pneumoniae* cells. (a) An overview of 2DTM: a cryoEM-like density was generated from an *M. pneumoniae* 50S model, a B-factor was applied and the resulting template used to identify locations and orientations of 50S in 2D images of *M. pneumoniae* cells with 2DTM implemented in *cis*TEM. Scale bar = 500Å (b) 20-Å filtered 3D reconstruction generated using the locations and orientations of 5,080 50S subunits detected in 220 images using 2DTM with a *M. pneumoniae* 50S template, showing clear density for the 30S ribosomal subunit (not included in the template). (c) Difference map showing the regions of the 3D reconstruction that differ from the 50S 2DTM search template. Arrows indicate additional density consistent with 70S ribosome structures. The difference map was generated with the same threshold as in (b). (d) A region of the difference map shown in (c), showing tRNAs in characteristic arrangements in the E, P and A sites of the 30S subunit. *M.p.*: *M. pneumoniae*

Figure 4: 2DTM using a *B. subtilis* 50S template reveals species-specific structures. (a) Molecular models of *B. subtilis* (red) and *M. pneumoniae* (blue) 50S ribosomal subunits aligned using Chimera. (b) Venn diagram showing the number of 50S subunits detected in the same dataset of 220 images of *M. pneumoniae* cells using the indicated template. (c) Boxplot showing the distribution of 2DTM SNR values of the locations quantified in the diagram in (b). The width of the box indicates the interquartile range, the middle line indicates the median and the whiskers indicate the range. The dashed vertical line indicates the 2DTM SNR threshold used. (d) 20-Å filtered 3D reconstruction generated using the locations and orientations of 1,172 50S subunits detected in 220 images using 2DTM with a *B. subtilis* 50S template, showing clear density for the 30S ribosomal subunit and L7/L12 (not included in the template). The threshold was selected to reflect the threshold used in **Figure 3b&c**. (e) Difference map showing the regions of the 3D reconstructions that differ from the 50S 2DTM search template. Arrows indicate additional density consistent with 70S ribosome structures. (f) Difference map as described in (e) (gray mesh), aligned to the *B. subtilis* 50S template (red) and both *M. pneumoniae* (*M.p.*, blue) and *B. subtilis* (red, not visible) molecular models. The difference map was generated with the same threshold as in (d). (g) 3D reconstruction as described in (d) (transparent gray), aligned to *M. pneumoniae* (*M.p.*, blue) and *B. subtilis* (*B.s.*, red) molecular models.

Figure 5: Comparison of ribosome detection by 2DTM and 3DTM. (a) Images of untilted cryo-EM grids of *M. pneumoniae* were collected with a total dose of 32 $e^-/Å^2$, followed by a tilt series of an overlapping region with a total dose of 129 $e^-/Å^2$ to reconstruct a tomogram. (b) 50S ribosomal subunits were identified in the 2D images by 2DTM with the *M. pneumoniae* 50S (left) and in the 3D tomogram by 3DTM using the 50S as a template (right). The 2DTM and 3DTM templates were aligned to ensure that the respective coordinate systems were aligned, and that the x,y coordinates of the detected 50S in each search could be aligned. (c) The proportion of 2DTM and 3DTM coordinates within 100 Å and 10° in each of the three Euler angles in 19 images was calculated using an SNR threshold that allowed either one false positive per image (upper), or detection of ~2 times more potential 50S targets (lower). (d) Plot showing the proportion of 2DTM targets that were also detected by 3DTM as a function of sample thickness. (e) Plot showing the proportion of 3DTM targets that were also detected by 2DTM as a function of sample thickness. (f) Plot showing the proportion of 2DTM 50S targets with a positional and rotational 3DTM match at the indicated 2DTM SNR threshold (dashed line). (g) Plot showing the number of expected false positives in the 2DTM search assuming a Gaussian noise model (black) and the observed number of 2DTM targets without a matching 3DTM target (blue) at the indicated 2DTM SNR.

Figure 6: 2DTM is specific, excludes non-ribosome particles and permits detection of ribosomes overlapping along the projection direction. (a) Pie chart showing the results of a comparison of a set of 1,380 3DTM coordinates initially identified by PyTom followed by 3D classification in *RELION* that identified 989 targets as 50S or 70S (left) and 391 targets as non-ribosomal particles (right), with a list of 652 50S 2DTM targets. Red indicates the proportion

aligning with 50S coordinates identified by 2DTM; gray indicates non-matching 3DTM coordinates. (b) Pie charts showing the proportion of 50S detected by 2DTM with the *M. pneumoniae* template with a bound 30S target as determined by performing a local search with *refine_template* (see Methods) (left), and the ratio of 3DTM 50S targets classified as 70S or 50S by *RELION* (right). (c) Histogram showing the distribution of the in-plane distance between matched 3DTM and 2DTM after 3DTM refinement of the subtomograms with *RELION*. (d) As in (c), showing the angular distance. (e) As in (c), showing the out-of-plane difference (z coordinate) before (blue) or after (yellow) 2DTM refinement of z-coordinates. (f) Plotted result from a 2DTM search showing template projections at the locations and Euler angles of detected 50S subunits (left), inset showing two overlapping 50S subunits when viewed parallel to the image plane. (g) As in (f), showing the plotted result from 3DTM of the same area aligned to show the same perspective. (h) The template projections in (f) rotated 90° to show the overlapping 50S perpendicular to the image plane. (i) As in (h), showing the result from 3DTM in the same area, faded density reflects a 50S in the tomogram that was not detected in the 2D image.

Figure 1 - Supplement 1: The 2DTM matching algorithm as implemented in *cis*TEM. Inputs to various stages are in hexagons. If the GPU is used, all memory allocations are handled by the TemplateMatchingCore class via calls to the underlying GpuImage class. The whitening filter and CTF are combined on the host and if needed, copied to the GPU once for each defocus plane searched. The inner loop (**Figure 2 - figure supplement 1**) is executed and results returned to the host.

Figure 2 - Supplement 1: (a) The 2DTM matching inner loop as implemented on the GPU in *cis*TEM. (b) Approximate percentage of run time for each step. The relative percentages can vary based on the automatic load balancing due to the combination of CUDA streams and mixed kernel launch configurations that restrain some low-complexity operations to a small subset of available streaming multi-processors via grid-stride loops.

Figure 3 - Supplement 1: (a) Boxplot showing the distribution of 2DTM SNR values in a representative image using an *M. pneumoniae* 50S template with the indicated B-factor applied. The boxes indicate the interquartile range, the middle line indicates the median, the whiskers indicate the 10-90th percentiles and the dots indicate 50S peaks with SNRs outside this range. (b) As in (a), showing the distribution of 2DTM SNRs in images whereby 8 (12.8 e/Å^2) or 20 (32 e/Å^2) frames with or without exposure weighting were used to generate the final image as indicated. (c) Bar chart showing the number of detected 50S when 13 defocus planes are searched, relative to the number when a single defocus plane is searched. Individual images of indicated thickness are shown separately. (d) Scatterplot showing the mean 2DTM SNR of 50S identified in images with >10 peaks relative to the mean defocus of the image calculated using *CTFFIND4*. (e) FSC obtained for the 3D reconstruction (shown in **Figure 3b**) calculated from the targets found by 2DTM.

Figure 4 - Supplement 1: (a) Fourier shell correlation (FSC) curve showing the correlation between the *M. pneumoniae* and *B. subtilis* templates. (b) FSC obtained for the 3D reconstruction calculated from the targets found by 2DTM (shown in **Figure 4d**). (c) Image showing the *B. subtilis* 50S template (red) aligned with the *M. pneumoniae* 50S model (blue), *B. subtilis* 50S model (not visible), the difference map from **Figure 4e** (gray mesh) and the difference map from **Figure 3c**. White arrow indicates the *M. pneumoniae*-specific C-terminal extension of L22, the black arrows indicate unattributed density that is common to both difference maps.

Figure 5 - Supplement 1: (a) Cross-correlation scores for the top 600 hits detected in 3DTM using a 50S template. (b) Classification of the subtomograms in *RELION*. (c) Bar chart showing the proportion of ranked 3DTM hits classified as false positives (yellow), 50S (green; used as a template) or 70S (purple) by classification in *RELION*. The ranked hits are binned in sets of 3.

Figure 5 - Supplement 2: (a) Scatterplot showing the mean 2DTM SNR of the 19 images searched in **Figure 5** relative to the sample thickness calculated from a tomogram. The size of each point is proportional to the total number of 50S detected with 2DTM as shown in the legend (right). (b) As in (a), showing the number of 50S detected with 3DTM in the area of the tomogram overlapping the 2D image. (c) Scatterplot showing the 2DTM SNR relative to the 3DTM cross-correlation (CC) score of all 576 targets identified in **Figure 5**.

Figure 6 - Supplement 1: Proportional bar chart showing the percentage of all 983 3DTM targets in the area of the tomogram overlapping with the 2D image or the 576 that were also detected by 2DTM, that were classified as 50S (orange) or 70S (blue) by *RELION*.

978 **References:**

979 Abeyrathne, P. D., Koh, C. S., Grant, T., Grigorieff, N., & Korostelev, A. A. (2016). Ensemble
980 cryo-EM uncovers inchworm-like translocation of a viral IRES through the ribosome.
981 *ELife*, *5*, 1–31. https://doi.org/10.7554/eLife.14874

982 Albert, S., Schaffer, M., Beck, F., Mosalaganti, S., Asano, S., Thomas, H. F., Plitzko, J. M.,
983 Beck, M., Baumeister, W., & Engel, B. D. (2017). Proteasomes tether to two distinct sites at
984 the nuclear pore complex. *Proceedings of the National Academy of Sciences*, *114*(52),
985 201716305. https://doi.org/10.1073/pnas.1716305114

986 Alberts, B. (1998). The Cell as a Collection of Protein Machines : Preparing the Next Generation
987 of Molecular Biologists. *Cell*, *92*, 291–294.

