bioRxiv preprint doi: https://doi.org/10.1101/838920; this version posted November 12, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Topaz-Denoise: general deep denoising models for cryoEM

Tristan Bepler[1,2], Alex J. Noble[3,*], and Bonnie Berger[2,4,*]

[1] Computational and Systems Biology, MIT, Cambridge, MA, USA

[2] Computer Science and Artificial Intelligence Laboratory, MIT, Cambridge, MA, USA

[3] National Resource for Automated Molecular Microscopy, Simons Electron Microscopy Center, New York Structural Biology Center, NY, NY, USA

[4] Department of Mathematics, MIT, Cambridge, MA, USA

* Corresponding authors: bab@mit.edu and anoble@nysbc.org

Abstract

Cryo-electron microscopy (cryoEM) is becoming the preferred method for resolving protein structure. Low signal-to-noise (SNR) in cryoEM images reduces the confidence and throughput of structure determination during several steps of data processing, resulting in impediments such as missing particle orientations. Denoising cryoEM images can not only improve downstream analysis but also accelerate the time-consuming data collection process by allowing lower electron dose micrographs to be used for analysis without compromising structural interpretability. Here, we present Topaz-Denoise, a deep learning method for reliably increasing the SNR of cryoEM images in seconds. By training on a dataset composed of thousands of micrographs collected across a wide range of imaging conditions, we are able to learn models capturing the complexity of the cryoEM image formation process. While this general idea has been deployed successfully in natural imaging, protein threading, and proteomics, it has yet to be applied systematically in cryoEM where the lack of ground truth signal has been a long-standing limitation. To address this, we make the key insight that forming paired, independent micrographs from even and odd camera movie frames enables us to train denoising models without observing ground truth signal. We demonstrate that our denoising model improves SNR by roughly 100x over raw micrographs and 1.8x over other methods. Notably, we show that denoising with our general model enables solving the first 3D single particle structure of clustered protocadherin, an elongated particle with previously-elusive views. Topaz-Denoise and pre-trained general models are now included in Topaz (https://github.com/tbepler/topaz), a free and open-source software package that focuses on particle picking, and has been integrated into the Appion cryoEM suite. We expect that Topaz-Denoise will be of broad utility to the cryoEM community for improving micrograph interpretability and accelerating analysis.

bioRxiv preprint doi: https://doi.org/10.1101/838920; this version posted November 12, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Introduction

Visualization of micrographs from cryo-electron microscopy (cryoEM) of biological specimens is primarily limited by the phase contrast of proteins, the low electron dose conditions required due to radiation damage accrued by the proteins, and the thickness of the ice. As researchers push towards smaller and smaller proteins, these issues hinder downstream analysis because these proteins become increasingly difficult to distinguish from noise. Certain orientations of larger, non-globular proteins can also have low signal, leading to missing views. The typical signal-to-noise ratio (SNR) of a cryoEM micrograph is estimated to be as high as 0.1[1], amongst the lowest in any imaging field, and no ground truth exists. Several steps during collection and processing of micrographs in single particle cryoEM rely on properly human-inspecting micrographs, identifying particles, and examining processed data. Conventional cryoEM methods for improving contrast in micrographs include downsampling, bandpass filtering, and Wiener filtering[2,3]. However, these methods do not address the specific noise properties of micrographs and often do not provide interpretable results, which is increasingly becoming an issue as researchers attempt to resolve small and non-globular proteins[4,5].

Image denoising has long been a topic of significant interest in the machine learning community[6–8]. Advances in deep neural networks have enabled substantial improvements in image restoration and inpainting (i.e. filling in missing pixels) by learning complex, non-linear priors over the applied image domain. However, these methods require ground truth images to provide supervision for learning the denoising model[9,10], and is hence limited to domains where ground truth is available. To overcome this barrier, Lehtinen et al.[11] presented a general machine learning (ML) framework, called Noise2Noise, for learning denoising models from paired noisy images rather than paired noisy and ground truth images. This method has been followed by several others for learning denoising models without ground truth[12–14]. These methods offer new approaches for training deep neural network models for denoising in challenging domains. In cryo-EM, neural network denoising software has only just started to emerge for dataset-by-dataset tomogram denoising[15,16] and single particle micrograph denoising[17]. However, there have not been any systematic evaluation of these methods to date nor general denoising models developed.

Here, we develop Topaz-Denoise, the first large-scale, publicly available denoising models for cryoEM. Conventional cryoEM denoising methods are ad-hoc filters that do not model the complex image generative process. However, deep denoising models typically require ground truth signal which is not available in cryoEM. We make the key insight that the individual movie frames collected by modern microscope cameras are many independent observations of the same underlying signal and, hence, can be used to learn denoising models directly via the Noise2Noise framework. Trained on thousands of micrographs across a variety of imaging conditions, these models provide robust denoising without the need to train on a dataset-by-dataset basis. We test and compare these denoising models on several micrographs of typical particles and of small particles, study improvements in SNR, and use denoising combined with Topaz particle picking[18] to obtain the first 3D single particle cryoEM structure of clustered protocadherin, an elongated particle with previously-elusive views. We also show, for the first time, that denoising enables more rapid data collection by allowing micrographs to be collected with a lower electron total dose (10%–25% typical exposure times) without sacrificing interpretability or downstream processing. Shorter exposure times allow for higher throughput

bioRxiv preprint doi: https://doi.org/10.1101/838920; this version posted November 12, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

microscope usage, which reduces research cost and increases research efficiency. These models are integrated into Topaz allowing easy access to the community.

Topaz-Denoise source code is freely available as part of Topaz (http://topaz.csail.mit.edu) and can be installed through Anaconda, Pip, Docker, Singularity, and SBGrid[19], and is now integrated into Appion[20]. As with Topaz, Topaz Denoise is designed to be modular and can easily be integrated into other cryoEM software suites. Topaz Denoise includes several pre-trained models and the ability for the user to train their own models. Topaz Denoise training and inference runs efficiently on a single GPU computer and is provided together with the standalone Topaz GUI to assist with command generation.

Methods

Training dataset preparation

To train the denoising model, we collected a large dataset of micrograph frames from public repositories[21] and internal datasets at the New York Structural Biology Center (NYSBC), as described in Supplementary Table 1. These micrograph frames were collected under a large variety of imaging conditions and contain data collected on FEI Krios, FEI Talos Arctica, and JEOL CRYOARM300 microscopes with Gatan K2 and FEI Falcon II cameras at both super-resolution (K2) and counting modes and at many defocus levels. Including several microscopes, cameras, and datasets allows for robust denoising parameters to be modelled across common microscope setups.

We form two general aggregated datasets, one we call “Large” and one called “Small”. The “Large” dataset contains micrographs from all individual datasets. To roughly balance the contribution of the individual datasets in these aggregate datasets, we randomly select up to 200 micrographs from each individual dataset for inclusion rather than all micrographs. The Small dataset contains micrographs from individual datasets selected by eye based on the denoising performance of individually-trained U-net denoising models.

The Noise2Noise framework requires paired noisy observations of the same underlying signal. We generate these pairs by splitting the micrograph frames into even/odd frames which represent independent observations. These even/odd micrograph frames are then summed directly to form the paired observations. Because micrographs are typically motion corrected before summing and this motion correction procedure can change the noise distribution of the micrographs, we also form aligned, summed micrograph pairs by aligning the even/odd micrograph frames with MotionCor2[22] using 5 by 5 patches and b-factor of 100. This resulted in 1,929 paired micrographs for the Small dataset and 3,439 paired micrographs for the Large dataset.

Model architectures

We adopt a U-Net model architecture[23] similar to that used by Lehtinen et al.[11] except that the input and output feature maps are 1-dimensional (n=1 to match monochrome micrographs) and we replace the first two width 3 convolutional layers of Lehtinen et al. with a single width 11 convolutional layer. This model contains five max pooling downsampling blocks and five nearest-neighbor upsampling blocks with skip connections between down- and up-sampling blocks at each spatial resolution. We refer to this as the U-net model. For comparison, we also

bioRxiv preprint doi: https://doi.org/10.1101/838920; this version posted November 12, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

consider a smaller U-net model with only 3 downsampling and upsampling blocks which we refer to as the U-net (small) model. We also compare with a fully convolutional neural network consisting of three convolutional layers of width 11x11 with 64 filters each and leaky rectified linear unit activations, termed FCNN, and an affine model with a single convolutional filter of width 31x31.

Loss functions and the Noise2Noise framework

The Noise2Noise framework takes advantage of the observation that we can learn models that recover statistics of the noise distribution given paired noisy observations of the same underlying signal. Given a ground truth signal, *y*, we observe images of this signal that have been corrupted by some probabilistic noise process, $x \sim Noise(y)$. Given paired noisy observations for matched signals, $x_a \sim Noise(y)$ and $x_b \sim Noise(y)$, we can learn a function that recovers statistics of this distribution. This is accomplished by finding parameters of the denoising function, *f* with parameters θ, such that the error between the denoised sample $f(x_a)$ and raw x_b are minimized. The form of this error function determines what statistics of the noise distribution we learn to recover. Given a dataset, *X*, containing many such image pairs, minimizing the L2 error over paired samples,

$$argmin_\theta \, E_{x_z, x_b \sim X}[\, \|f(x_a) - x_b\|_2^2 \,],$$

finds *f* with mean-seeking behaviour. Minimizing the L1 error over paired samples,

$$argmin_\theta \, E_{x_z, x_b \sim X}[\, \|f(x_a) - x_b\|_1 \,],$$

finds *f* with median-seeking behaviour. Finally, minimizing the L0 error over paired samples,

$$argmin_\theta \, E_{x_z, x_b \sim X}[\, \|f(x_a) - x_b\|_0 \,],$$

finds *f* with mode-seeking behaviour. This last objective is not differentiable and requires a smoothing term to minimize with standard gradient descent. We refer the reader to Lehtinen et al.[11] for details on this training objective.

Training details

For neural networks, weights are initialized using the default initialization in PyTorch[24]. For affine models, weights are initialized to zero. All models are trained using the Adagrad variant of stochastic gradient descent[24] with a learning rate of 0.001 for 100 epochs. We train on 800 by 800 patches randomly sampled from each micrograph using a minibatch size of 4. As data augmentation during, these patches are randomly rotated and mirrored. Images are first normalized at the whole micrograph level by subtracting the mean pixel intensity and dividing by the standard deviation of the pixel intensities. Models were trained on a single NVIDIA V100 GPU with 32 GB of VRAM. Training took about 15 hours per model.

Inference details

Given a trained denoising model, we denoise full size micrographs. When operating on a GPU, RAM constraints may require denoising to be performed in patches. Here, we denoise in patches of 4,000 by 4,000 pixels. In order to avoid artifacts that can occur at the patch edges when stitched together, we include padding of 500 pixels around each patch when denoising. Whole micrographs are first normalized by subtracting the mean and dividing by the standard

bioRxiv preprint doi: https://doi.org/10.1101/838920; this version posted November 12, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

deviation of the pixel intensities. The pixel intensities of the denoised micrograph are then restored by multiplying by the standard deviation and adding back the mean. Given the trained denoising model, inference is fast. We are able to denoise 4k by 4k K2 images at a speed of about 3 seconds/micrograph on a single NVIDIA 1080 Ti.

Signal-to-noise quantification

We quantify the SNR of raw micrographs and processed micrographs based on paired labeled signal and background regions. To this end, we hand labeled 20 signal and paired background regions across up to 10 micrographs from each dataset. We sought to label a variety of signal regions and to select paired background regions as close as possible to each signal region. Labeling was performed with reference to low-pass filtered micrographs in order to prevent any possible bias towards our denoising models. Given *N* signal, background region pairs, $x_s{}^i$, $x_b{}^i$, indexed by *i*, we calculate the mean and variance of each background region, $\mu_b{}^i$ and $v_b{}^i$. From this, the signal for each region pair is defined as $s^i = (\mu_s{}^i - \mu_b{}^i)^2$ where $\mu_s{}^i$ is the mean of signal region *i*. We then calculate the average SNR in dB for the regions,

$$SNR = \frac{10}{N} \sum_{i=1}^{N} log_{10}(s^i) - log_{10}(v^i),$$

which is reported for each dataset given raw and denoised micrographs.

Short exposure micrograph processing

To quantify our ability to improve interpretability of low electron dose micrographs, we selected between five and ten random micrographs for the four datasets presented (EMPIAR-10234, 18sep08d, 19jan04d, and 19may10e). Micrographs from each dataset were split into five frame-fractionated subsets using IMOD's newstack program[25] to simulate short exposures: 10%, 25%, 50%, 75%, and 100%. Frames were aligned with Motioncor2 using 5x5 patches and dose weighting. For each dataset, SNR quantification was performed as previously described.

Short exposure apoferritin processing

To quantify downstream results from frame titration, 100 random independently frame-aligned fractionated micrographs of 19jan04d were prepared as using Motioncor2 without dose weighting. CTF estimation of the resulting 500 frame aligned micrographs was performed using CTFFind4[26] from within Appion[20]. 9,373 particles were picked from the micrographs using the first 10% of frames, an initial model was created in Cryosparc, and the particles were refined through homogeneous refinement. The same particle picks and initial model were then used to extract and process the 25%, 50%, 75%, and 100% subsets through de novo homogeneous refinement while retaining each independent micrograph CTF estimation. 3DFSC[27] plots were then generated from the results.

EMPIAR-10234 clustered protocadherin single particle processing

We processed the EMPIAR-10234 clustered protocadherin dataset in two seperate ways to test the whether picking in denoised micrographs was advantageous: First by using the particle

bioRxiv preprint doi: https://doi.org/10.1101/838920; this version posted November 12, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

picks provided by the data owner, and second by manually picking on the denoised micrographs.

