MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy

To the Editor: In recent years single-particle cryo-electron microscopy (cryo-EM) has enabled groundbreaking advancements in obtaining atomic-resolution structures of macromolecules[1,2]. Central to this success has been the broad application of direct electron detector cameras with high output frame rates, which has enabled the recording of images of frozen hydrated biological samples as dose-fractionated stacks of subframes (movies). Beam-induced sample motion that blurs the captured images can then be corrected by registering identical features in the subframes, followed by summing the registered subframes to produce a motion-corrected image[3]. Here we describe MotionCor2, a software tool for anisotropic correction of beam-induced motion (**Supplementary Software** and http://msg.ucsf.edu/em/software/index.html).

Beam-induced sample motion can be decomposed into two components, uniform whole-frame motion and nonuniform local motions that are idiosyncratic and vary across the image. We previously developed an algorithm[4], implemented in the program MotionCorr, that made use of redundant measurements of image shifts between all subframes to derive a least-squares estimate of relative motions between neighboring subframes. It provided an efficient correction of whole-frame image motions with sufficient accuracy for determination of near-atomic-resolution 3D reconstructions[5]. Alternative strategies assume that particles located near each other have similar trajectories, with particle motion either being corrected early[7] or as part of 3D refinement[6]. A commonly used strategy nowadays is to correct whole-frame motion first in raw micrographs and track particle motions toward the end of the refinement of final reconstructions.

Here we develop an algorithm based on an experimentally validated model that describes the sample motion as a local deformation that varies smoothly throughout the exposure (**Supplementary Methods**). Tilting experiments show that the nonuniform motions seen in movie stacks are projections of complex 3D sample deformations onto the image plane (**Fig. 1a** and **Supplementary Fig. 1a**). We use a time-varying 2D polynomial function to describe these projections. The image is first divided into patches, and motions within each patch are iteratively determined (**Supplementary Fig. 1b**). The resultant shifts are used to fit the 2D polynomial functions that smoothly vary with time. Each image subframe is subsequently remapped using this smooth polynomial function at each individual pixel (**Fig. 1b**) and summed with or without radiation-damage weighting. This algorithm is implemented as MotionCor2, a program running on Linux using multiple graphics processing units. It combines the correction of both uniform whole-frame motion and anisotropic local motion, and it streamlines all the necessary preprocessing steps including bad pixel detection and correction before the normal cryo-EM processing pipeline (**Supplementary Figs. 2** and **3**).

We tested the performance of MotionCor2 using two previously acquired single-particle cryo-EM data sets, the archaeal 20S proteasome[4] and the rat TRPV1 ion channel[5]; and we compared the results with the results from MotionCorr. MotionCor2 significantly improves high-resolution Thon ring signals and leads to better correlation with simulated contrast transfer functions (**Supplementary Fig. 4**). 3D reconstructions of the 20S proteasome were determined from images that were motion corrected by MotionCorr, Unblur[8] and MotionCor2, followed by refinement and reconstruction with RELION (**Fig. 1c–e**). Using MotionCorr and Unblur with dose weighting, but without further particle polishing, we obtained 3D reconstructions of 2.7(6)-Å and 2.6(9)-Å nominal resolutions. Using Unblur followed by particle polishing improved the resolution to 2.5(7) Å. Using MotionCor2 with dose weighting, we obtained a 3D reconstruction of 2.5(0)-Å resolution, a resolution comparable to that achieved by combining Unblur with particle polishing. We also tested applying the RELION per-frame B-factor weighting scheme instead of dose weighting, but with MotionCor2-determined shifts. This resulted in a slightly improved nominal resolution of 2.4(6) Å, which was identical to the nominal resolution of the reconstruction obtained from using particle polishing after MotionCor2 correction. This suggests that further tracking individual particle trajectories after local motion correction by MotionCor2 provides minimal, if any, improvement in resolution, and that further improvements in the dose-weighting scheme may be possible. Refinement of the atomic model against the 2.4(6)-Å map resulted in improved model validation statistics, consistent with our claim of significantly better resolution (**Supplementary Table 1** and **Supplementary Data**).