988 Allegretti, M., Zimmerli, C. E., Rantos, V., Wilfling, F., Ronchi, P., Fung, H. K. H., Lee, C. W.,
989 Hagen, W., Turoňová, B., Karius, K., Börmel, M., Zhang, X., Müller, C. W., Schwab, Y.,
990 Mahamid, J., Pfander, B., Kosinski, J., & Beck, M. (2020). In-cell architecture of the
991 nuclear pore and snapshots of its turnover. *Nature*, *586*(7831), 796–800.
992 https://doi.org/10.1038/s41586-020-2670-5

993 Asano, S., Fukuda, Y., Beck, F., Aufderheide, A., Förster, F., Danev, R., & Baumeister, W.
994 (2015). A molecular census of 26. *Science (New York, N.Y.)*, *347*(6220), 439–443.

995 Bartesaghi, A., Sprechmann, P., Liu, J., Randall, G., Sapiro, G., & Subramaniam, S. (2008).
996 Classification and 3D averaging with missing wedge correction in biological electron
997 tomography. *Journal of Structural Biology*, *162*(3), 436–450.
998 https://doi.org/10.1016/j.jsb.2008.02.008

999 Bartesaghi, Alberto, Lecumberry, F., Sapiro, G., & Subramaniam, S. (2012). Protein secondary
1000 structure determination by constrained single-particle cryo-electron tomography. *Structure*,
1001 *20*(12), 2003–2013. https://doi.org/10.1016/j.str.2012.10.016

1002 Beck, M., & Baumeister, W. (2016). Cryo-Electron Tomography: Can it Reveal the Molecular
1003 Sociology of Cells in Atomic Detail? *Trends in Cell Biology*, *26*(11), 825–837.
1004 https://doi.org/10.1016/j.tcb.2016.08.006

1005 Behrmann, E., Loerke, J., Budkevich, T. V., Yamamoto, K., Schmidt, A., Penczek, P. A., Vos,
1006 M. R., Bürger, J., Mielke, T., Scheerer, P., & Spahn, C. M. T. T. (2015). Structural
1007 Snapshots of Actively Translating Human Ribosomes. *Cell*, *161*(4), 845–857.
1008 https://doi.org/10.1016/j.cell.2015.03.052

1009 Berman, H. M., Battistuz, T., Bhat, T. N., Bluhm, W. F., Bourne, P. E., Burkhardt, K., Feng, Z.,
1010 Gilliland, G. L., Iype, L., Jain, S., Fagan, P., Marvin, J., Padilla, D., Ravichandran, V.,
1011 Schneider, B., Thanki, N., Weissig, H., Westbrook, J. D., & Zardecki, C. (2002). The
1012 protein data bank. *Acta Crystallographica Section D: Biological Crystallography*, *58*(6 I),
1013 899–907. https://doi.org/10.1107/S0907444902003451

1014 Bharat, T. A. M., & Scheres, S. H. W. (2016). Resolving macromolecular structures from

electron cryo-tomography data using subtomogram averaging in RELION. *Nature Protocols*, *11*(11), 2054–2065. https://doi.org/10.1038/nprot.2016.124

Bohm, J., Frangakis, A., Hegerl, R., Nickell, S., Typke, D., & Baumeister, W. (2000). Toward detecting and identifying macromolecules in a cellular context: template matching applied to electron tomograms. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(26), 14245–14250. https://doi.org/10.1073/pnas.230282097

Braunger, K., Pfeffer, S., Shrimal, S., Gilmore, R., Berninghausen, O., Mandon, E. C., Becker, T., Förster, F., & Beckmann, R. (2018). Structural basis for coupling protein transport and N-glycosylation at the mammalian endoplasmic reticulum. *Science*, *360*(6385), 215–219. https://doi.org/10.1126/science.aar7899

Brilot, A. F., Chen, J. Z., Cheng, A., Pan, J., Harrison, S. C., Potter, C. S., Carragher, B., Henderson, R., & Grigorieff, N. (2012). Beam-induced motion of vitrified specimen on holey carbon film. *Journal of Structural Biology*, *177*(3), 630–637. https://doi.org/10.1016/j.jsb.2012.02.003

Cai, S., Böck, D., Pilhofer, M., & Gan, L. (2018). The in situ structures of mono-, di-, and trinucleosomes in human heterochromatin. *Molecular Biology of the Cell*, *29*(20), 2450–2457. https://doi.org/10.1091/mbc.e18-05-0331

Campbell, M. G., Cheng, A., Brilot, A. F., Moeller, A., Lyumkis, D., Veesler, D., Pan, J., Harrison, S. C., Potter, C. S., Carragher, B., & Grigorieff, N. (2012). Movies of ice-embedded particles enhance resolution in electron cryo-microscopy. *Structure*, *20*(11), 1823–1828. https://doi.org/10.1016/j.str.2012.08.026

Diaconu, M., Kothe, U., Schlünzen, F., Fischer, N., Harms, J. M., Tonevitsky, A. G., Stark, H., Rodnina, M. V., & Wahl, M. C. (2005). Structural basis for the function of the ribosomal L7/12 stalk in factor binding and GTpase activation. *Cell*, *121*(7), 991–1004. https://doi.org/10.1016/j.cell.2005.04.015

Eibauer, M., Hoffmann, C., Plitzko, J. M., Baumeister, W., Nickell, S., & Engelhardt, H. (2012). Unraveling the structure of membrane proteins in situ by transfer function corrected cryo-electron tomography. *Journal of Structural Biology*, *180*(3), 488–496. https://doi.org/10.1016/j.jsb.2012.09.008

Frangakis, A. S., Böhm, J., Förster, F., Nickell, S., Nicastro, D., Typke, D., Hegerl, R., & Baumeister, W. (2002). Identification of macromolecular complexes in cryoelectron tomograms of phantom cells. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(22), 14153–14158. https://doi.org/10.1073/pnas.172520299

Frank, J. (2006). *Electron Tomography: Methods for Three-Dimensional Visualization of Structures in the Cell*. https://doi.org/10.1007/978-0-387-69008-7_4

Fu, X., Himes, B. a, Ke, D., Rice, W. J., Ning, J., & Zhang, P. (2014). Controlled Bacterial Lysis for Electron Tomography of Native Cell Membranes. *Structure (London, England : 1993)*, *22*(12), 1875–1882. https://doi.org/10.1016/j.str.2014.09.017

Grant, T., & Grigorieff, N. (2015). Measuring the optimal exposure for single particle cryo-EM using a 2.6 Å reconstruction of rotavirus VP6. *ELife*, *4*, e06980. https://doi.org/10.7554/eLife.06980

Grant, T., Rohou, A., & Grigorieff, N. (2018). *cis TEM , user-friendly software for single-particle image processing*. 1–24.

Grigorieff, N. (2021a). *bfactor.exe*. https://grigoriefflab.umassmed.edu/bfactor

Grigorieff, N. (2021b). *diffmap.exe*. https://grigoriefflab.umassmed.edu/diffmap

Grigorieff, N., & Harrison, S. C. (2011). Near-atomic resolution reconstructions of icosahedral viruses from electron cryo-microscopy. *Current Opinion in Structural Biology*, *21*(2), 265–273. https://doi.org/10.1016/j.sbi.2011.01.008

Grosjean, H., Breton, M., Sirand-Pugnet, P., Tardy, F., Thiaucourt, F., Citti, C., Barré, A., Yoshizawa, S., Fourmy, D., de Crécy-Lagard, V., & Blanchard, A. (2014). Predicting the Minimal Translation Apparatus: Lessons from the Reductive Evolution of Mollicutes. *PLoS Genetics*, *10*(5). https://doi.org/10.1371/journal.pgen.1004363

Grünewald, K., Medalia, O., Gross, A., Steven, A. C., & Baumeister, W. (2002). Prospects of electron cryotomography to visualize macromolecular complexes inside cellular compartments: Implications of crowding. *Biophysical Chemistry*, *100*(1–3), 577–591. https://doi.org/10.1016/S0301-4622(02)00307-1

Gubins, I., Schot, G. Van Der, Veltkamp, R. C., Förster, F., Du, X., Zeng, X., Zhu, Z., Chang, L., Xu, M., Moebel, E., Lai, T. M., Han, X., Terashi, G., Kihara, D., Himes, B. A., Wan, X., Zhang, J., Gao, S., Hao, Y., … Zhang, F. (2019). SHREC ' 19 Track : Classification in Cryo-Electron Tomograms. *Eurographics Workshop on 3D Object Retrieval*, 1–6.