The picking method used by the data owner are described in Brasch et al.[28] Briefly, 1,540 particles were manually picked by the data owner from 87 raw micrographs and used to train a Topaz picking model, resulting in 14,569 particles. The following reconstruction workflow was performed in CryoSparc v2[29] using C1 symmetry in every step and using frame-summed particles for consistency. 2D classification was performed three times to remove obvious non-particle classes, resulting in 13,739 particles. Ab-initio reconstruction with 2 classes was performed, resulting in one good class with 10,010 particles. Homogeneous refinement was performed resulting in the final reconstruction.

The picking method we used is as follows. Frame-summed micrographs were denoised with the Topaz-Denoise v0.2.1 L2 model, proprocessed with `topaz preprocess` while binning by a factor of 4, and 1,023 particles were manually picked *not* by the data owner from 215 denoised micrographs. A Topaz picking model was trained using the particle coordinates on raw micrographs, resulting in 59,273 particles. The following reconstruction workflow was performed in CryoSparc v2 using C1 symmetry in every step and using frame-summed particles for consistency. 2D classification was performed three times to remove obvious non-particle classes, resulting in 44,303 particles. Ab-initio reconstruction with 2 classes was performed, resulting in one good class with 23,695 particles. Heterogeneous refinement with 2 classes was performed, resulting in one good class with 16,049 particles. Homogeneous refinement was performed resulting in the final reconstruction.

Clustered protocadherin low particle number single particle processing

Denoising and picking were performed as described in the previous section. Then 1,000 random particles were chosen and processed through CryoSparc v2 ab-initio reconstruction six times using the raw particles and six times using the particles denoised by the v0.2.1 L2 model. Comparisons between the full 3D map and each set of six ab-initio models were made in UCSF Chimera[30].

Results

1. Denoising with Topaz improves micrograph interpretability and SNR

We develop a general cryoEM micrograph denoising model by training a neural network using the Noise2Noise framework on dozens of representative datasets of commonly used imaging conditions (Figure 1, Methods). By learning the denoising model directly from data, we avoid making specific assumptions about the noise-generating process leading to superior denoising performance.

Denoising with Topaz improves micrograph interpretability by eye on several datasets and improves SNR measurements in a quantitative analysis. Our model correctly smoothes background areas while preserving structural features better than conventional methods (i.e. affine or low-pass filtering) (Figure 1, Supplemental Figures 1-4). Given this known smoothing behavior of micrograph areas containing primarily noise, we find that denoising allows for identification of structured background features from noise. Figure 1 shows two micrographs where the background areas between particles are flattened after denoising, while

bioRxiv preprint doi: https://doi.org/10.1101/838920; this version posted November 12, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Supplemental Figure 5 shows microtubules with known small proteins in background areas properly retained after denoising. Our denoising model has the combined advantage of reducing ambiguity as to whether the background of a micrograph is generally free from contamination, allowing researchers to identify small and/or low density particle views, for example as applied to micrographs from Mao et al.[31] (Supplemental Figure 6,7). In these types of scenarios, visual assessment of denoised micrographs compared to raw micrographs increases protein density confidence, increases confidence of background content, and reduces the eye strain for researchers.

We quantitatively assess denoising performance by measuring the SNR of raw micrographs, micrographs denoised with our model, and micrographs denoised with conventional methods. To this end, we manually annotated paired signal and background regions on micrographs from 10 different datasets (Supplemental Figure 8). We then calculated the average SNR (in dB) for each method using these regions[32]. We present a comparison of four different denoising model architectures (affine, FCNN, U-net (small), and U-net) trained with L1 and L2 losses on either the small or large datasets (Supplemental Table 2). We find only minor differences between L1 and L2 models, with L1 loss being slightly favored overall. Furthermore, we find that the training dataset is important. Intriguingly, the affine, FCNN, and U-net (small) models all perform better than the full U-net model when trained on the small dataset and perform better than their equivalents trained on the large dataset. The best performing model overall, however, is the full U-net model trained on the large dataset. Furthermore, we find that this model outperforms conventional low-pass filtering denoising on all datasets except for one, where they perform equivalently (EMPIAR-10005).

A summary comparison is presented in Table 1, where we report SNR results on each dataset for the best overall performing low-pass filter (16x binning) with the L2 U-net trained on the large dataset and the L1 affine model trained on the small dataset. Our pretrained U-net improves SNR by >2 dB on average over low-pass filtering and improves SNR by roughly 20 dB (100 fold) over the raw micrographs. Furthermore, the model generalizes well across different imaging parameters, improving SNR on micrographs collected on K2 and Falcon III cameras as well as micrographs collected in super-resolution and counting modes.

2. Denoising enables shorter exposure imaging

To simulate shorter exposure times at the microscope, we truncated frames of several datasets used during frame alignment and summed to the first 10%, 25%, 50%, and 75% of the frames. These datasets were collected with a total dose of between 40 and 69 e-/Å^2. We denoised each short exposure with our general U-net model and compare both visually and quantitatively to low-pass filtering and to the raw micrographs without denoising.

Figure 2 shows denoised and lowpass filtered example micrographs of each subset along with the raw micrographs. Visual analysis and our SNR analysis suggests that between 10% and 25% of the exposure time is comparable to the full, raw micrographs (Figure 2, Supplemental Figure 9 for FFTs, Supplemental Figures 10-12). This corresponds to between 4.0 and 16.7 e-/Å^2. 3D reconstructions of frame titrations of identical apoferritin particles from 19jan04d suggests that a total dose of about 16.7 e-/Å^2 is required for accurate CTF estimation (Supplemental Figure 13). Furthermore, roughly double the electron dose is required for low-pass filtering to match the SNR of our neural denoised micrographs. This could allow a

bioRxiv preprint doi: https://doi.org/10.1101/838920; this version posted November 12, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

factor of two or more savings in exposure time. Such a significant reduction in exposure time substantially increases the efficiency of cryoEM collection sessions, allowing for microscopes to operate at higher throughput.

3. Denoising enables picking of difficult particle projections

We denoised micrographs of particles with particularly difficult-to-identify projections, clustered protocadherin (EMPIAR-10234) to test whether denoising enables these views to be picked more completely than without denoising. Figure 3 shows a representative micrograph before and after denoising (Supplemental Figure 9 for FFTs). Before denoising, many particle top-views were indistinguishable by eye from noise (Figure 3a, inset). After denoising, top-views in particular became readily identifiable (Figure 3a, inset and circled in green).

We manually picked 1,023 particles while attempting to balance the percentage of side, oblique, and top-views of the particle in our picks. Using these picks, we trained a Topaz picking model as described in the Methods. The resulting model was used to infer a total of 16,049 particles after CryoSparc 2D classification and 3D heterogeneous refinement. Using only the raw micrographs for initial manual picking, the data owner picked 1,540 particles to train a Topaz model as described in Brasch et al.[28] that inferred 10,010 particles after CryoSparc 2D classification and 3D refinement. Using denoised micrographs allowed us to pick over 60% more real particles and substantially increase the percentage of top- and oblique-views.

Figure 3b shows the resulting 3D structured using raw particles. The tertiary features of the EC-domains are better-resolved in the particles picked using the U-net denoising model, and the angular particle coverage is more filled in and complete (Figure 3c, Supplemental Figure 14). The side-views allow the data owner to more accurately interpret how the symmetry of the cadherin dimers is broken upon interaction compared to the previous sub-tomogram average model in Brasch et al.[28] Interestingly, CryoSparc ab-initio reconstruction using a minimal set of denoised particles is less reliable than using the same set of raw particles (Supplemental Figure 15). Four or five of the six ab-initio reconstructions using the raw particles resulted in the correct overall structure, while only one of the six ab-initio reconstructions using the denoised particles resulted in the correct overall structure

Conclusion

CryoEM has long been hampered by the ability for researchers to confidently identify protein particles in all represented orientations from behind sheets of noise. Several bottlenecks in the general cryoEM workflow may preclude protein structure determination due to low SNR, such as differentiating protein from noise during picking, picking homogeneous subsets of particles, picking sufficient numbers of particles in all represented orientations, and obtaining a sufficient number of particles for 3D refinement. The initial stages of *de novo* protein structure determination are particularly affected by these issues. To ameliorate these potentially critical issues, we present Topaz-Denoise, a Noise2Noise convolutional neural network for learning and removing significant noise information from cryoEM images. By employing a network trained on dozens of datasets to account for varying sample, microscope, and collection parameters, we show and provide robust general denoising models. We show empirically that our U-net denoising models result in higher SNR relative to affine models and low-pass filters.

bioRxiv preprint doi: https://doi.org/10.1101/838920; this version posted November 12, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Topaz-Denoise enables visual identification of low SNR particle views, as exemplified by the clustered protocadherin dataset where denoising allowed for a more representative and complete 3D reconstruction. Moreover, due to the considerable increase in SNR of denoised single particle micrographs, exposure time may be reduced without sacrificing the ability to pick particles reliably, thus enabling an increase in collection efficiency. We expect micrograph denoising in Topaz to become a standard component of the micrograph analysis pipeline due to its performance and modularity.

Code availability statement

Source code for Topaz Denoise is publicly available as part of Topaz (v0.2.0 and above) on GitHub at https://github.com/tbepler/topaz. Topaz is installable through Anaconda, Pip, Docker, Singularity, SBGrid, and source. Topaz is licensed under the GNU General Public License v3.0.

Data availability statement

The general models used in this manuscript are included as options in Topaz Denoise. Over 100 NYSBC dataset frames used for some of the models have been deposited to EMPIAR-XXXXX. The clustered protocadherin model from manual picking on denoised micrographs has been deposited to EMD-YYYY.

Acknowledgements

The authors wish to thank Simons Electron Microscopy Center (SEMC) OPs for many of the test datasets used in training and the authors of EMPIAR entries XXXXX, YYYYY, … for additional training datasets. The authors thank Dr. Julia Brasch for sharing her processing experience with EMPIAR-10234. The authors wish to thank Sargis Dallakyan for integrating Topaz Denoise into Appion. The authors wish to thank Anchi Cheng, Mykhailo Kopylov, Bridget Carragher, and Clinton Potter for helpful discussions.

T.B. and B.B. were supported by NIH grant R01-GM081871. A.J.N. was supported by a grant from the NIH National Institute of General Medical Sciences (NIGMS) (F32GM128303). The cryoEM work was performed at SEMC and National Resource for Automated Molecular Microscopy located at NYSBC, supported by grants from the Simons Foundation (SF349247), NYSTAR, and the NIH NIGMS (GM103310) with additional support from the Agouron Institute (F00316) and NIH (OD019994).

Author contributions

T.B., A.J.N., and B.B. conceived of this project. T.B. developed and implemented the models. T.B. and A.J.N. processed and analyzed the data and model results. T.B. and A.J.N. processed

bioRxiv preprint doi: https://doi.org/10.1101/838920; this version posted November 12, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

and analyzed single particle results. T.B., A.J.N., and B.B. designed the experiments and wrote the manuscript.

bioRxiv preprint doi: https://doi.org/10.1101/838920; this version posted November 12, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

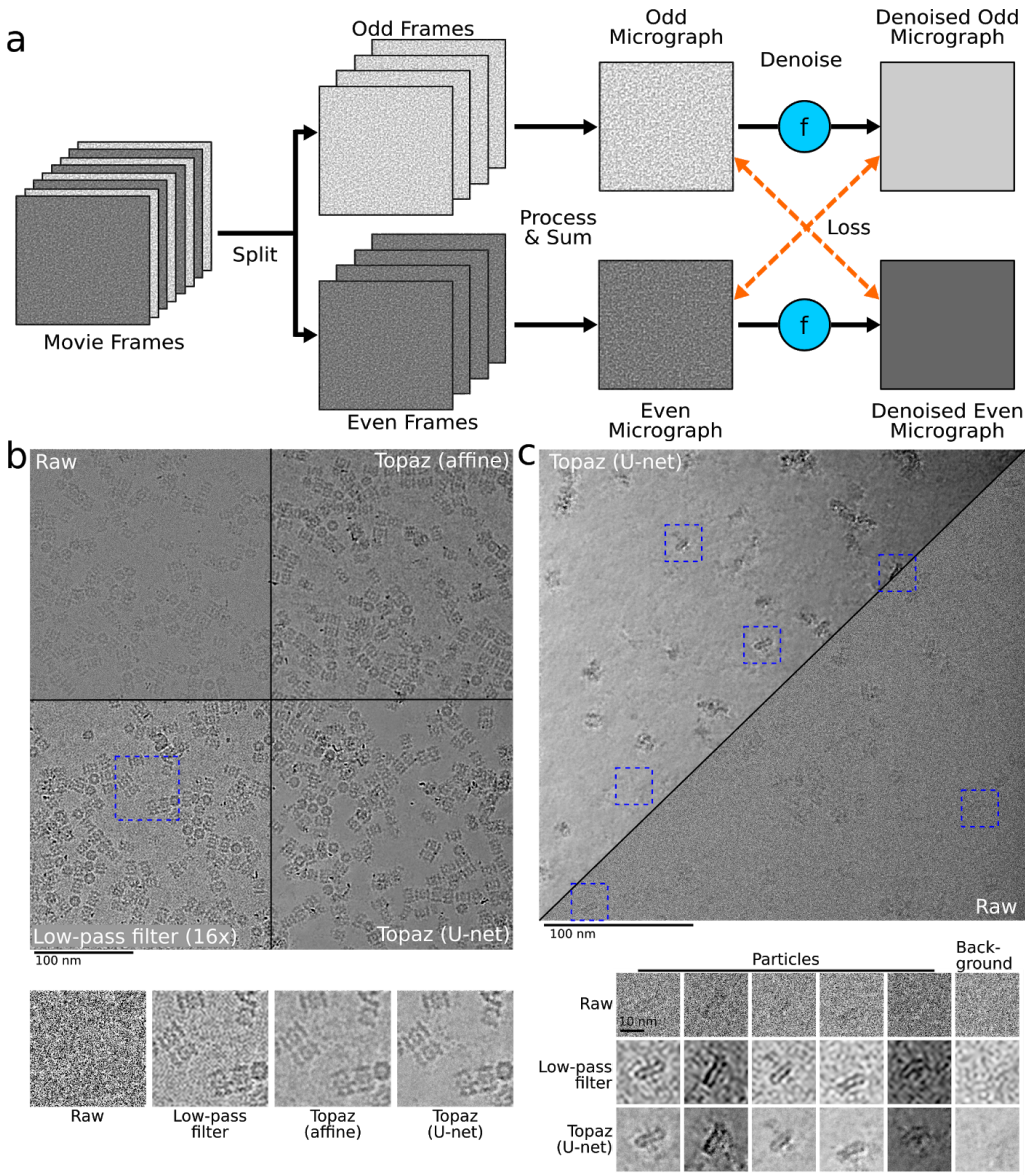


Figure 1 | Illustration of the training framework and comparison of denoising methods on two example micrographs. **(a)** The Noise2Noise method requires paired noisy observations of the same underlying signal. We generate these pairs from movie frames collected in the normal cryoEM process, because each movie frame is an independent sample of the same signal. These are first split into even/odd movie frames. Then, each is processed and summed independently following standard micrograph processing protocols. The resulting even and odd micrographs are denoised with the denoising model (denoted here as *f*). Finally, to calculate the loss, the odd denoised micrograph is compared with the raw even micrograph and vice versa. **(b)** Micrograph from EMPIAR-10025 split into four quadrants showing the raw micrographs, lowpass filtered micrograph by a binning factor of 16, and results of denoising with our affine and U-net models. Particles become clearly visible in the lowpass filtered and denoised micrographs, but the U-net denoising shows strong additional smoothing of background noise. A detail view of the micrograph is highlighted in blue and helps to illustrate the improved background smoothing provided by our U-net denoising model. **(c)** Micrograph from EMPIAR-10261 split into the U-net