Similarly, we reprocessed our previous raw micrographs from the TRPV1 ion channel and redetermined its 3D reconstruction, improving the resolution from 3.5 Å to 3.1(5) Å (**Supplementary Fig. 5**). The improvements are particularly obvious in some transmembrane regions, where extra densities associated with the TRPV1 protein are seen to have well-defined features that can only now be interpreted as lipid molecules. We also tested whether MotionCor2 enabled motion correction on images recorded with low defocus at 200 kV, again with superior results (**Supplementary Figs. 6** and **7**).

Lastly, we tested the performance of MotionCor2 on movie stacks of a tomographic tilt series collected from frozen hydrated *Drosophila* centrioles, where a particle-tracking strategy is not applicable. Again, a comparison was made between MotionCor2 and Unblur in terms of restoring Thon rings of tilted images using the cross-correlation metrics from CTFFIND4. The results showed that local motion correction substantially improved Thon rings of every tilt image in this tomographic tilt series (**Supplementary Fig. 8**).

Overall, MotionCor2 is extremely robust and sufficiently accurate at correcting local motions so that the very time-consuming and

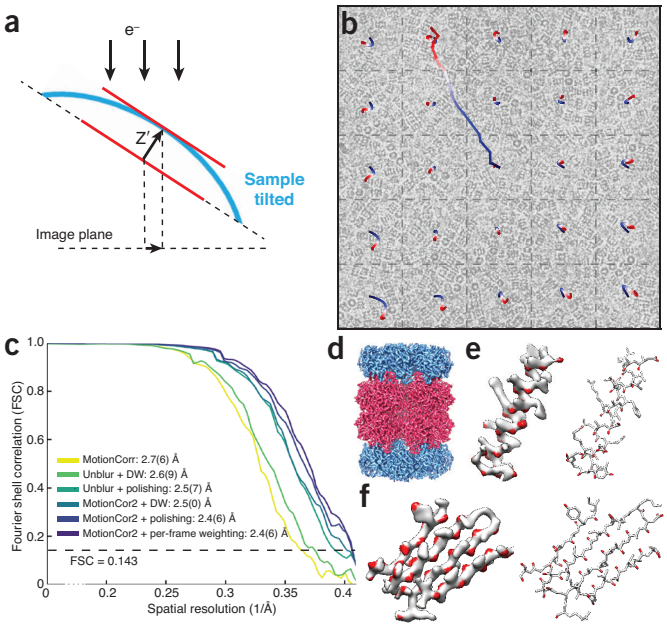

Figure 1 | Local motion correction by MotionCor2. (**a**) Schematic drawing illustrates that when the sample is tilted the observable motion in the image plan is the projection of *z*-motion produced by doming of the sample under electron beam. (**b**) Image of frozen hydrated archaeal 20S proteasome overlaid with the traces of global motion based upon whole-frame alignment (long trace originated from the center of image) and each patch predicted from the polynomial function. (**c**) Fourier shell correlation (FSC) curves of 3D reconstructions determined using micrographs corrected by Unblur with dose weighting, Unblur followed by particle polishing, correction by MotionCor2 with dose weighting, MotionCor2 followed by particle polishing and MotionCor2 with per-frame B-factor weighting. (**d**) 3D reconstruction of archaeal 20S proteasome filtered to 2.5-Å resolution and sharpened by a temperature factor of −103.8 Å^2. (**e**) Density of an α helix from the map, with resolved oxygen atom functional groups colored in red. Visualization of main chain carbonyls requires resolution below 3 Å. The refined atomic model is shown side by side for comparison. (**f**) As in **e**, but showing a β sheet.

computationally intensive particle polishing in RELION can be skipped. Importantly, it also works on a wide range of data sets, including cryo tomographic tilt series.

Data availability statement. The refined coordinates of archaeal 20S proteasome and the density maps of both archaeal 20S proteasome and TRPV1 are included as **Supplementary Data**.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper (doi:10.1038/nmeth.4193).