Hagen, W. J. H., Wan, W., & Briggs, J. A. G. (2017). Implementation of a cryo-electron tomography tilt-scheme optimized for high resolution subtomogram averaging. *Journal of Structural Biology*, *197*(2), 191–198. https://doi.org/10.1016/j.jsb.2016.06.007

Hasler, L., Heymann, J. B., Engel, A., Kistler, J., & Walz, T. (1998). 2D crystallization of membrane proteins: Rationales and examples. *Journal of Structural Biology*, *121*(2), 162–171. https://doi.org/10.1006/jsbi.1998.3960

Henderson, R. (2013). Avoiding the pitfalls of single particle cryo-electron microscopy: Einstein from noise. *Proceedings of the National Academy of Sciences*, *110*(45), 18037–18041. https://doi.org/10.1073/pnas.1314449110

Himes, B. A., Grant, T., & Grigorieff, N. (2021). *GPU accelerated 2DTM in cisTEM - prerelease code*. https://doi.org/DOI: 10.5281/zenodo.4603401

Himes, B. A., & Grigorieff, N. (2021). Cryo-TEM simulations of amorphous radiation-sensitive samples using multislice wave propagation. *BioRxiv*. https://doi.org/https://doi.org/10.1101/2021.02.19.431636

Himes, B. A., & Zhang, P. (2018). emClarity: software for high-resolution cryo-electron tomography and subtomogram averaging. *Nature Methods*, *15*(11), 955–961. https://doi.org/10.1038/s41592-018-0167-z

Hrabe, T., Chen, Y., Pfeffer, S., Cuellar, L. K., Mangold, A., & Förster, F. (2012). PyTom : A python-based toolbox for localization of macromolecules in cryo-electron tomograms and subtomogram analysis. *Journal of Structural Biology*, *178*(2), 177–188. https://doi.org/10.1016/j.jsb.2011.12.003

Huynh, D. Q. (2009). Metrics for 3D rotations: Comparison and analysis. *Journal of Mathematical Imaging and Vision*, *35*(2), 155–164. https://doi.org/10.1007/s10851-009-0161-2

Kuhlbrandt, W. (2014). The Resolution Revolution. *Science*, *343*(6178), 1443–1444. https://doi.org/10.1126/science.1251652

Kühner, S., Van Noort, V., Betts, M. J., Leo-Madas, A., Batisse, C., Rode, M., Yamada, T., Maier, T., Bader, S., Beltran-Alvarez, P., Castaño-Diez, D., Chen, W. H., Devos, D., Güell, M., Norambuena, T., Racke, I., Rybin, V., Schmidt, A., Yus, E., … Gavin, A. C. (2009). Proteome organization in a genome-reduced bacterium. *Science*, *326*(5957), 1235–1240. https://doi.org/10.1126/science.1176343

Li, X., Mooney, P., Zheng, S., Booth, C. R., Braunfeld, 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. *Nature Methods*. https://doi.org/10.1038/nmeth.2472

Ling, C., & Ermolenko, D. N. (2016). Structural insights into ribosome translocation. *Wiley Interdisciplinary Reviews: RNA*, *7*(5), 620–636. https://doi.org/10.1002/wrna.1354

Lyumkis, D., Brilot, A. F., Theobald, D. L., & Grigorieff, N. (2013). Likelihood-based classification of cryo-EM images using FREALIGN. *Journal of Structural Biology*, *183*(3), 377–388. https://doi.org/10.1016/j.jsb.2013.07.005

Mahamid, J., Pfeffer, S., Schaffer, M., Villa, E., Danev, R., Kuhn Cuellar, L., Forster, F., Hyman, A. A., Plitzko, J. M., Baumeister, W., Cuellar, L. K., Förster, F., Hyman, A. A., Plitzko, J. M., Baumeister, W., Kuhn Cuellar, L., Forster, F., Hyman, A. A., Plitzko, J. M., & Baumeister, W. (2016). Visualizing the molecular sociology at the HeLa cell nuclear periphery. *Science*, *351*(6276), 969–972. https://doi.org/10.1126/science.aad8857

Marko, M., Hsieh, C., Schalek, R., Frank, J., & Mannella, C. (2007). Focused-ion-beam thinning of frozen-hydrated biological specimens for cryo-electron microscopy. *Nature Methods*, *4*(3), 215–217. https://doi.org/10.1038/nmeth1014

Mastronarde, D. N. (2005). Automated electron microscope tomography using robust prediction of specimen movements. *Journal of Structural Biology*, *152*(1), 36–51. https://doi.org/10.1016/j.jsb.2005.07.007

Mastronarde, D. N., & Held, S. R. (2017). Automated tilt series alignment and tomographic reconstruction in IMOD. *Journal of Structural Biology*, *197*(2), 102–113. https://doi.org/10.1016/j.jsb.2016.07.011

McDonough, R. N. (1995). *Detection of signals in noise* (2nd ed.). Academic Press, Inc.

Moebel, E., Martinez-Sanchez, A., Larivière, D., Fourmentin, E., Ortiz, J., Baumeister, W., & Kervrann, C. (2020). *Deep Learning Improves Macromolecules Localization and Identification in 3D Cellular Cryo-Electron Tomograms*. https://doi.org/10.1101/2020.04.15.042747

O'Reilly, F. J., Xue, L., Graziadei, A., Sinn, L., Lenz, S., Tegunov, D., Blötz, C., Singh, N., Hagen, W. J. H., Cramer, P., Stülke, J., Mahamid, J., & Rappsilber, J. (2020). In-cell architecture of an actively transcribing-translating expressome. *Science*, *369*(6503), 554–557. https://doi.org/10.1126/science.abb3758

Oikonomou, C. M., & Jensen, G. J. (2017). Cellular Electron Cryotomography : Toward Structural Biology In Situ. *Annu Rev Biochem*, *April*, 1–24. https://doi.org/10.1146/annurev-biochem-061516-044741

Pettersen, E. F., & Goddard, T. D. (2004). UCSF Chimera—a visualization system for exploratory research and analysis. *Journal of …*. https://doi.org/10.1002/jcc.20084

Pfeffer, S., Delarue, M., Brittingham, G. P., Pfeffer, S., Surovtsev, I. V, Pinglay, S., Kennedy, K. J., Schaffer, M., & Gutierrez, J. I. (2018). mTORC1 Controls Phase Separation and the Biophysical Properties of the Cytoplasm by Tuning Article mTORC1 Controls Phase Separation and the Biophysical Properties of the Cytoplasm by Tuning Crowding. *Cell*, *174*(2), 338-349.e20. https://doi.org/10.1016/j.cell.2018.05.042

Pfeffer, S., & Mahamid, J. (2018). Unravelling molecular complexity in structural cell biology. *Current Opinion in Structural Biology*, *52*, 111–118. https://doi.org/10.1016/j.sbi.2018.08.009

Rickgauer, J. P., Choi, H., Lippincott-Schwartz, J., & Denk, W. (2020). Label-free single-instance protein detection in vitrified cells. *BioRxiv*. https://doi.org/10.1101/2020.04.22.053868

Rickgauer, J. P., Grigorieff, N., & Denk, W. (2017). Single-protein detection in crowded molecular environments in cryo-EM images. *ELife*, *6*, 1–22. https://doi.org/10.7554/eLife.25648

Rigort, A., Bäuerlein, F. J. B., Villa, E., Eibauer, M., Laugks, T., Baumeister, W., & Plitzko, J. M. (2012). Focused ion beam micromachining of eukaryotic cells for cryoelectron tomography. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(12), 4449–4454. https://doi.org/10.1073/pnas.1201333109

Rohou, A., & Grigorieff, N. (2015). CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *Journal of Structural Biology*, *192*(2), 216–221.

https://doi.org/10.1016/j.jsb.2015.08.008

Rosenthal, P. B., & Henderson, R. (2003). Optimal Determination of Particle Orientation, Absolute Hand, and Contrast Loss in Single-particle Electron Cryomicroscopy. *Journal of Molecular Biology*, *333*(4), 721–745. https://doi.org/10.1016/j.jmb.2003.07.013

Sachse, C., Chen, J. Z., Coureux, P., Stroupe, M. E., Fändrich, M., & Grigorieff, N. (2008). Fresh Look at Tobacco Mosaic Virus. *Journal of Molecular Biology*, *371*(3), 812–835.

Sanchez, R. M., Zhang, Y., Chen, W., Dietrich, L., & Kudryashev, M. (2020). Subnanometer-resolution structure determination in situ by hybrid subtomogram averaging - single particle cryo-EM. *Nature Communications*, *11*(1), 1–7. https://doi.org/10.1038/s41467-020-17466-0

Scheres, S. H. W. (2012). RELION: Implementation of a Bayesian approach to cryo-EM structure determination. *Journal of Structural Biology*, *180*(3), 519–530. https://doi.org/10.1016/j.jsb.2012.09.006

Scheres, S. H. W. (2015). Semi-automated selection of cryo-EM particles in RELION-1.3. *Journal of Structural Biology*, *189*(2), 114–122. https://doi.org/10.1016/j.jsb.2014.11.010

Scheres, S. H. W., & Chen, S. (2012). Prevention of overfitting in cryo-EM structure determination. *Nature Methods*, *9*(9), 853–854. https://doi.org/10.1038/nmeth.2115

Schur, F. K. M., Hagen, W. J. H., de Marco, A., & Briggs, J. a G. (2013). Determination of protein structure at 8.5Åresolution using cryo-electron tomography and sub-tomogram averaging. *Journal of Structural Biology*, *184*(3), 394–400. https://doi.org/10.1016/j.jsb.2013.10.015

Sigworth, F. J. (2004). Classical detection theory and the cryo-EM particle selection problem. *Journal of Structural Biology*, *145*(1–2), 111–122. https://doi.org/10.1016/j.jsb.2003.10.025

Sigworth, F. J., Doerschuk, P. C., Carazo, J. M., & Scheres, S. H. W. (2010). An introduction to maximum-likelihood methods in cryo-EM. In *Methods in Enzymology* (1st ed., Vol. 482, Issue C). Elsevier Inc. https://doi.org/10.1016/S0076-6879(10)82011-7

Sohmen, D., Chiba, S., Shimokawa-Chiba, N., Innis, C. A., Berninghausen, O., Beckmann, R., Ito, K., & Wilson, D. N. (2015). Structure of the Bacillus subtilis 70S ribosome reveals the basis for species-specific stalling. *Nature Communications*, *6*, 1–10. https://doi.org/10.1038/ncomms7941

Spahn, C. M. T., Penczek, P. A., Leith, A., & Frank, J. (2000). A method for differentiating proteins from nucleic acids in intermediate-resolution density maps: Cryo-electron microscopy defines the quaternary structure of the Escherichia coli 70S ribosome. *Structure*, *8*(9), 937–948. https://doi.org/10.1016/S0969-2126(00)00185-4

Stewart, A., & Grigorieff, N. (2004). Noise bias in the refinement of structures derived from single particles. *Ultramicroscopy*, *102*(1), 67–84. https://doi.org/10.1016/j.ultramic.2004.08.008