bioRxiv preprint doi: https://doi.org/10.1101/838920; this version posted November 12, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

denoised and raw micrographs along the diagonal. Detail views of five particles and one background patch are boxed in blue. The Topaz U-net reveals particles and reduces background noise.

	EMPIAR datasets			NYSBC K2 datasets				NYSBC Falcon3 datasets			
Method	10261	10005	10025	Protocadherin	18sep08d	19jan04d	19may10e	18aug17l	18sep06d	18sep19l	Overall
Affine (Topaz)	5.49	1.29	0.72	4.83	4.51	**8.87**	12.02	10.65	6.90	9.15	6.44
U-net (Topaz)	**7.17**	**1.72**	**1.07**	**5.94**	**6.06**	8.43	**13.07**	**15.17**	**7.37**	**13.24**	**7.92**
Low-pass	5.19	-0.12	-0.40	4.22	3.53	6.87	9.99	9.04	6.95	8.71	5.40
Raw	-17.14	-20.13	-24.15	-14.47	-15.40	-11.73	-5.44	-6.33	-3.64	-5.63	-12.41

Table 1 | Comparison of denoising methods based on estimated SNR (in dB, larger is better). SNR was estimated from 20 paired signal and background regions selected for each dataset. In each column, the best performing model is highlighted. We report denoising results on aligned micrographs for the NYSBC K2 and Falcon3 datasets. All datasets were collected in electron counting modes, except for 18sep06d which was collected using Falcon III integrating mode. Our U-net denoising model performs best overall and is best on all except for the 19jan04d dataset where our affine denoising model slightly outperforms it. We report low-pass filtering by a binning factor of 16 on all datasets, which we found to give better SNR overall compared to Gaussian low-pass filtering.

bioRxiv preprint doi: https://doi.org/10.1101/838920; this version posted November 12, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

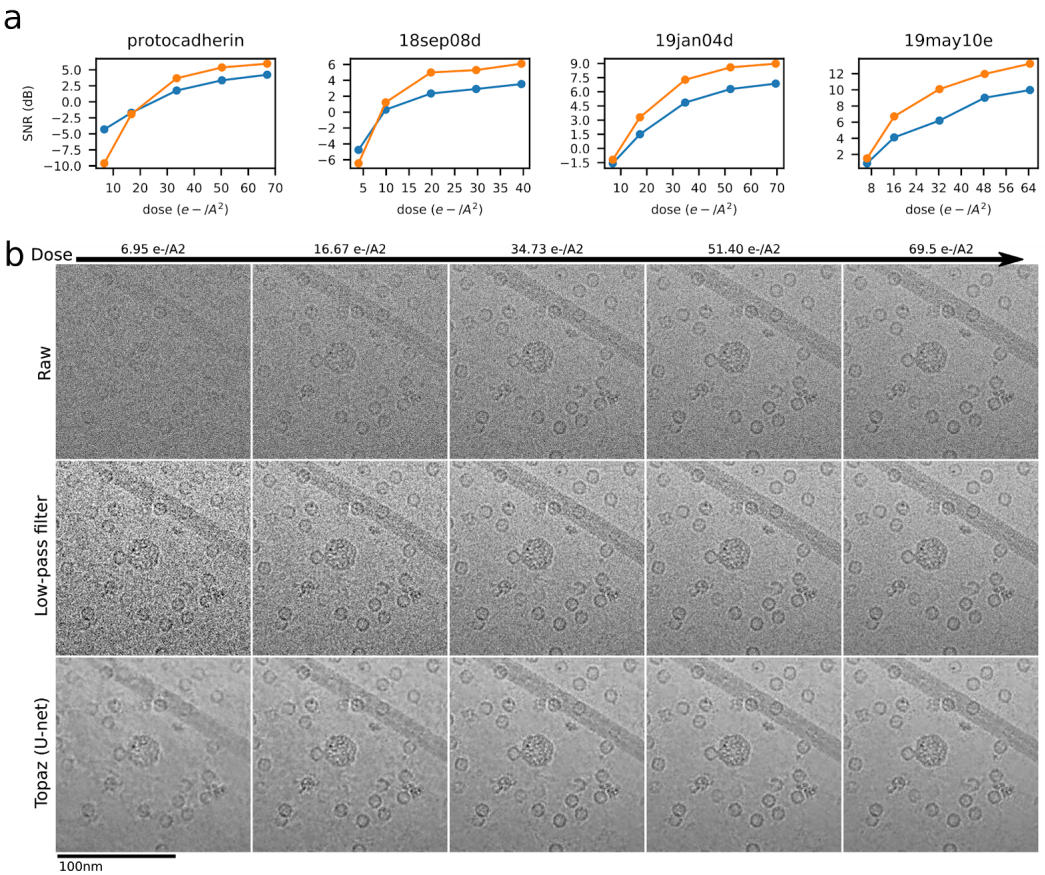


Figure 2 | Denoising in Topaz enhances SNR of short exposure micrographs. **(a)** SNR (dB) as a function of electron dose in low-pass filtered micrographs by a binning factor of 16 (blue), and U-net denoised micrographs (orange) in the four NYSBC K2 datasets. Our U-net denoising model enhances the SNR of micrographs across almost all dosages in all four datasets. SNR can be enhanced by a factor of 1.5x or more at 20 e-/A^2. **(b)** Example section of a micrograph from the 19jan04d dataset of apoferritin, β-galactosidase, and TMV (full micrograph in Supplemental Figure 3,4) showing the raw micrograph, low-pass filtered micrograph, and U-net denoised micrograph over increasing dose. Particles are clearly visible at the lowest dose in the denoised micrograph and background noise is substantially reduced by Topaz denoising.

bioRxiv preprint doi: https://doi.org/10.1101/838920; this version posted November 12, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

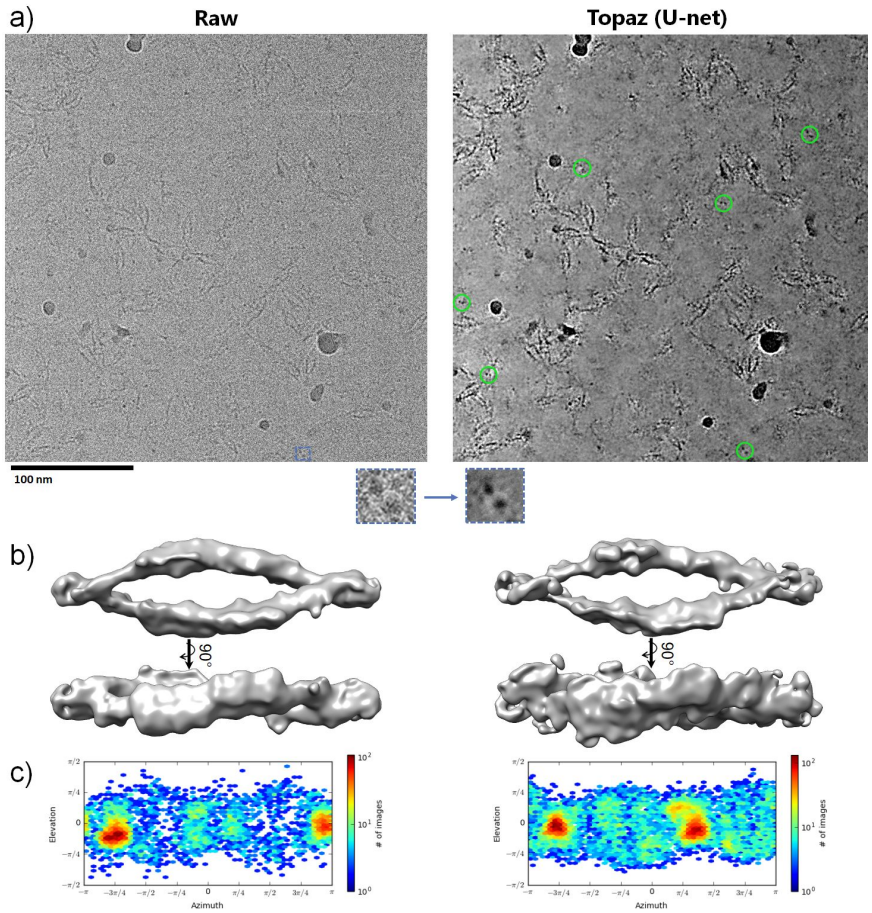


Figure 3 | Denoising in Topaz improves interpretability and picking of difficult particle projections. **(a)** A raw micrograph (left) and Topaz denoised micrograph (right) of the clustered protocadherin dataset (EMPIAR-10234) with a top-view boxed out (insets). Denoising allows for top-views to be clearly identified (green circles, right) and subsequently used to increase the confidence of particle picking. **(b)** Topaz picking training on raw micrographs using 1,540 manually picked particles from the raw micrographs resulted in the reconstruction on the left. Topaz picking training on the raw micrographs using 1,023 manually picked particles from the denoised micrographs resulted in the reconstruction on the right. Manually picking on denoised micrographs resulted in 60% more particles in the 3D reconstruction. **(c)** Particle distributions for each reconstruction. 3DFSC plots are shown in Supplemental Figure 14.

bioRxiv preprint doi: https://doi.org/10.1101/838920; this version posted November 12, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

References

1. Baxter, W. T., Grassucci, R. A., Gao, H. & Frank, J. Determination of signal-to-noise ratios and spectral SNRs in cryo-EM low-dose imaging of molecules. *J. Struct. Biol.* **166**, 126–132 (2009).
2. Sindelar, C. V. & Grigorieff, N. An adaptation of the Wiener filter suitable for analyzing images of isolated single particles. *J. Struct. Biol.* **176**, 60–74 (2011).
3. Penczek, P. A. Chapter Two - Image Restoration in Cryo-Electron Microscopy. in *Methods in Enzymology* (ed. Jensen, G. J.) vol. 482 35–72 (Academic Press, 2010).
4. Merk, A. *et al.* Breaking Cryo-EM Resolution Barriers to Facilitate Drug Discovery. *Cell* **165**, 1698–1707 (2016).
5. Herzik, M. A., Wu, M. & Lander, G. C. High-resolution structure determination of sub-100 kDa complexes using conventional cryo-EM. *Nat. Commun.* **10**, 1–9 (2019).
6. Milanfar, P. A Tour of Modern Image Filtering: New Insights and Methods, Both Practical and Theoretical. *IEEE Signal Process. Mag.* **30**, 106–128 (2013).
7. Buades, A., Coll, B. & Morel, J.-. A non-local algorithm for image denoising. in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)* vol. 2 60–65 vol. 2 (2005).
8. Dabov, K., Foi, A., Katkovnik, V. & Egiazarian, K. Image Denoising by Sparse 3-D Transform-Domain Collaborative Filtering. *IEEE Trans. Image Process.* **16**, 2080–2095 (2007).
9. Xie, J., Xu, L. & Chen, E. Image Denoising and Inpainting with Deep Neural Networks. in *Advances in Neural Information Processing Systems 25* (eds. Pereira, F., Burges, C. J. C., Bottou, L. & Weinberger, K. Q.) 341–349 (Curran Associates, Inc., 2012).

bioRxiv preprint doi: https://doi.org/10.1101/838920; this version posted November 12, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

10. Jain, V. & Seung, S. Natural Image Denoising with Convolutional Networks. in *Advances in Neural Information Processing Systems 21* (eds. Koller, D., Schuurmans, D., Bengio, Y. & Bottou, L.) 769–776 (Curran Associates, Inc., 2009).

11. Lehtinen, J. *et al.* Noise2Noise: Learning Image Restoration without Clean Data. *ArXiv180304189 Cs Stat* (2018).

12. Zhussip, M., Soltanayev, S. & Chun, S. Y. Extending Stein's unbiased risk estimator to train deep denoisers with correlated pairs of noisy images. *ArXiv190202452 Cs* (2019).