ACKNOWLEDGMENTS

We thank X. Li for helpful discussion during the initial stage of this work. We also thank M. Braunfeld for supporting the cryo-EM facility at UCSF, G. Greenan (Department of Biochemistry and Biophysics, University of California San Francisco) for providing his cryo-tomographic tilt series collected on a *Drosophila* centriole, and C. Kennedy for supporting the computational infrastructure for processing cryo-EM data. This work was supported in part by grants from National Institute of Health—R01GM031627 to D.A.A. and P01GM111126, P50GM082250, R01GM082893 and R01GM098672 to Y.C. Y.C. and D.A.A. are Investigators of Howard Hughes Medical Institute.

AUTHOR CONTRIBUTIONS

S.Q.Z. implemented the algorithm and wrote all codes for MotionCor2. D.A.A. contributed to algorithm development. S.Q.Z, E.P., Y.C. and D.A.A. designed experiments to evaluate the performance of MotionCor2. K.A.V. designed initial experiments for camera defect correction. S.Q.Z. and E.P. carried out image processing. E.P. and J.-P.A. collected low-defocus cryo-EM images of 20S proteasomes. S.Q.Z., E.P., Y.C. and D.A.A. wrote the manuscript. All authors participated in discussion and revision of the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Shawn Q Zheng[1,2], Eugene Palovcak[1], Jean-Paul Armache[1], Kliment A Verba[1], Yifan Cheng[1,2] & David A Agard[1,2]

[1]Department of Biochemistry and Biophysics, University of California, San Francisco, San Francisco, California, USA. [2]Howard Hughes Medical Institute, University of California, San Francisco, San Francisco, California, USA.
e-mail: ycheng@ucsf.edu or agard@msg.ucsf.edu.

Published online 27 February 2017; doi:10.1038/nmeth.4193

1. Cheng, Y. *Cell* **161**, 450–457 (2015).
2. Kühlbrandt, W. *Science* **343**, 1443–1444 (2014).
3. Brilot, A.F. *et al. J. Struct. Biol.* **177**, 630–637 (2012).
4. Li, X. *et al. Nat. Methods* **10**, 584–590 (2013).
5. Liao, M., Cao, E., Julius, D. & Cheng, Y. *Nature* **504**, 107–112 (2013).
6. Bai, X.C., Fernandez, I.S., McMullan, G. & Scheres, S.H. *eLife* **2**, e00461 (2013).
7. Rubinstein, J.L. & Brubaker, M.A. *J. Struct. Biol.* **192**, 188–195 (2015).
8. Grant, T. & Grigorieff, N. *eLife* **4**, e06980 (2015).

Automatic tracing of ultra-volumes of neuronal images

To the Editor: Despite substantial advancement in the automatic tracing of neuronal morphology in recent years[1–3], it is challenging to apply the existing algorithms to large image data sets containing billions or even trillions of voxels. Most neuron-tracing methods published to date were not designed to handle such data. We introduce UltraTracer (**Fig. 1**), a solution designed to extend any base neuron-tracing algorithm to allow the tracing of ever-growing data volumes. We applied this approach to neuron-tracing algorithms with different design principles and tested it on human and mouse neuron data sets that have hundreds of billions of voxels. Results indicate that UltraTracer is scalable, accurate, and more efficient than other state-of-the-art approaches.

The core algorithm of UltraTracer (**Fig. 1**) reconstructs a neuron structure from the available image data on the basis of a formulation of maximum-likelihood estimation. The underlying assumption is that the occurrence of a specific neuron structure could be modeled using the joint probability of all of its subparts given the image. Briefly, UltraTracer iteratively factorizes the joint probability based on progressive maximization of conditional probabilities of the occurrence of salient and continuous subparts of a neuron (**Supplementary Note**). Therefore, UltraTracer explores an image by following where the neurite signal goes, on the basis of either adaptive windows generated based on the already reconstructed neuron structure or certain domain knowledge (prior information or statistics) of neuron morphology, to help refine the choice of the next tracing subarea (**Supplementary Note**). This process repeats until the neuron structure grows as completely as possible. We designed the UltraTracer software to quickly extract an arbitrary subvolume of interest from large neuron image files (**Supplementary Note**), and thus smoothly traced an image archive without the need to load a large number of image voxels into computer memory.