1199 Strunk, K. M., Wang, K., Ke, D., Gray, J. L., & Zhang, P. (2012). Thinning of large mammalian
1200 cells for cryo-tem characterization by cryo-fib milling. *Journal of Microscopy*, *247*(3), 220–
1201 227. https://doi.org/10.1111/j.1365-2818.2012.03635.x

1202 Subramaniam, S. (2013). Structure of trimeric HIV-1 envelope glycoproteins. *PNAS*, *110*(45),
1203 4172–4174. https://doi.org/10.1073/pnas.1313802110

1204 Tang, G., Peng, L., Baldwin, P. R., Mann, D. S., Jiang, W., Rees, I., & Ludtke, S. J. (2007).
1205 EMAN2 : An extensible image processing suite for electron microscopy. *Journal of*
1206 *Structural Biology*, *157*, 38–46. https://doi.org/10.1016/j.jsb.2006.05.009

1207 Tegunov, D., Xue, L., Dienemann, C., Cramer, P., & Mahamid, J. (2021). Multi-particle cryo-
1208 EM refinement with M visualizes ribosome-antibiotic complex at 3.5 Å in cells. *Nature*
1209 *Methods*, *18*(2), 186–193. https://doi.org/10.1038/s41592-020-01054-7

1210 Tichelaar, W., Hagen, W. J. H., Gorelik, T. E., Xue, L., & Mahamid, J. (2020). TEM bright field
1211 imaging of thick specimens: nodes in Thon ring patterns. *Ultramicroscopy*, *216*(November
1212 2019), 113023. https://doi.org/10.1016/j.ultramic.2020.113023

1213 Van Heel, M. (2013). Finding trimeric HIV-1 envelope glycoproteins in random noise. *PNAS*,
1214 *110*(45), 4175–4177. https://doi.org/10.1073/pnas.1314353110

1215 Voortman, L. M., Vulović, M., Maletta, M., Voigt, A., Franken, E. M., Simonetti, A., Peters, P.
1216 J., van Vliet, L. J., & Rieger, B. (2014). Quantifying resolution limiting factors in
1217 subtomogram averaged cryo-electron tomography using simulations. *Journal of Structural*
1218 *Biology*, *187*(2), 103–111. https://doi.org/10.1016/j.jsb.2014.06.007

1219 Wan, W., & Briggs, J. A. G. (2016). Cryo-Electron Tomography and Subtomogram Averaging.
1220 In *The Resolution Revolution: Recent Advances In cryoEM* (1st ed., Vol. 579). Elsevier Inc.
1221 https://doi.org/10.1016/bs.mie.2016.04.014

1222 Winkler, H., Zhu, P., Liu, J., Ye, F., Roux, K. H., & Taylor, K. A. (2009). Tomographic
1223 subvolume alignment and subvolume classification applied to myosin V and SIV envelope
1224 spikes. *Journal of Structural Biology*, *165*(2), 64–77.
1225 https://doi.org/10.1016/j.jsb.2008.10.004

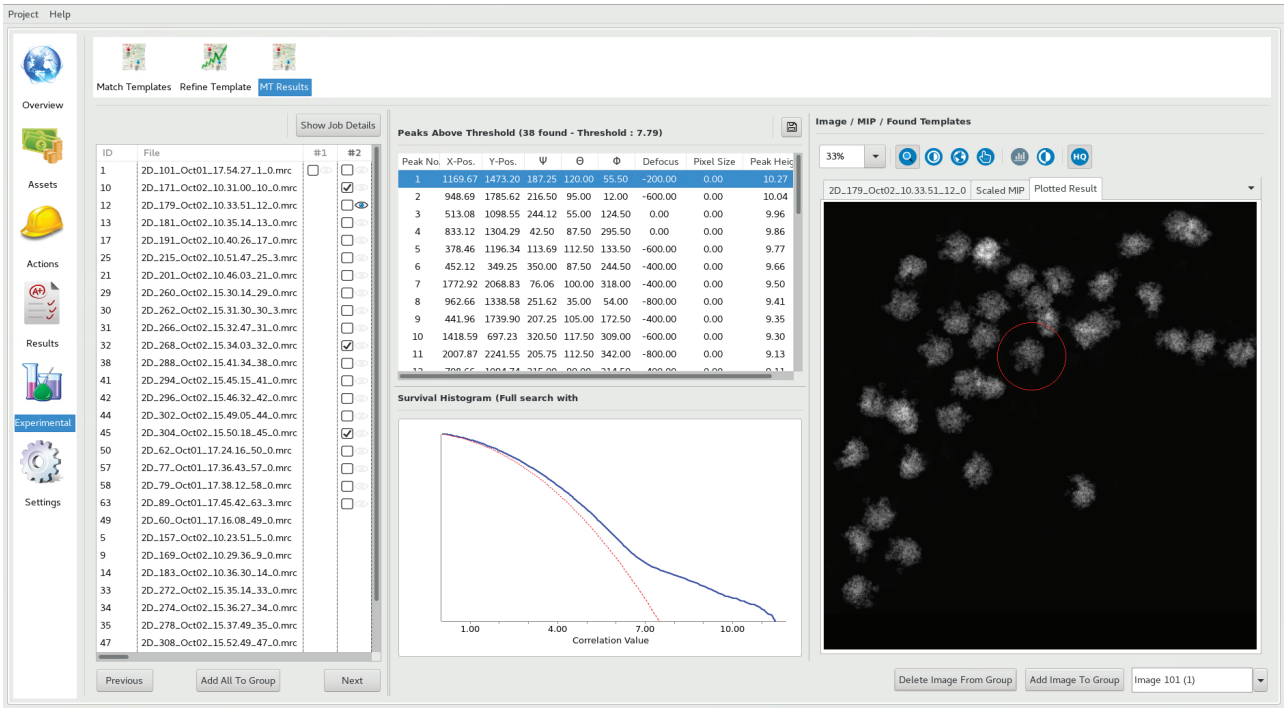
1226 Zhang, P. (2019). Advances in cryo-electron tomography and subtomogram averaging and
1227 classification. *Current Opinion in Structural Biology*, *58*, 249–258.
1228 https://doi.org/10.1016/j.sbi.2019.05.021

1229 Zivanov, J., Nakane, T., Forsberg, B. O., Kimanius, D., Hagen, W. J. H., Lindahl, E., & Scheres,
1230 S. H. W. (2018). New tools for automated high-resolution cryo-EM structure determination
1231 in RELION-3. *ELife*, *7*, 1–22. https://doi.org/10.7554/eLife.42166

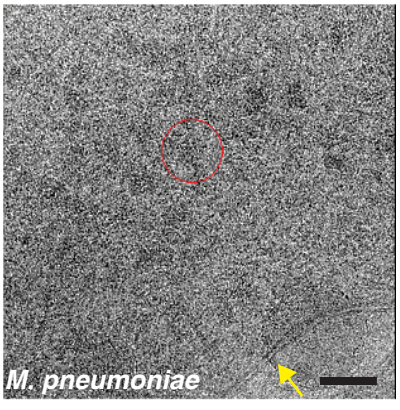
1232

Figure 1

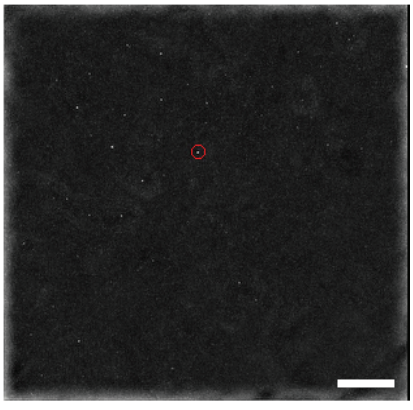
(a)



(b)



(c)



(d)