13. Batson, J. & Royer, L. Noise2Self: Blind Denoising by Self-Supervision. *ArXiv190111365 Cs Stat* (2019).

14. Krull, A., Buchholz, T.-O. & Jug, F. Noise2Void - Learning Denoising from Single Noisy Images. *ArXiv181110980 Cs* (2018).

15. Buchholz, T., Jordan, M., Pigino, G. & Jug, F. Cryo-CARE: Content-Aware Image Restoration for Cryo-Transmission Electron Microscopy Data. in *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)* 502–506 (2019). doi:10.1109/ISBI.2019.8759519.

16. Buchholz, T.-O. *et al.* Chapter 13 - Content-aware image restoration for electron microscopy. in *Methods in Cell Biology* (eds. Müller-Reichert, T. & Pigino, G.) vol. 152 277–289 (Academic Press, 2019).

17. Tegunov, D. & Cramer, P. Real-time cryo-electron microscopy data preprocessing with Warp. *Nat. Methods* **16**, 1146–1152 (2019).

18. Bepler, T. *et al.* Positive-unlabeled convolutional neural networks for particle picking in cryo-electron micrographs. *Nat. Methods* 1–8 (2019) doi:10.1038/s41592-019-0575-8.

19. Morin, A. *et al.* Collaboration gets the most out of software. *eLife* **2**, e01456 (2013).

20. Lander, G. C. *et al.* Appion: An integrated, database-driven pipeline to facilitate EM

bioRxiv preprint doi: https://doi.org/10.1101/838920; this version posted November 12, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

image processing. *J. Struct. Biol.* **166**, 95–102 (2009).

21. Iudin, A., Korir, P. K., Salavert-Torres, J., Kleywegt, G. J. & Patwardhan, A. EMPIAR: a public archive for raw electron microscopy image data. *Nat. Methods* **13**, 387–388 (2016).

22. Zheng, S. Q. *et al.* MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).

23. Ronneberger, O., Fischer, P. & Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015* (eds. Navab, N., Hornegger, J., Wells, W. M. & Frangi, A. F.) 234–241 (Springer International Publishing, 2015).

24. Paszke, A. *et al.* Automatic differentiation in PyTorch. (2017).

25. Kremer, J. R., Mastronarde, D. N. & McIntosh, J. R. Computer Visualization of Three-Dimensional Image Data Using IMOD. *J. Struct. Biol.* **116**, 71–76 (1996).

26. Rohou, A. & Grigorieff, N. CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* **192**, 216–221 (2015).

27. Tan, Y. Z. *et al.* Addressing preferred specimen orientation in single-particle cryo-EM through tilting. *Nat. Methods* **14**, 793–796 (2017).

28. Brasch, J. *et al.* Visualization of clustered protocadherin neuronal self-recognition complexes. *Nature* **569**, 280–283 (2019).

29. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017).

30. Pettersen, E. F. *et al.* UCSF Chimera--a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).

31. Mao, Y. *et al.* Molecular architecture of the uncleaved HIV-1 envelope glycoprotein trimer. *Proc. Natl. Acad. Sci.* **110**, 12438–12443 (2013).

bioRxiv preprint doi: https://doi.org/10.1101/838920; this version posted November 12, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

32. Haider, S. A. *et al.* Fluorescence microscopy image noise reduction using a stochastically-connected random field model. *Sci. Rep.* **6**, 20640 (2016).

bioRxiv preprint doi: https://doi.org/10.1101/838920; this version posted November 12, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Supplement

Supplemental Table 1: Training datasets

Dataset Name	Number of Micrographs	Description
NYSBC zero defocus	671	NYSBC ladder dataset collected at zero defocus
NYSBC ice images	900	NYSBC micrographs taken of vitreous ice
EMPIAR-10025	196	T20S proteasome
EMPIAR-10005	499	TRPV1 dataset taken on a K2 direct electron detector
EMPIAR-10210	170	mouse MDA5-dsRNA filaments
EMPIAR-10243	142	heparin-induced 2N4R tau filaments
EMPIAR-10244	642	RNA polymerase II transcribing a nucleosome
EMPIAR-10248	971	Apoferritin by CRYOARM300 with cold-FEG
EMPIAR-10249	596	Horse liver alcohol dehydrogenase movies obtained using Talos Arctica operating at 200 kV equipped with a K2
EMPIAR-10250	181	Human methemoglobin movies obtained using Talos Arctica operating at 200 kV equipped with a K2
EMPIAR-10252	153	Catalytic subunit of protein kinase A bound to ATP, manganese, and IP20 movies obtained using Talos Arctica operating at 200 kV equipped with a K2
EMPIAR-10257	295	NDH the complex I-like molecule of photosynthesis
EMPIAR-10258	199	LRRC8A-DCPIB in MSP1E3D1 nanodiscs
EMPIAR-10259	198	apo-LRRC8A in MSP2N2 nanodiscs
EMPIAR-10261	1461	ProTx2-bound Nav1.7 VSD2-NavAb chimeric channel
EMPIAR-10031	512	MAVS CARD C1 filaments, Falcon2 direct electron detector
EMPIAR-10061	397	beta-galactosidase in complex with a cell-permeant inhibitor
EMPIAR-10028	600	Plasmodium falciparum 80S ribosome bound to the anti-protozoan drug emetine
Small	1929	Contains micrographs from datasets: EMPIAR-10005, -10025, -10061, -10244, -10249, -10250, -10252, -10257, -10258, and -10261
Large	3439	Contains micrographs from all individual datasets

Supplemental Table 1 | List of datasets included in model training. The individual datasets with number of micrographs from each and brief descriptions are provided in the first block. The second block describes the camera/imaging mode specific datasets. These are composed of all micrographs from subsets of the individual datasets. The last block describes the two general datasets. The “Small” dataset is composed of micrographs from a subset of the individual datasets that we found to give best performing models by eye. The “Large” dataset contains micrographs from all individual datasets. For the “Small” and “Large” datasets, individual datasets with more than 200 micrographs were subsampled to only include 200 images. This serves to approximately balance the contributions of each contained dataset.

bioRxiv preprint doi: https://doi.org/10.1101/838920; this version posted November 12, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

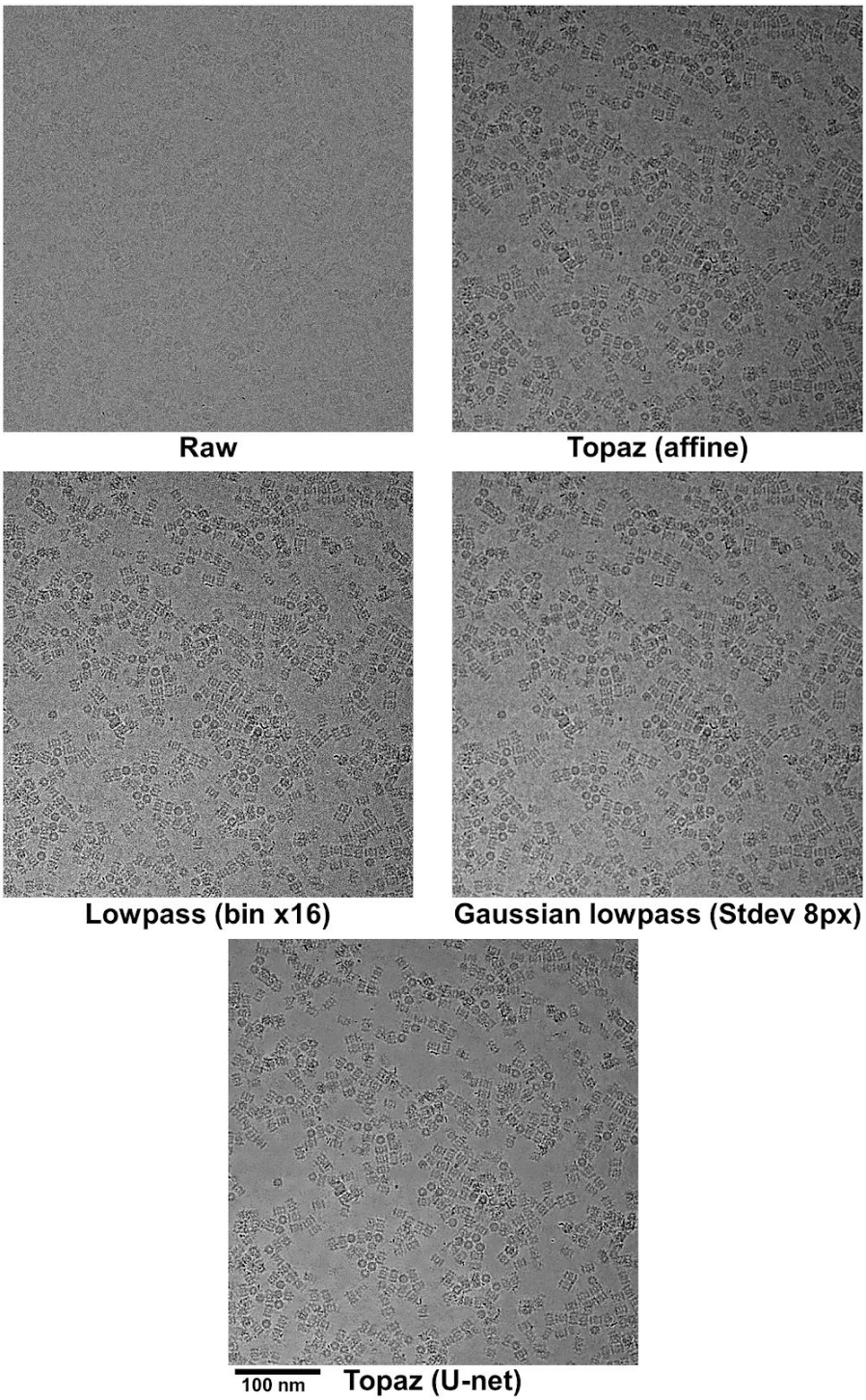
Supplemental Table 2: SNR comparison for model architectures, loss functions, and training datasets

Training Dataset	Model architecture	Objective	EMPIAR datasets			NYSBC K2 datasets				NYSBC Falcon3 datasets			Overall
			10261	10005	10025	Protocadherin	18sep08d	19jan04d	19may10e	18aug17l	18sep06d	18sep19l	
Small	Affine	L1	5.49	1.29	0.72	4.83	4.51	8.87	12.02	10.65	6.90	9.15	6.44
		L2	5.28	1.04	0.40	4.64	4.33	8.64	11.88	10.43	6.89	9.02	6.25
	FCNN	L1	5.45	1.25	-0.05	5.37	5.85	8.78	12.77	12.02	7.90	10.89	7.02
		L2	5.17	0.94	-0.59	4.82	5.67	8.62	12.31	11.30	7.76	10.29	6.63
	U-Net (small)	L1	5.99	0.65	0.12	5.89	**7.11**	9.71	14.17	12.05	**8.25**	10.93	7.49
		L2	5.32	0.49	0.27	5.55	6.85	**9.74**	**14.26**	11.23	7.49	10.36	7.16
	U-Net	L1	5.24	0.79	2.51	5.64	5.38	9.38	13.53	9.91	6.49	9.59	6.85
		L2	5.72	0.86	**3.01**	5.41	5.22	8.64	13.06	9.36	6.61	9.16	6.70
Large	Affine	L1	4.30	-0.24	-1.09	3.76	3.36	7.23	11.34	9.32	6.56	8.45	5.30
		L2	4.26	-0.27	-1.15	3.74	3.33	7.20	11.32	9.32	6.57	8.46	5.28
	FCNN	L1	4.69	-0.74	-1.57	3.95	5.25	7.36	11.93	10.26	8.04	9.72	5.89
		L2	3.78	-0.47	-1.57	3.76	4.79	6.81	11.21	10.15	7.34	9.62	5.54
	U-Net (small)	L1	6.47	0.51	-1.00	5.18	4.82	7.04	12.29	13.62	8.04	10.83	6.78
		L2	6.00	0.84	-0.46	5.19	5.20	7.90	12.34	12.36	7.20	10.38	6.69
	U-Net	L1	6.95	1.88	0.83	**6.33**	6.14	9.35	13.87	13.11	7.35	12.76	7.86
		L2	**7.17**	1.72	1.07	5.94	6.06	8.43	13.07	**15.17**	7.37	**13.24**	**7.92**
Lowpass		4	-5.28	-11.17	-11.92	-5.28	-6.08	-2.97	3.44	-1.04	-1.02	-0.29	-4.16
		8	0.30	-5.52	-5.89	-0.23	-0.89	2.03	7.84	4.30	2.87	4.57	0.94
		16	5.19	-0.12	-0.40	4.22	3.53	6.87	9.99	9.04	6.95	8.71	5.40
		32	3.92	**1.89**	0.08	1.65	1.63	8.25	10.13	8.58	1.90	7.92	4.59
		64	2.34	-2.22	-0.83	-3.25	5.60	5.18	11.12	10.23	6.58	-0.02	3.47
Raw			-17.14	-20.13	-24.15	-14.47	-15.40	-11.73	-5.44	-6.33	-3.64	-5.63	-12.41

Supplemental Table 2 | Comparison of denoising methods based on estimated SNR (in dB, larger is better). SNR was estimated from 20 paired signal and background regions selected for each dataset. In each column, the best performing model is highlighted. We report denoising results on aligned and dose weighted micrographs for the NYSBC K2 and Falcon3 datasets.All datasets were collected in electron counting modes, except for 18sep06d which was collected using Falcon III integrating mode. The U-net denoising model trained on the “Large” dataset with L2 loss performs best on average. For the low-pass filter baselines, the amount of filtering is reported in the “Objective” column. The SNR of the raw micrographs is reported in the last row.

bioRxiv preprint doi: https://doi.org/10.1101/838920; this version posted November 12, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

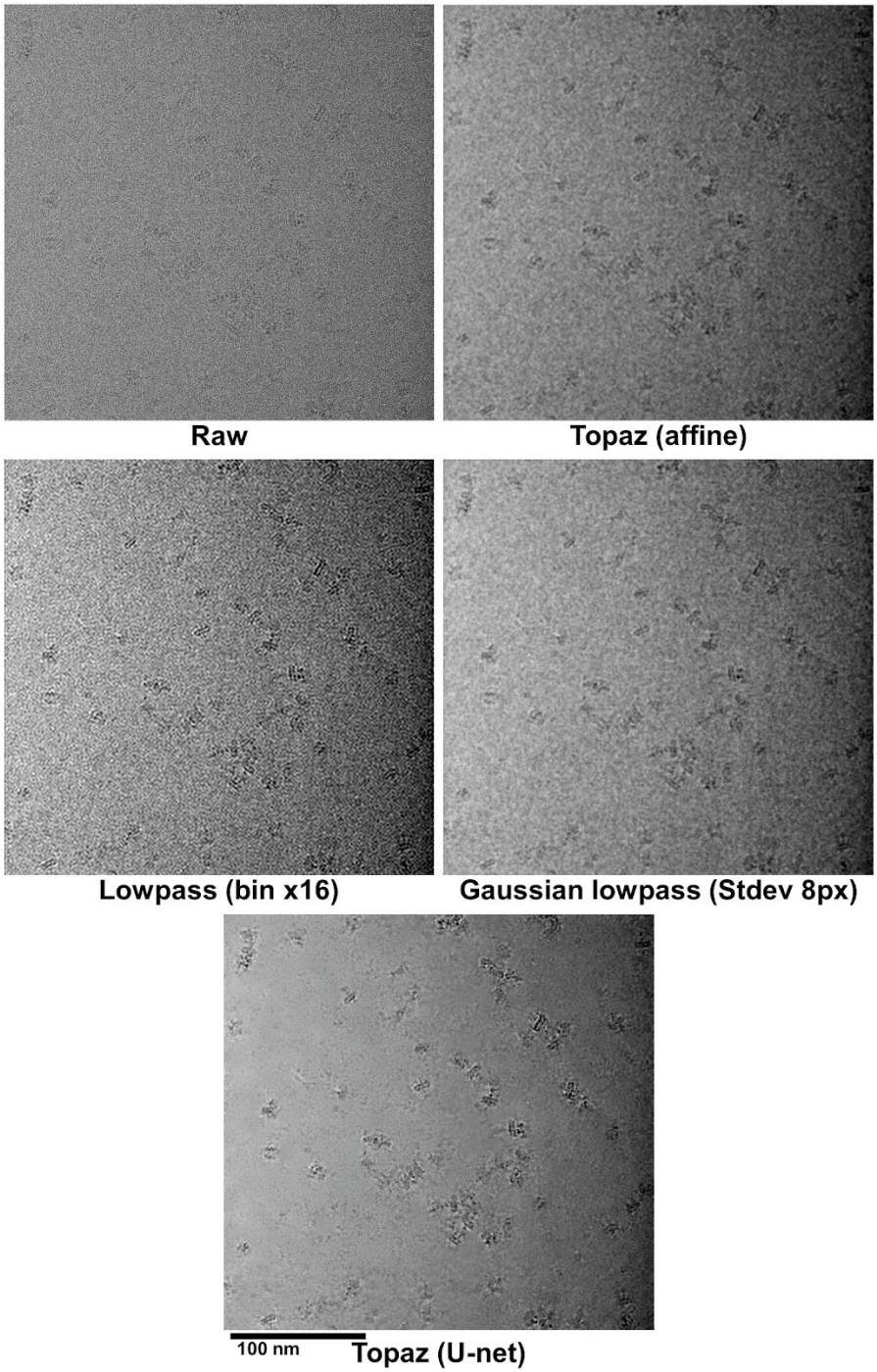
Supplemental Figure 1: Comparison of denoising methods



Supplemental Figure 1 | Figure 2a micrograph (pixel size: 0.6575 Å) processed in four different ways: Topaz affine denoising model, low-pass binning by Fourier cropping by a factor of 16 then padding, Gaussian low-pass filtering with a standard deviation of 8 pixels, and our Topaz U-net denoising model.

bioRxiv preprint doi: https://doi.org/10.1101/838920; this version posted November 12, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

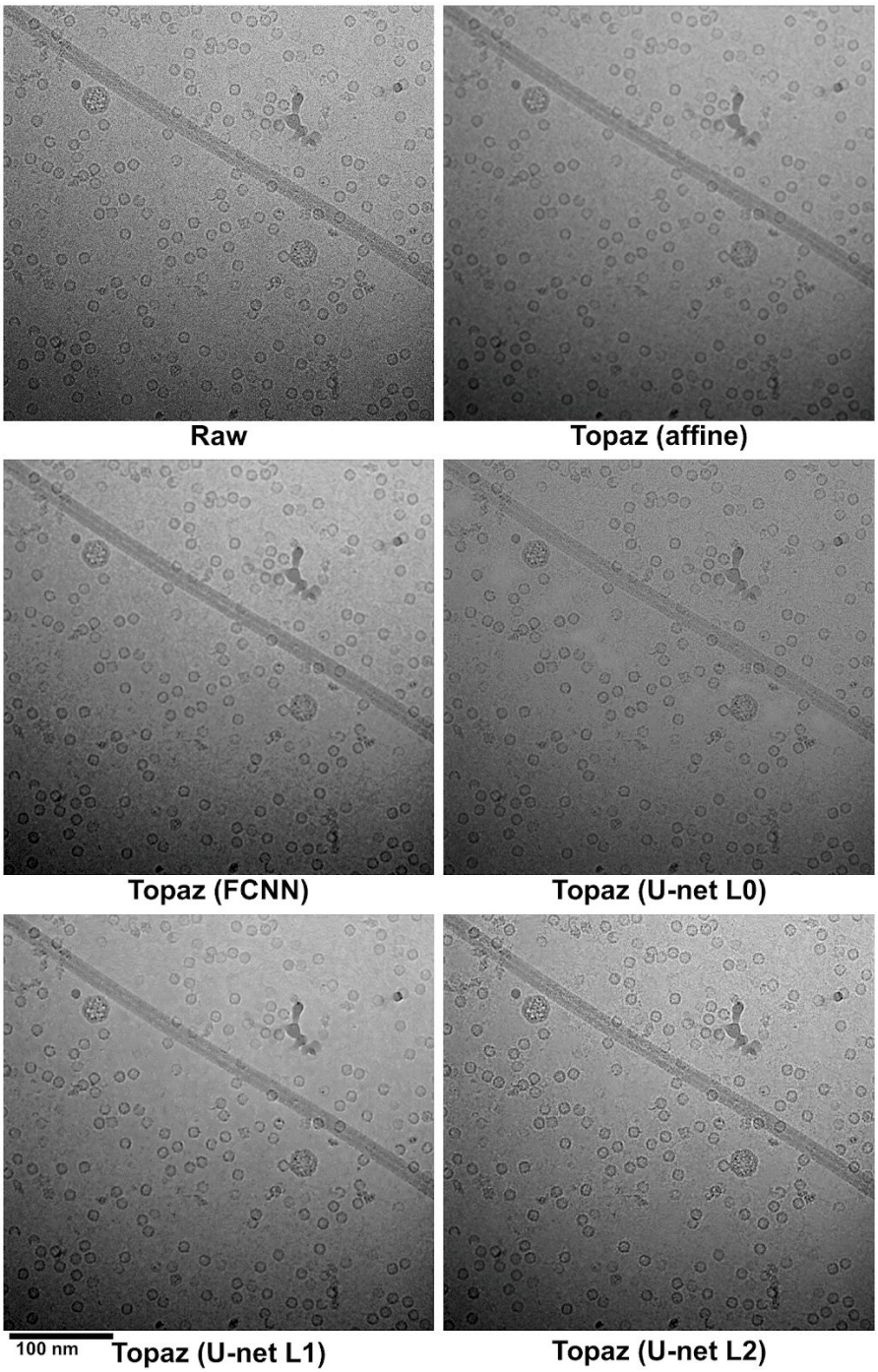
Supplemental Figure 2: Comparison of denoising methods on EMPIAR-10261



Supplemental Figure 2 | Figure 2b micrograph (pixel size: 0.849 Å) processed in four different ways: Topaz affine denoising model, low-pass binning by Fourier cropping by a factor of 16 then padding, Gaussian low-pass filtering with a standard deviation of 8 pixels, and our Topaz U-net denoising model.

bioRxiv preprint doi: https://doi.org/10.1101/838920; this version posted November 12, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

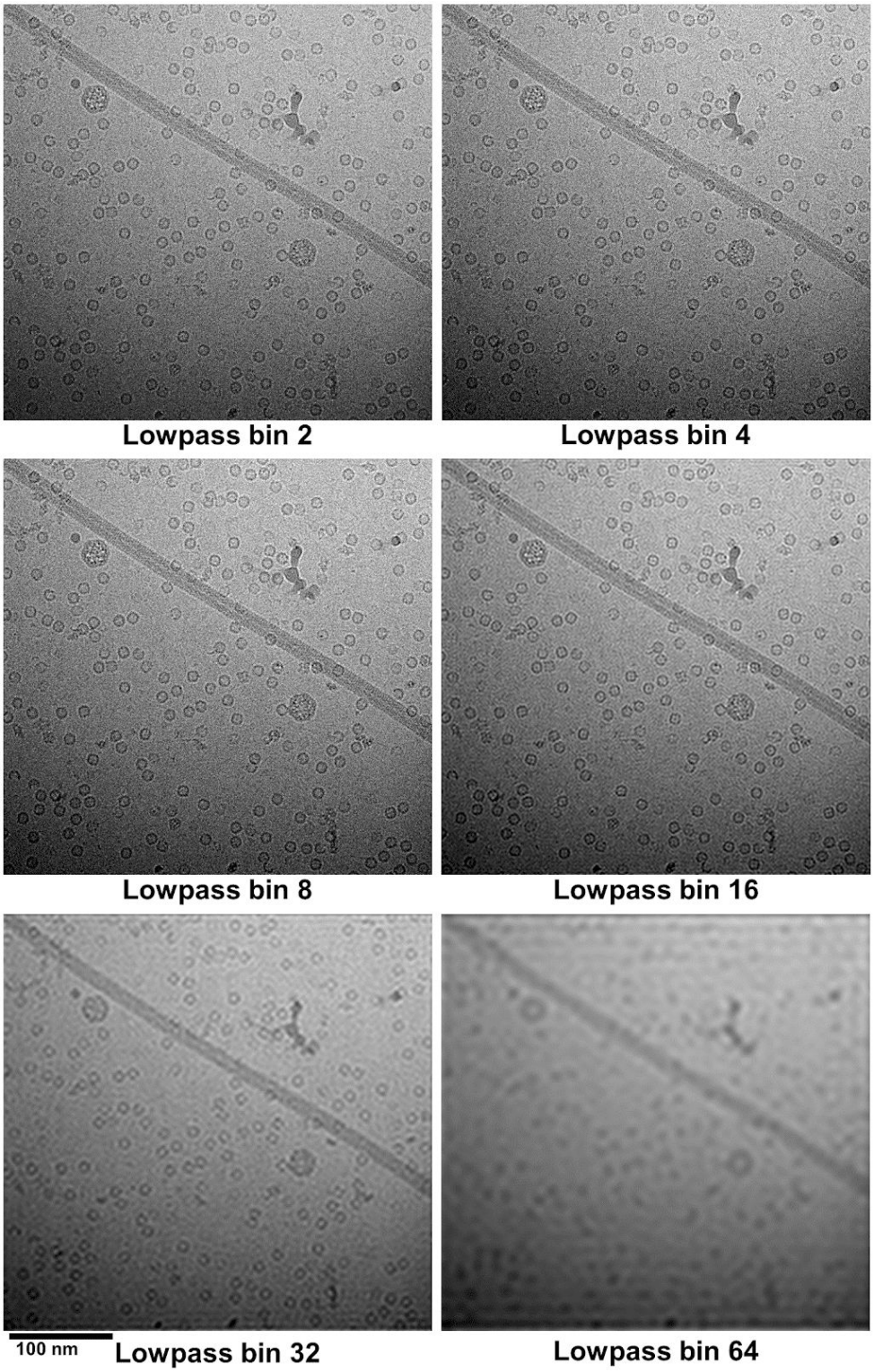
Supplemental Figure 3: Comparison of Topaz neural denoising models on 19jan04d



Supplemental Figure 3 | Comparison between denoising models on a micrograph (pixel size 1.10 Å) of apoferritin, β-galactosidase, and TMV. The raw image denoised with the affine model, FCNN model, U-net with mode-seeking L0 loss, U-net with median-seeking L1 loss, and U-net with mean-seeking L2 loss are shown.

bioRxiv preprint doi: https://doi.org/10.1101/838920; this version posted November 12, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