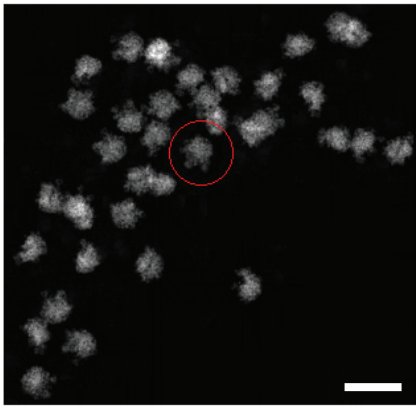


Figure 1.S1

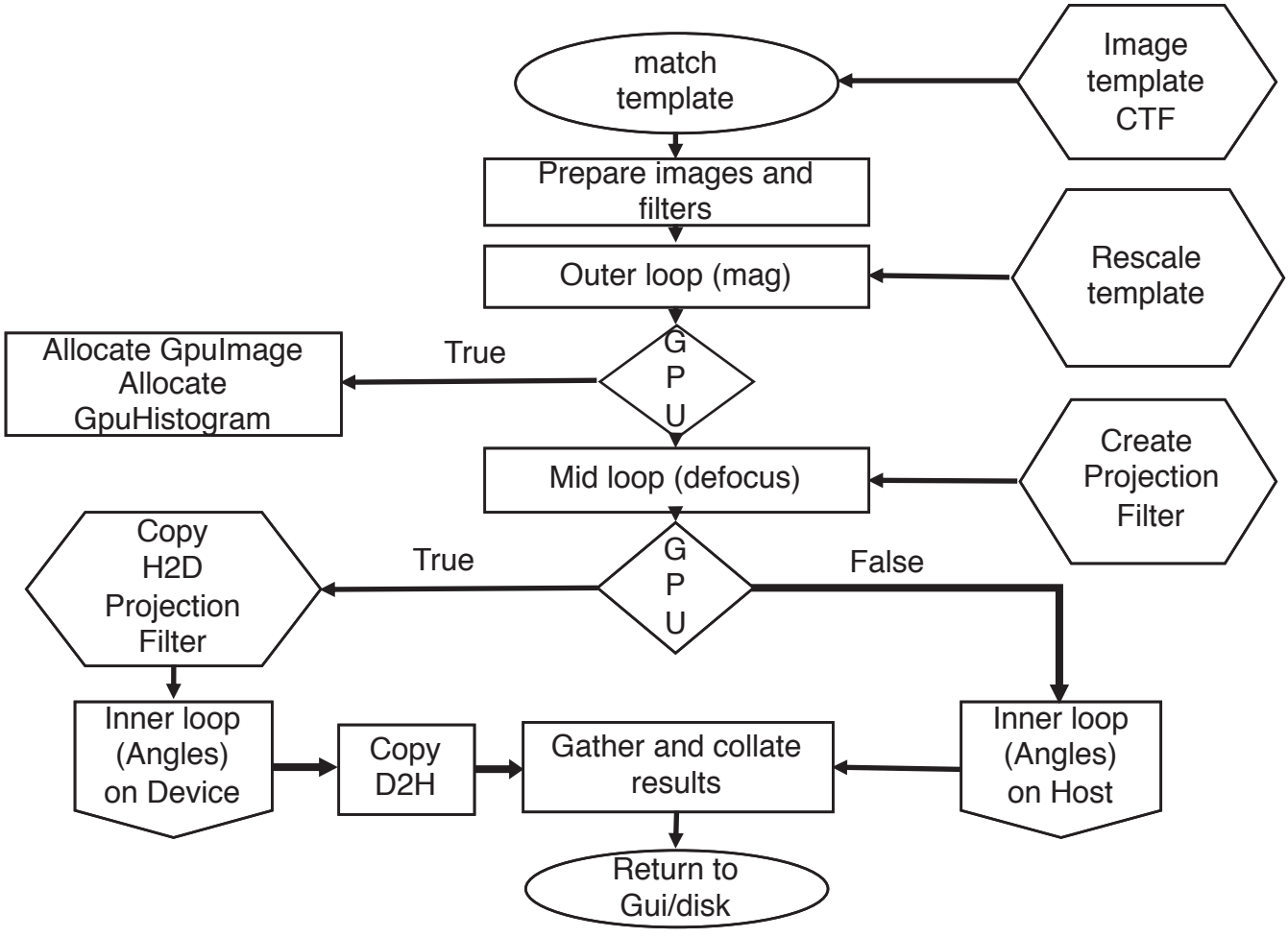


Figure 2

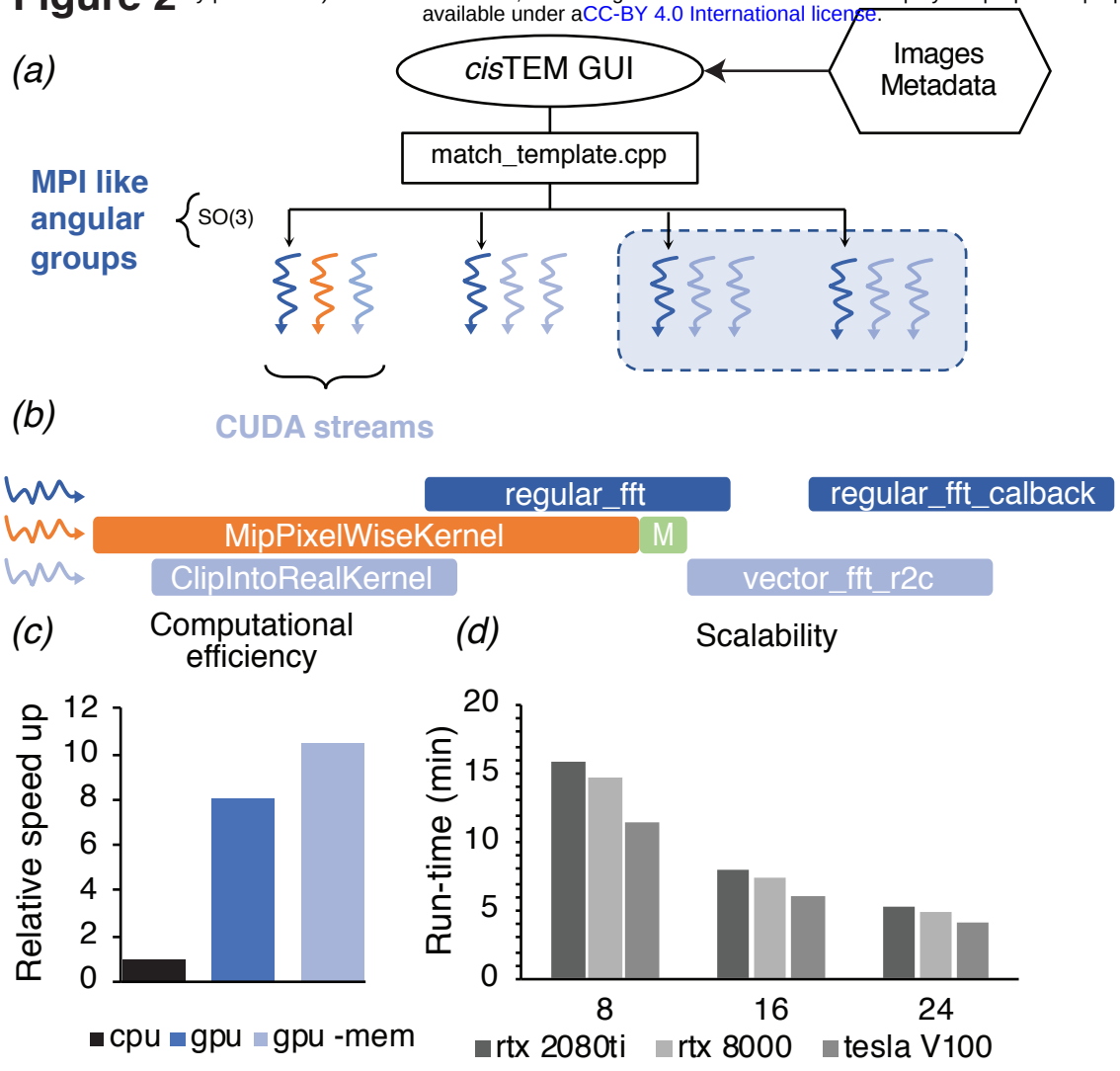


Figure 2: S1