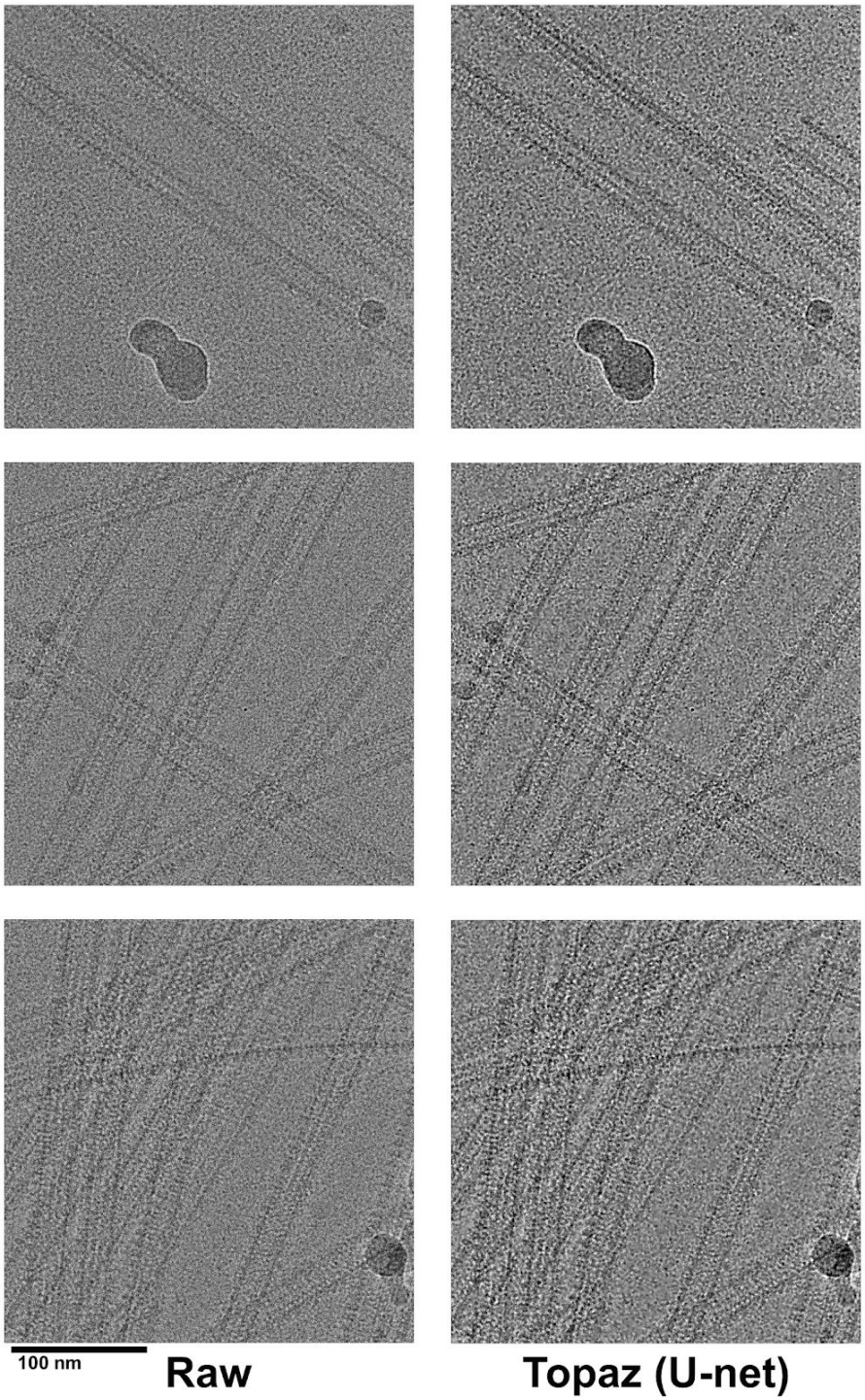
Supplemental Figure 4: Comparison of lowpass binning on 19jan04d



Supplemental Figure 4 | Comparison between low-pass binning by Fourier cropping. The raw micrograph in Supplemental Figure 3 is low-pass filtered by factors of 2, 4, 8, 16, 32, and 64.

bioRxiv preprint doi: https://doi.org/10.1101/838920; this version posted November 12, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

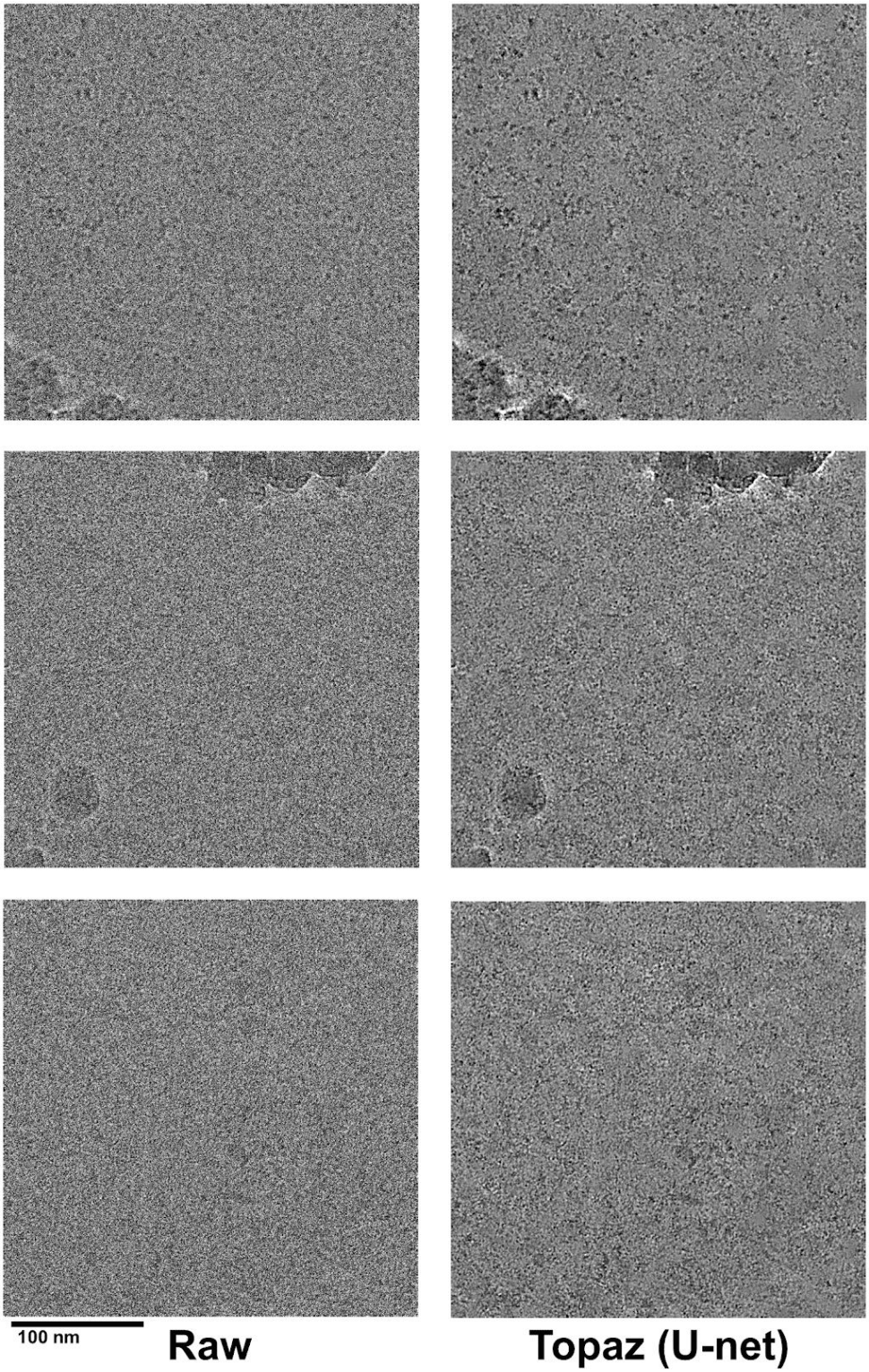
Supplemental Figure 5: Denoising sample with known background contaminants



Supplemental Figure 5 | Denoising raw micrographs (18sep15a) of microtubules with known background contaminant proteins, kinesin and tubulin. Topaz denoising appropriately accentuates features of the background proteins instead of smoothing them out.

bioRxiv preprint doi: https://doi.org/10.1101/838920; this version posted November 12, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

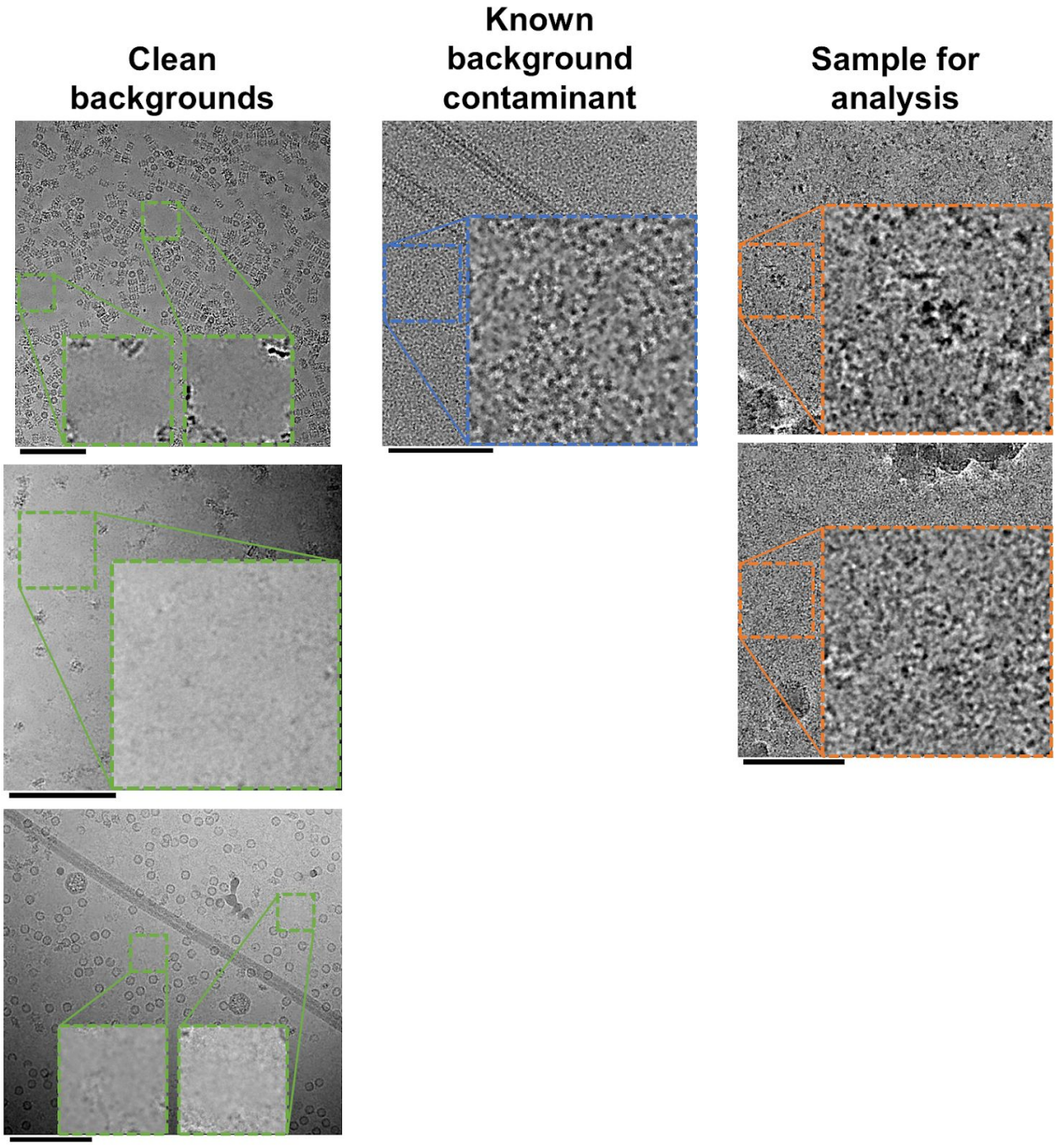
Supplemental Figure 6: Denoising EMPIAR-10003



Supplemental Figure 6 | Denoising of EMPIAR-10003 raw images (left) using the U-net model (right). Possible regions of proteins are particularly apparent in the top image.

bioRxiv preprint doi: https://doi.org/10.1101/838920; this version posted November 12, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

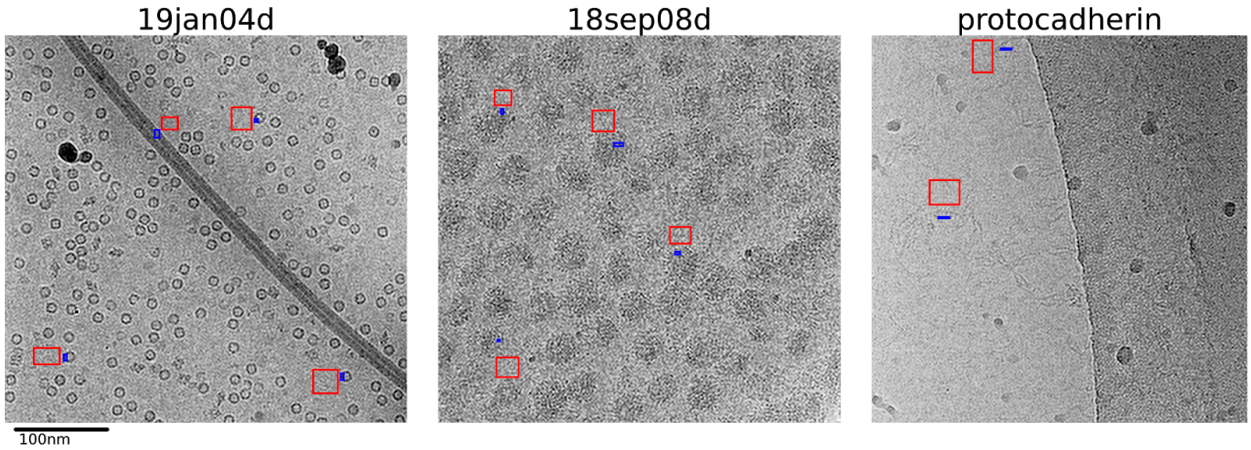
Supplemental Figure 7: Comparison of backgrounds after denoising



Supplemental Figure 7 | Denoising as a complementary method for analyzing background proteins and contamination in sample/grid preparations. Left: Three micrographs with nearly clean backgrounds (green insets). Middle: A micrograph of microtubules with known kinesin and tubulin background contaminant (blue inset). Right: Two micrographs from the EMPIAR-10003 dataset with the centers magnified (orange insets). All micrographs are denoised using our Topaz U-net model. Insets are magnified by 2x. Scalebars are 100nm.

bioRxiv preprint doi: https://doi.org/10.1101/838920; this version posted November 12, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

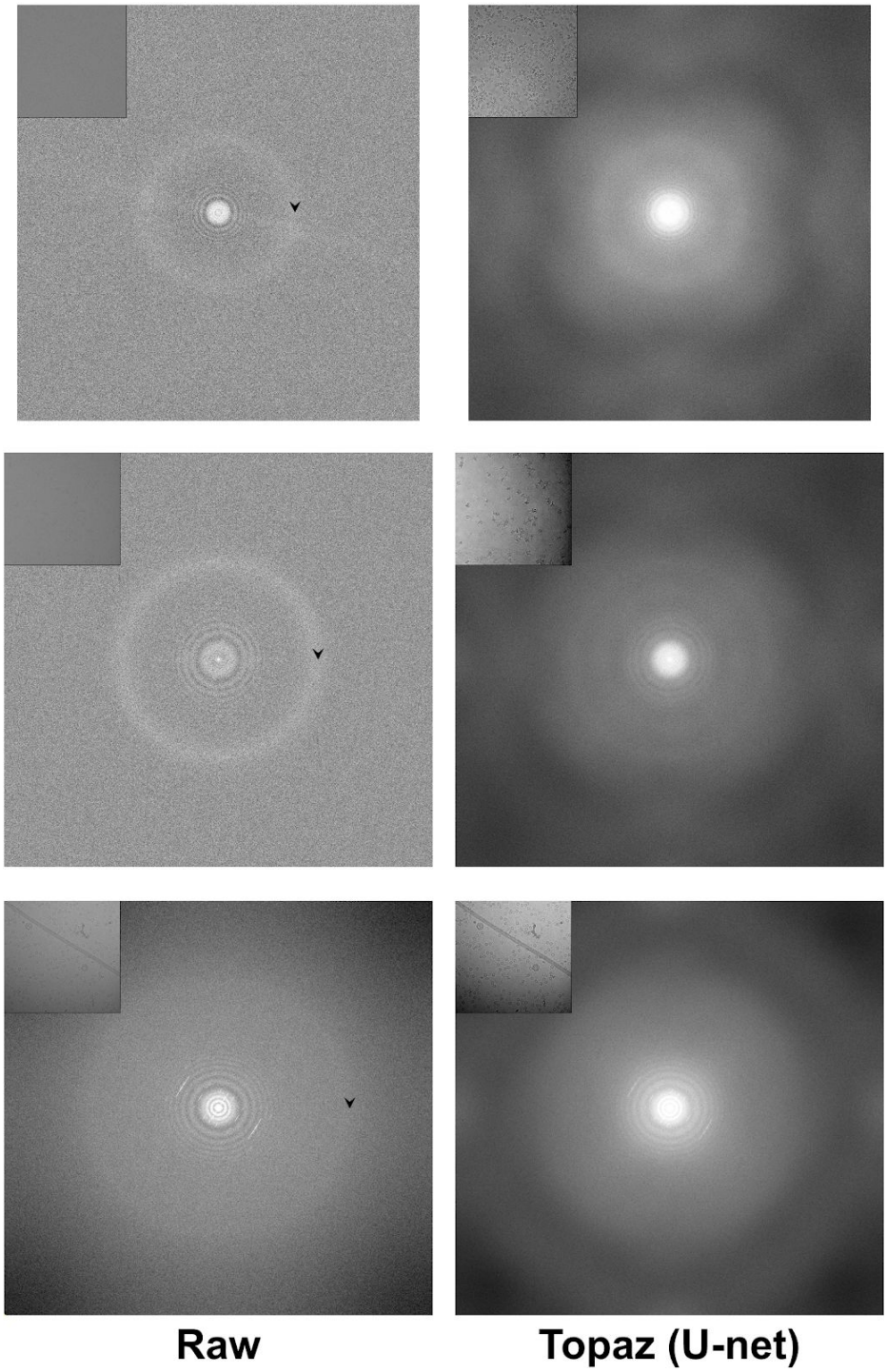
Supplemental Figure 8: Paired signal and background for SNR quantification



Supplemental Figure 8 | Example micrographs from 19jan04d, 18sep08d, and the protocadherin dataset showing labeled signal (blue) and background (red) regions overlayed over low-pass filtered images. Signal and background regions were selected close together to match local background properties as best as possible to each signal region.

bioRxiv preprint doi: https://doi.org/10.1101/838920; this version posted November 12, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

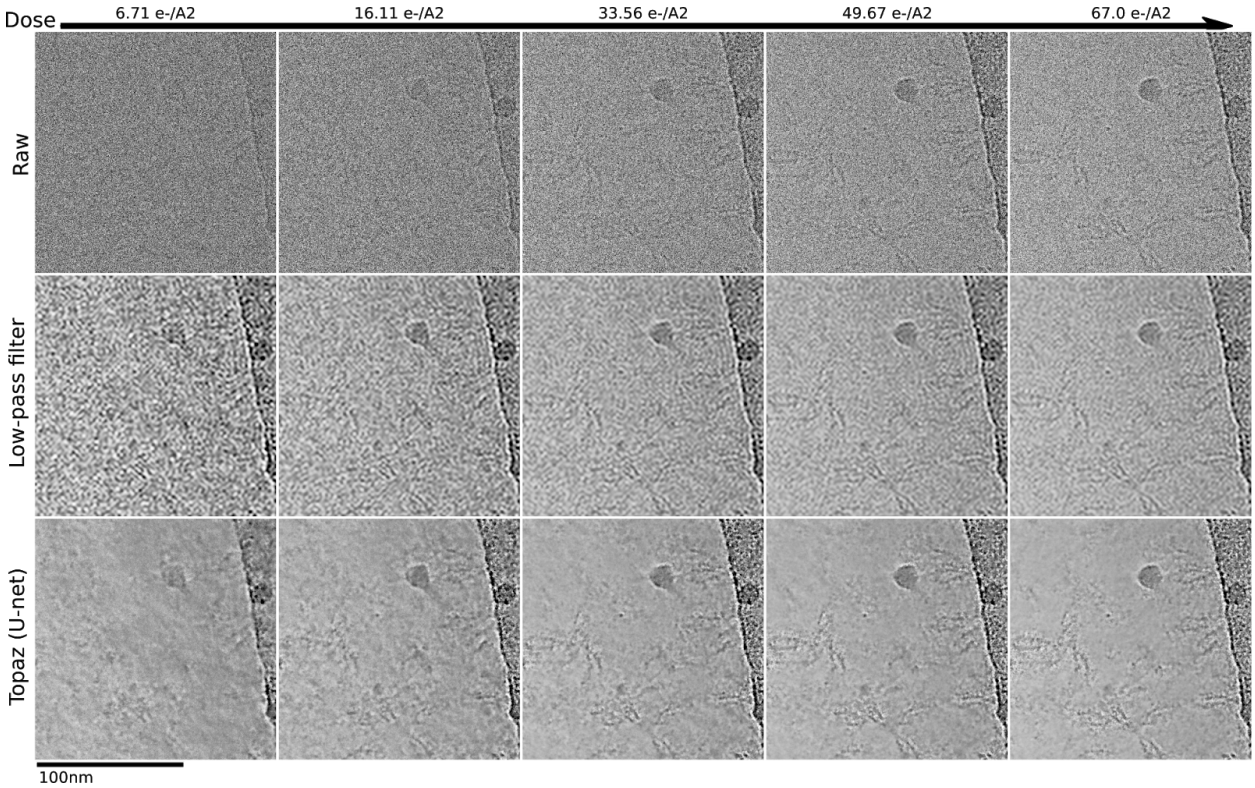
Supplemental Figure 9: Fourier transforms of raw and denoised micrographs



Supplemental Figure 9 | Fourier transforms of previously shown raw and denoised micrographs (insets). Top: Figure 2a, EMPIAR-10025. Middle: Figure 2b, EMPIAR-10261. Bottom: Figure 3, 19jan04d. Arrows show the location of the ~3.7 Å ice ring.

bioRxiv preprint doi: https://doi.org/10.1101/838920; this version posted November 12, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

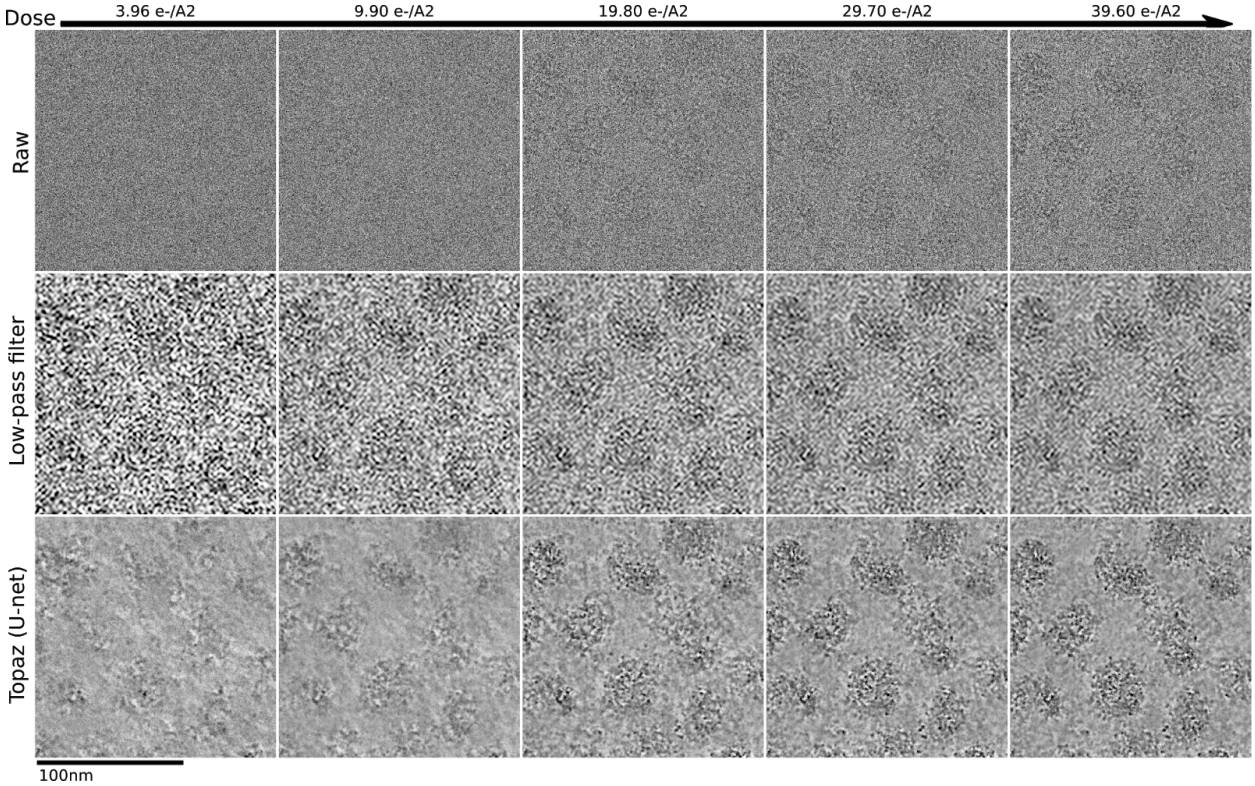
Supplemental Figure 10: Short exposure detail of protocadherin



Supplemental Figure 10 | Detail of protocadherin micrograph denoised and raw over increasing dose.

bioRxiv preprint doi: https://doi.org/10.1101/838920; this version posted November 12, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

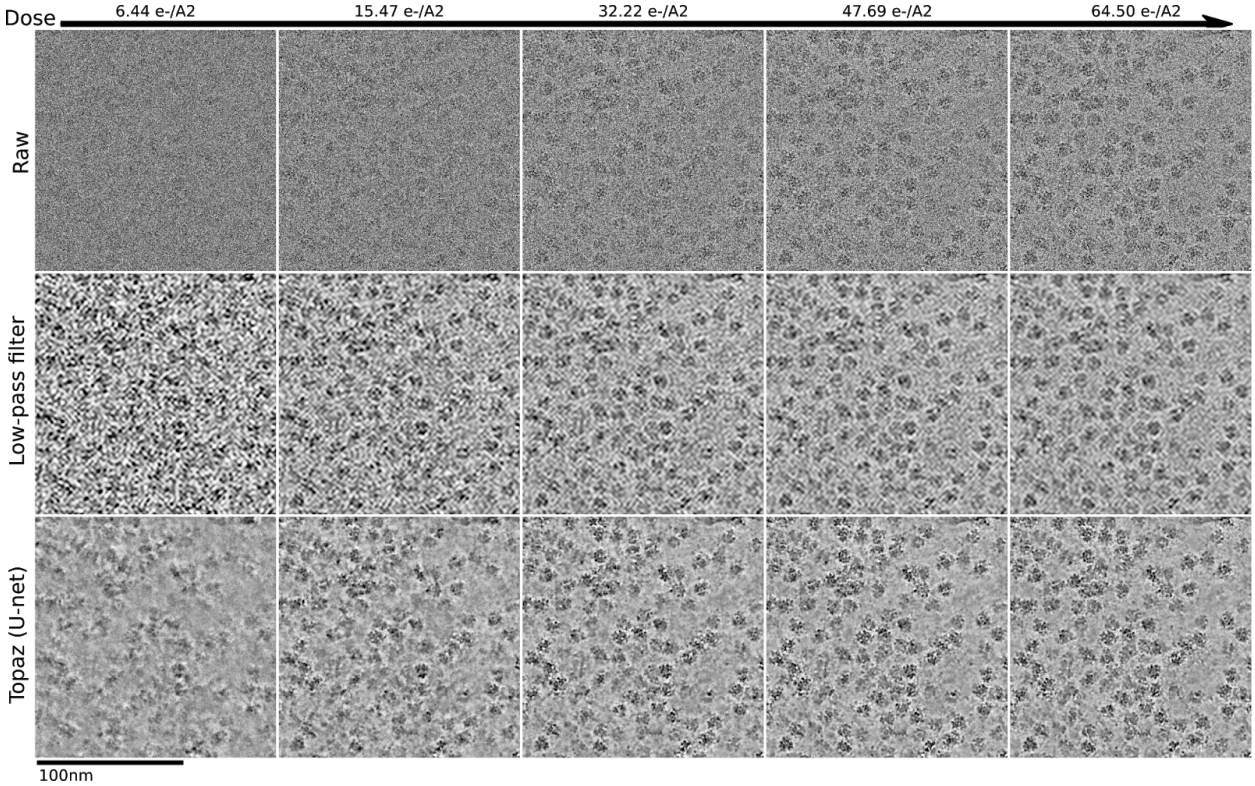
Supplemental Figure 11: Short exposure detail of 18sep08d (VLPs)