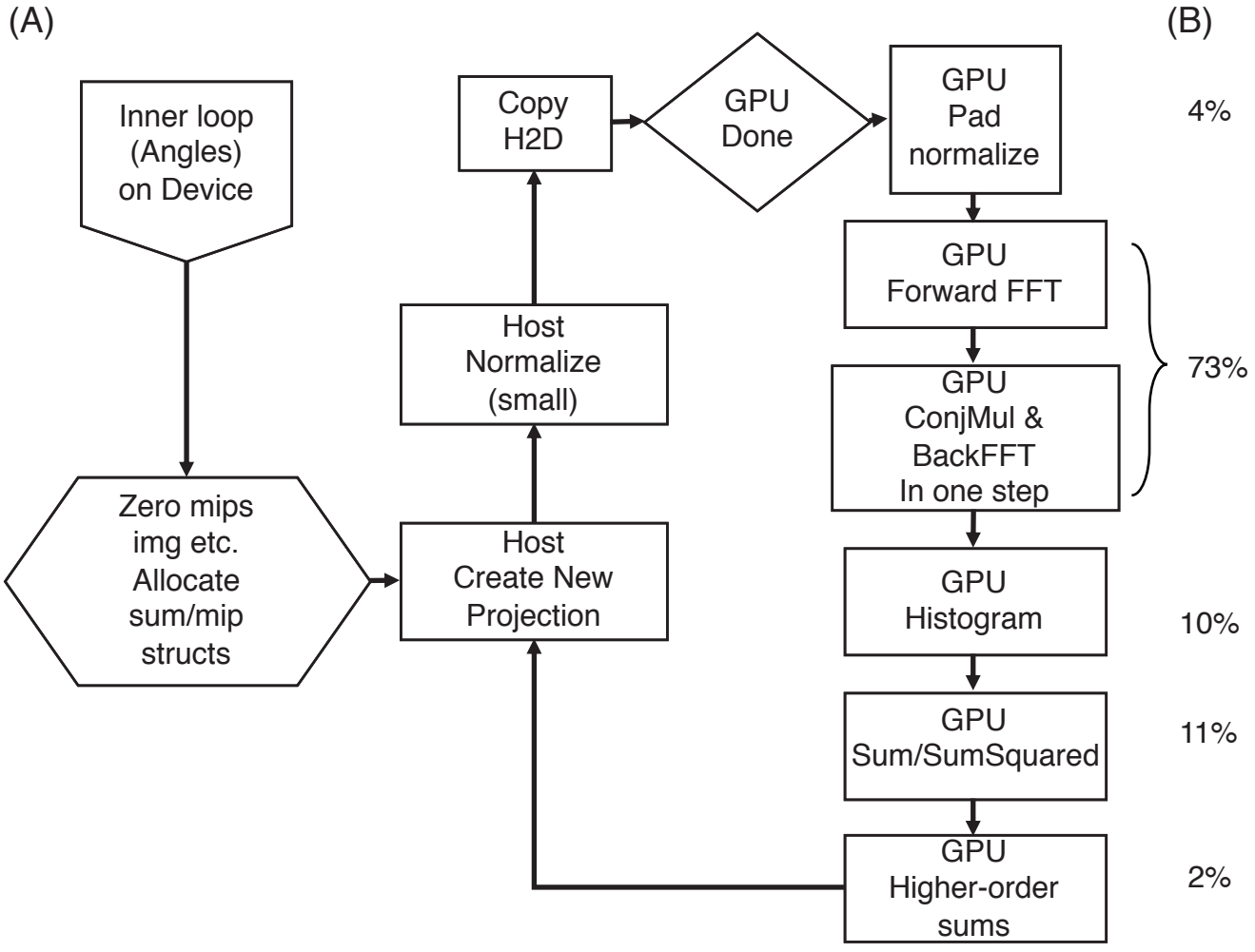


Figure 3

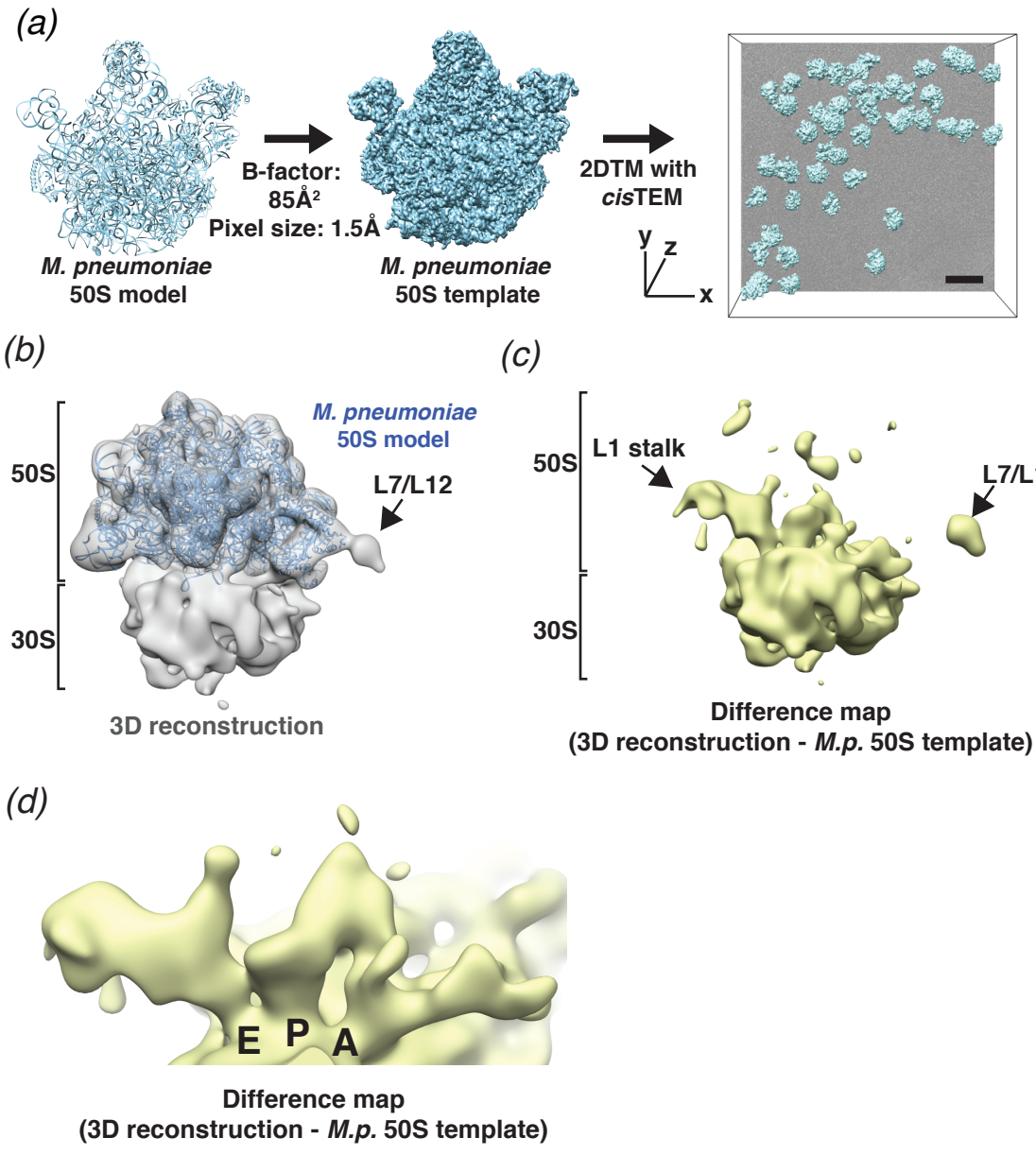


Figure 3:S1

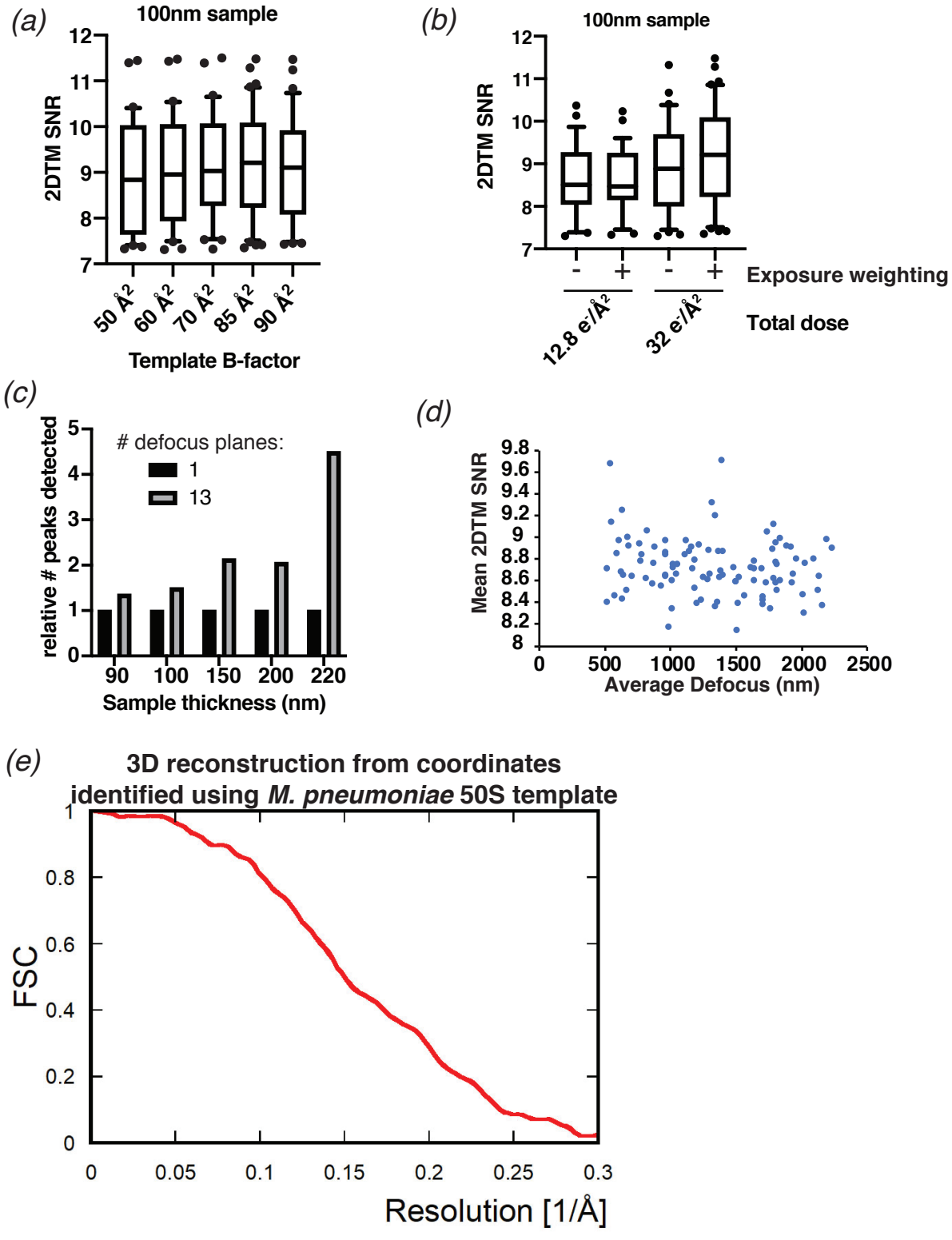


Figure 4

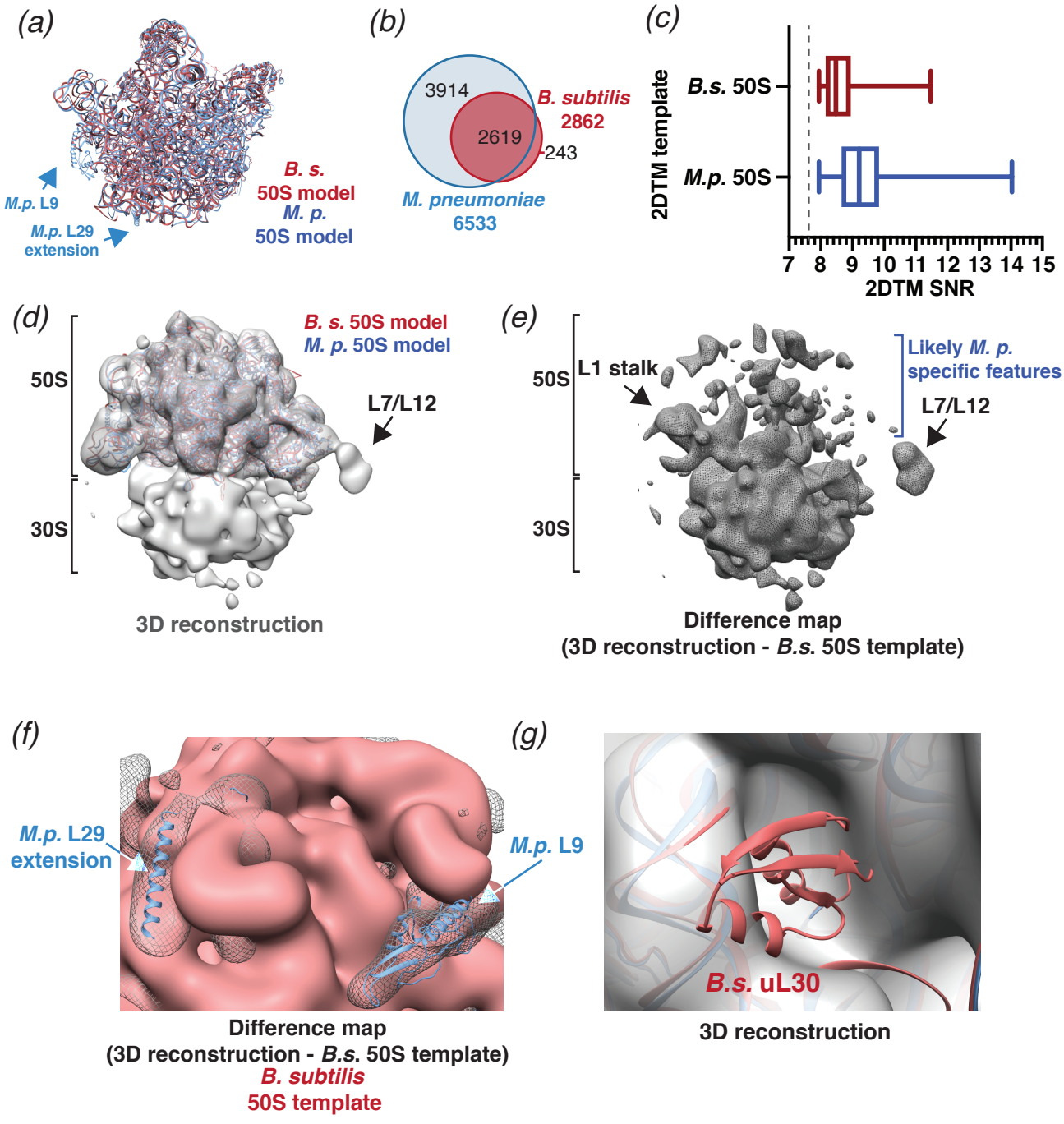


Figure 4.S1

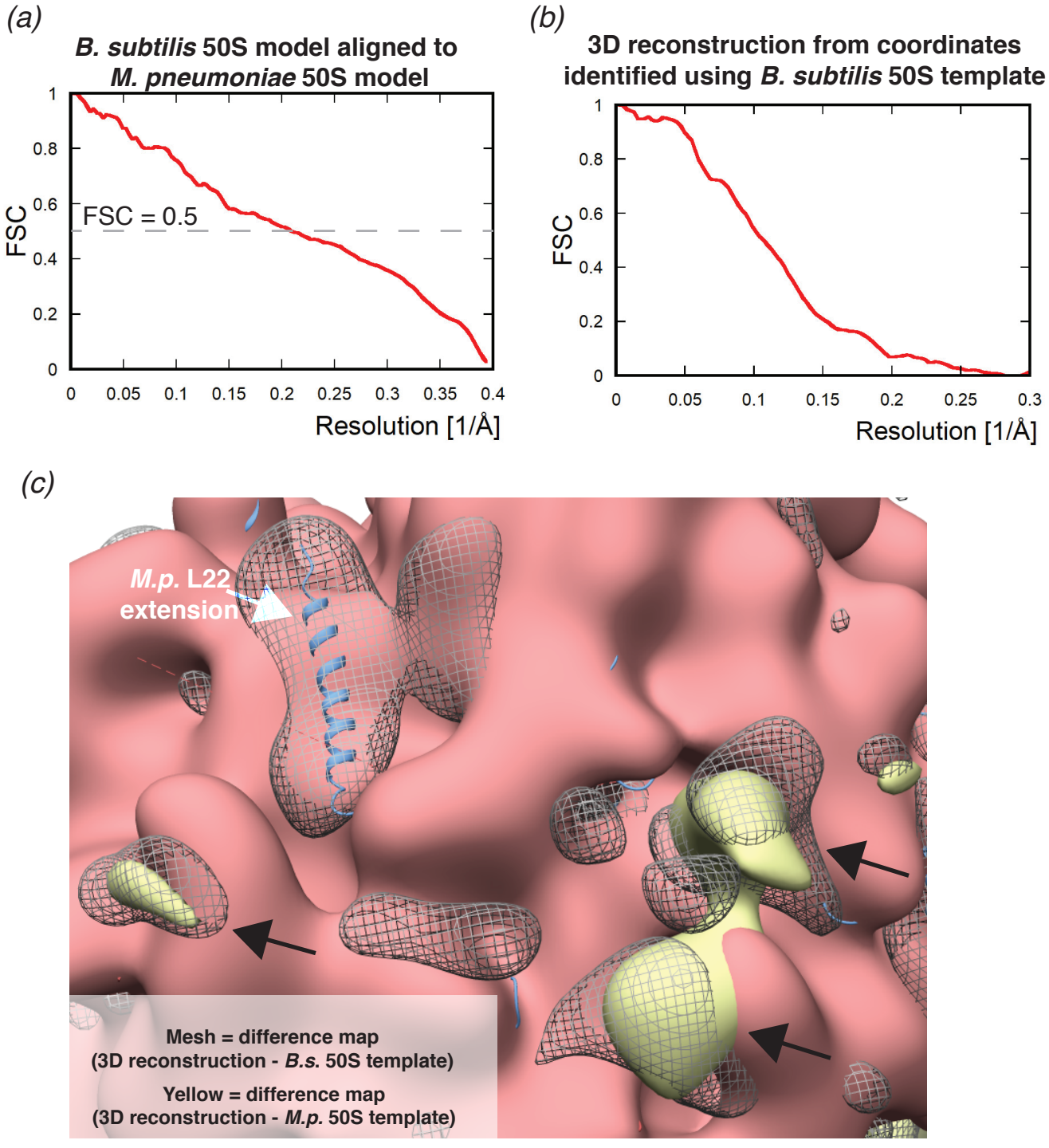


Figure 5

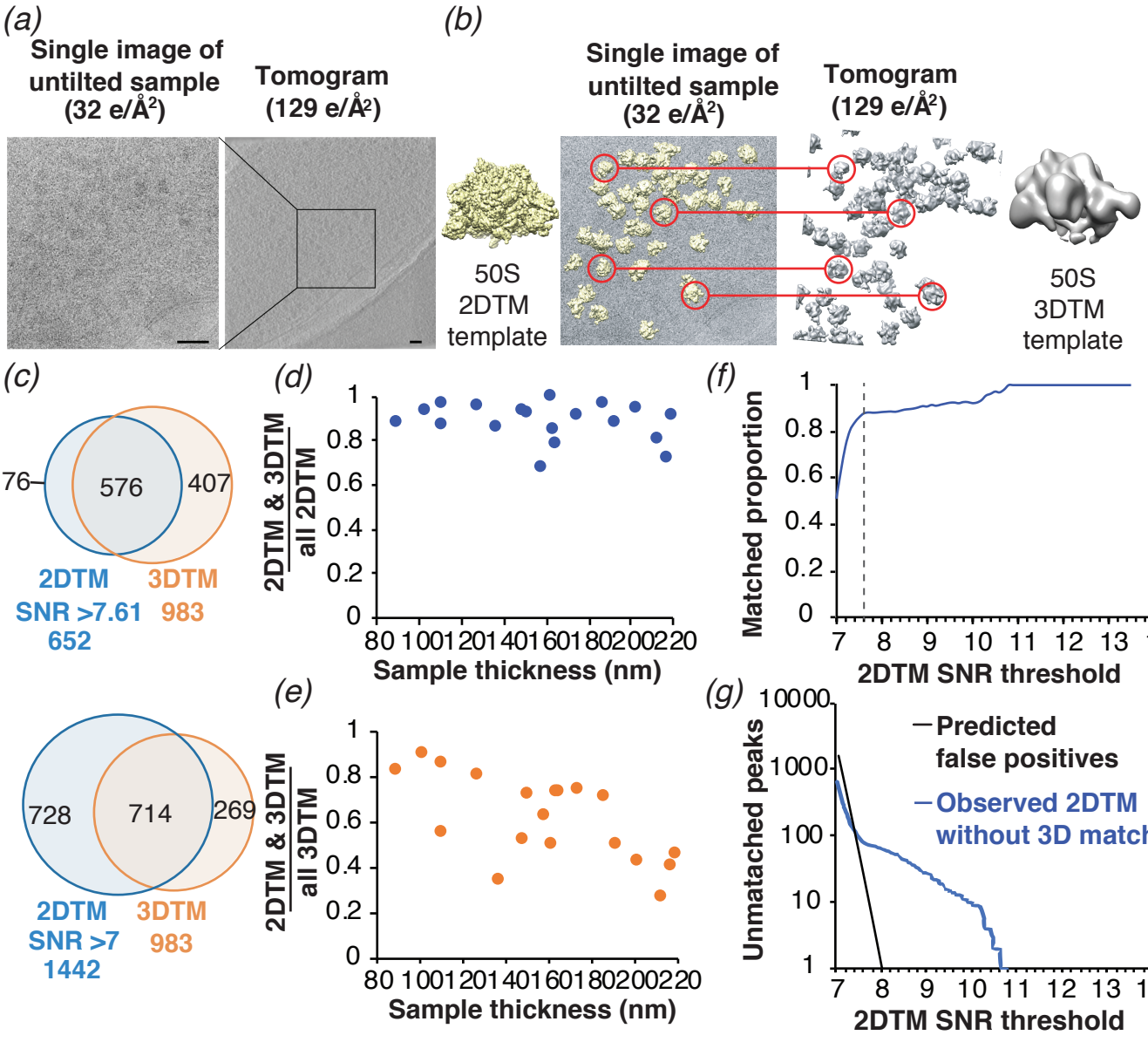


Figure 5- S1