Supplemental Figure 11 | Detail of 18sep08d micrograph denoised and raw over increasing dose.

bioRxiv preprint doi: https://doi.org/10.1101/838920; this version posted November 12, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

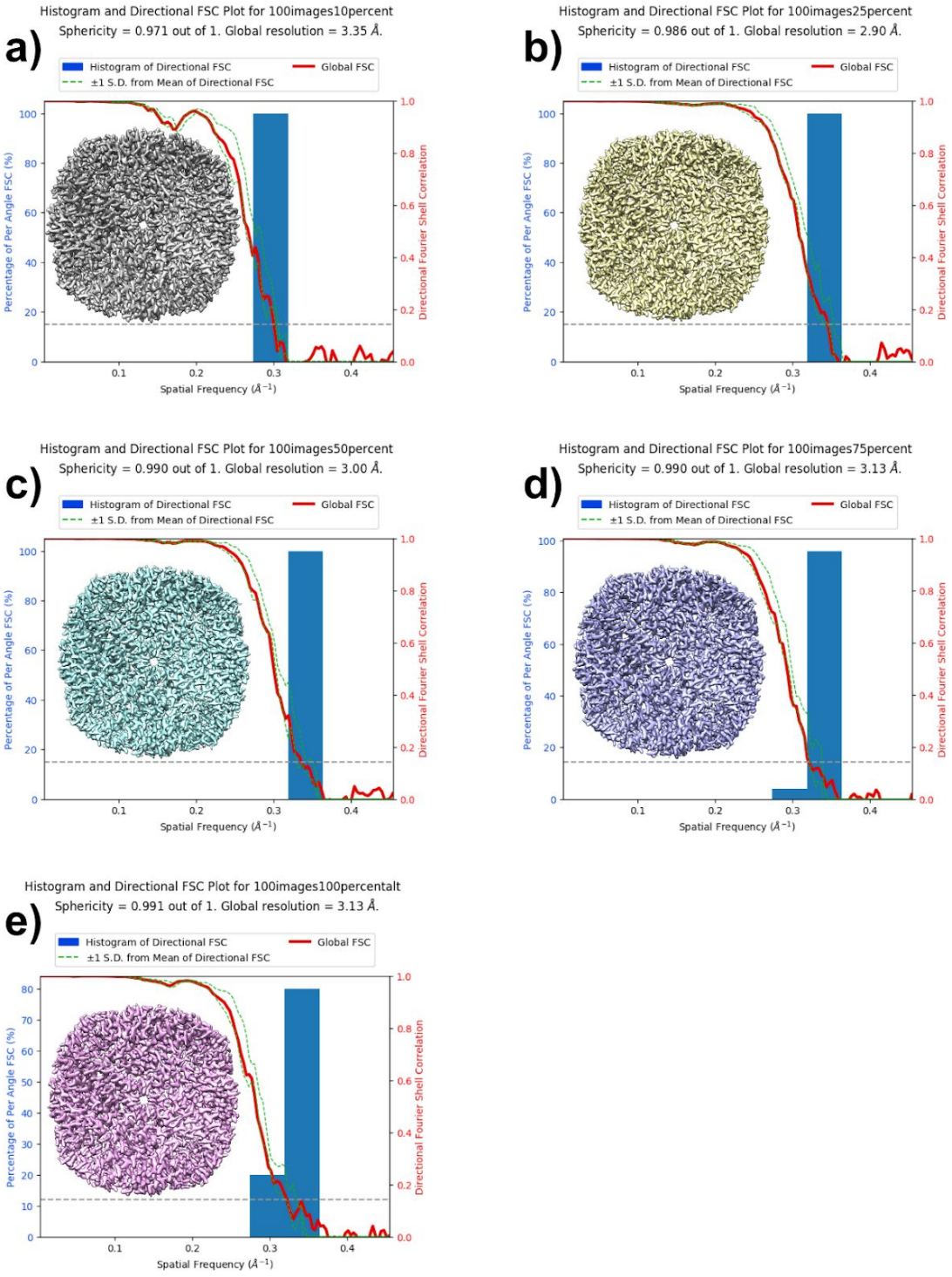
Supplemental Figure 12: Short exposure detail of 19may10e



Supplemental Figure 12 | Detail of 19may10e micrograph denoised and raw over increasing dose.

bioRxiv preprint doi: https://doi.org/10.1101/838920; this version posted November 12, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

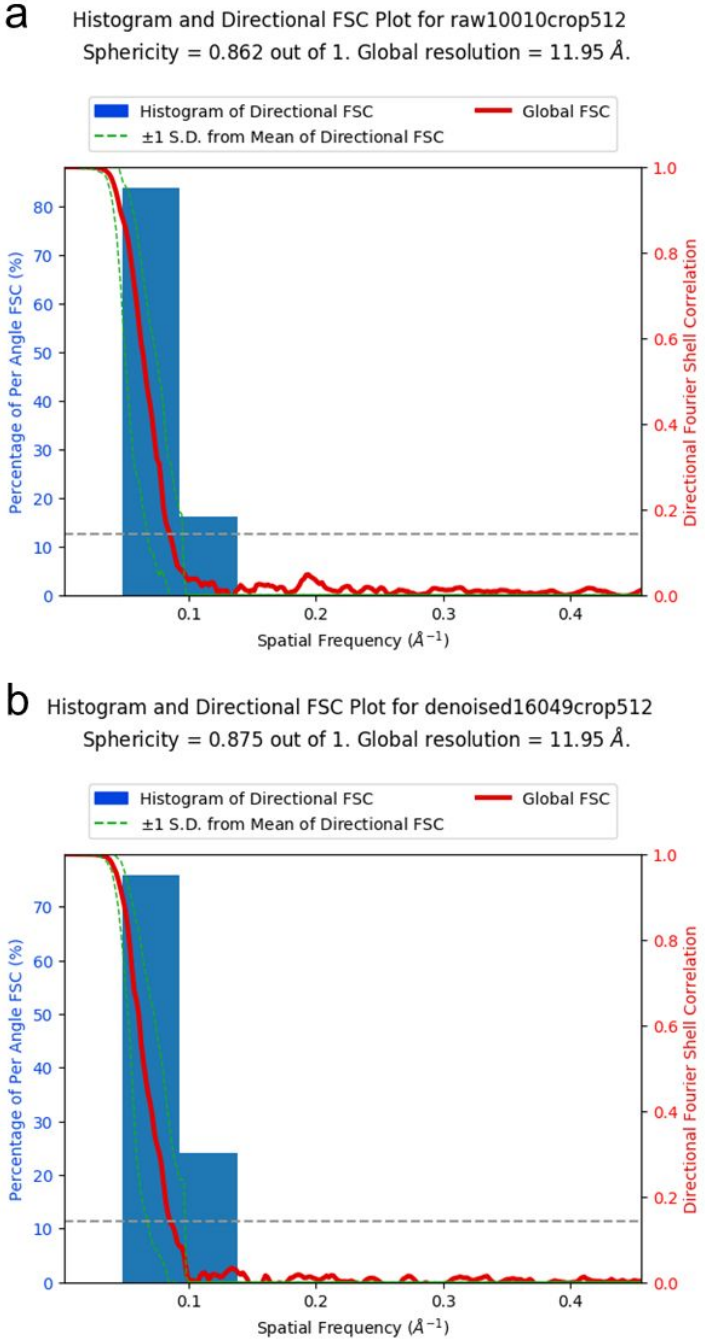
Supplemental Figure 13: Short exposure 3D reconstructions of apoferritin



Supplemental Figure 13 | 3DFSC plots of apoferritin particles fractionated by frames/exposure time: **(a)** 6.95 e-/Å^2, **(b)** 16.67 e-/Å^2, **(c)** 34.73 e-/Å^2, **(d)** 51.40 e-/Å^2, and **(e)** 69.50 e-/Å^2.

bioRxiv preprint doi: https://doi.org/10.1101/838920; this version posted November 12, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

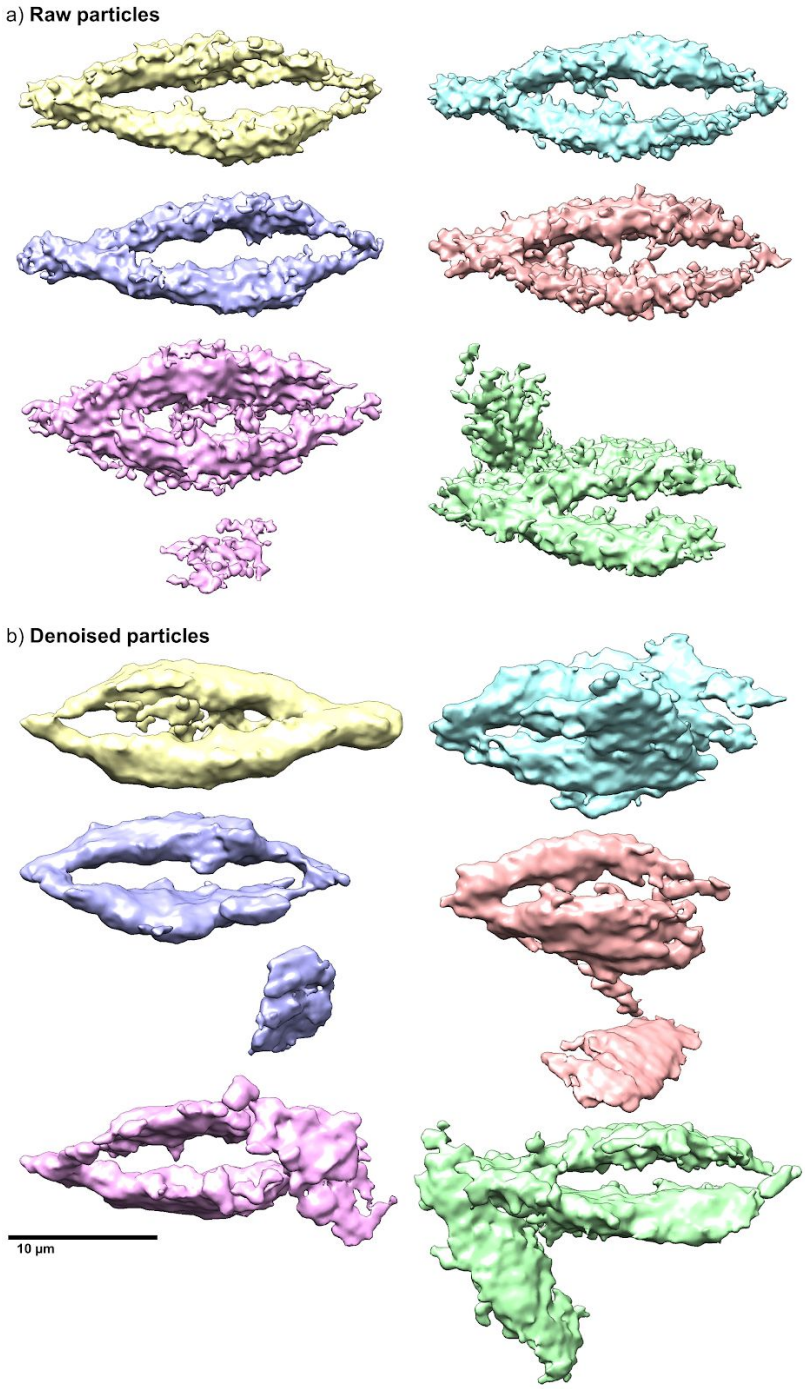
Supplemental Figure 14: Denoising improves manual picking completion for difficult particles



Supplemental Figure 14 | 3DFSC plots of clustered protocadherin (EMPIAR-10234). **(a)** Using the particle picks reported on in Brasch et al., 2019, resulting in 10,010 particles in the 3D reconstruction. **(b)** Using the particle picks reported on in the Methods (from manually picking on denoised micrographs prior to Topaz picking training on raw micrographs), resulting in 16,049 particles in the 3D reconstruction.

bioRxiv preprint doi: https://doi.org/10.1101/838920; this version posted November 12, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Supplemental Figure 15: Denoised particles are less reliable for ab-initio model generation



Supplemental Figure 15 | 1,000 random particles processed through CryoSparc ab-initio reconstruction using raw particles **(a)** and denoised particles **(b)**. 4 out of 6 reconstructions using raw particles result in the correct structure (a), while at best 1 out of 6 reconstructions using denoised particles result in the correct structure (b).