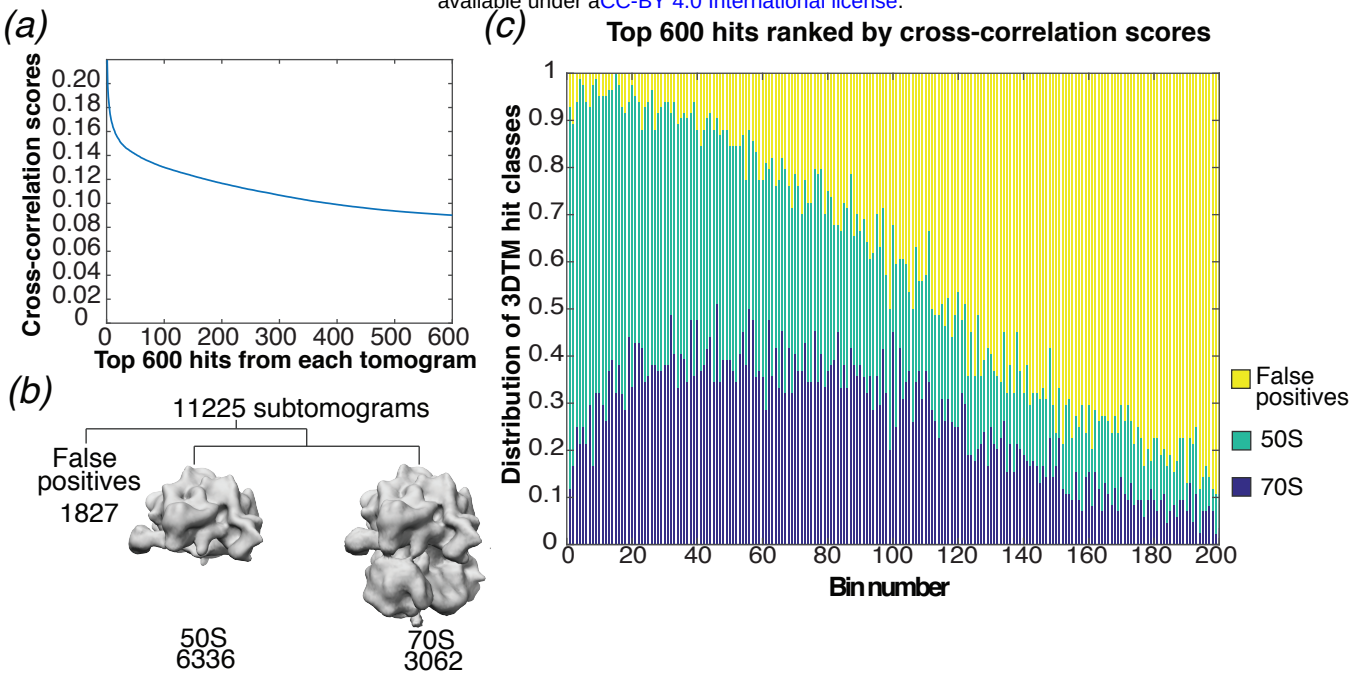


Figure 5-S2

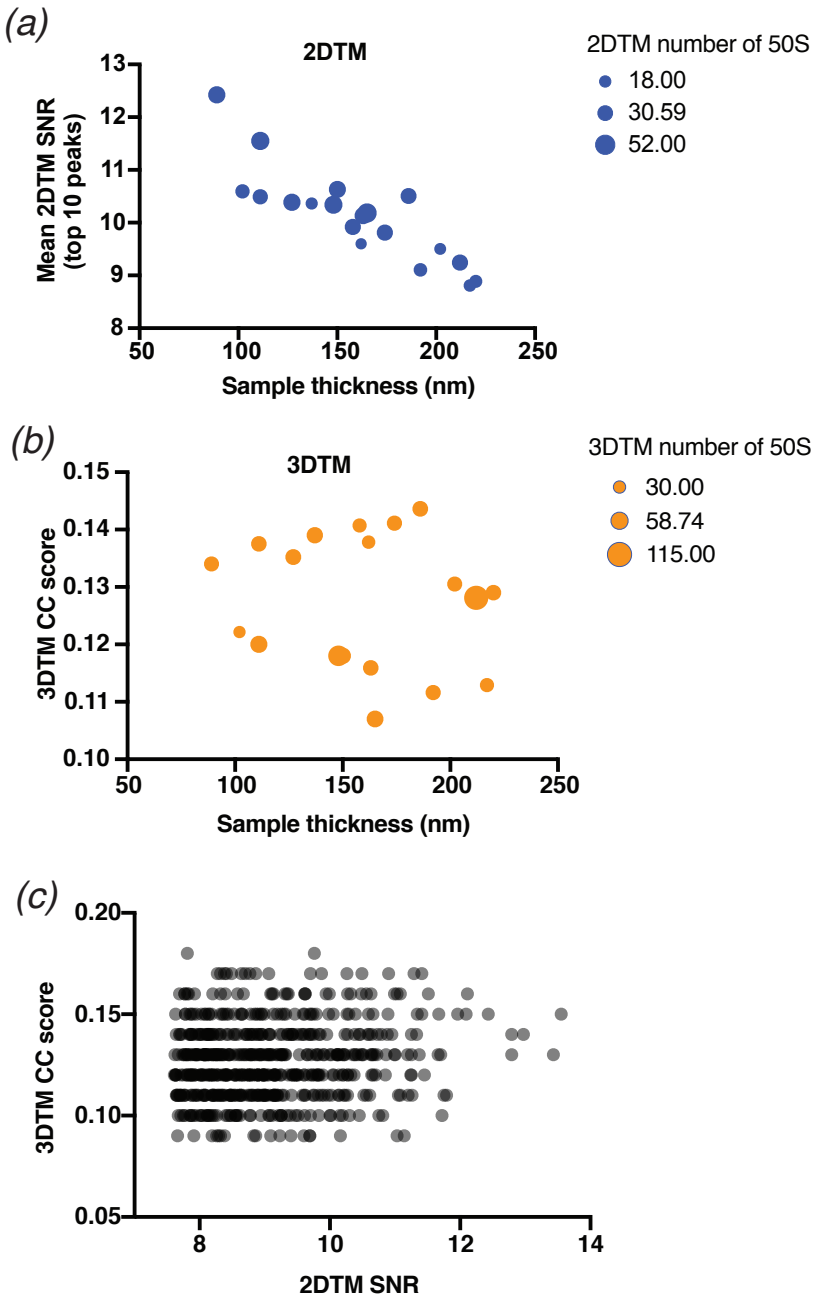


Figure 6

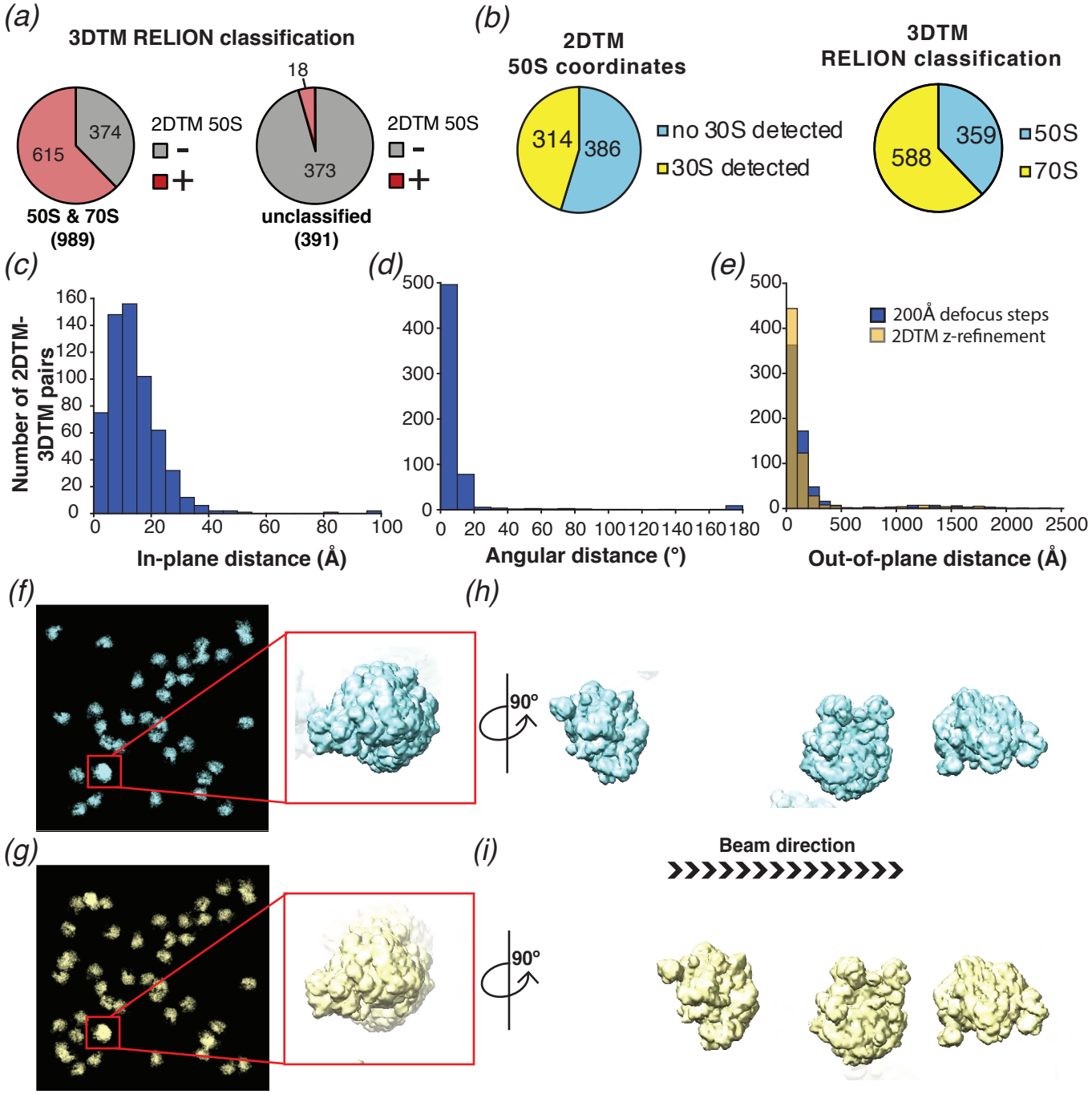


Figure S6: S1