Published in final edited form as:
Nat Methods. 2018 October ; 15(10): 777–780. doi:10.1038/s41592-018-0140-x.

Multicolor single particle reconstruction of protein complexes

Christian Sieben[1,3,*,#], **Niccolò Banterle**[2,*], **Kyle M. Douglass**[1], **Pierre Gönczy**[2,3], and **Suliana Manley**[1,3,#]

[1]Laboratory for Experimental Biophysics, Institute of Physics, École Polytechnique Fédérale de Lausanne (EPFL), 1015 Lausanne, Switzerland [2]Swiss Institute for Experimental Cancer Research, School of Life Sciences, École Polytechnique Fédérale de Lausanne (EPFL), 1015 Lausanne, Switzerland [3]Swiss National Centre for Competence in Research (NCCR) in Chemical Biology

Abstract

Single-particle reconstruction (SPR) from electron microscopy images is widely used in structural biology, but lacks direct information on protein identity. To address this limitation, we developed a computational and analytical framework that reconstructs and co-aligns multiple proteins from 2D superresolution fluorescence images. We demonstrate our method by generating multicolor 3D reconstructions of several proteins within the human centriole, revealing their relative locations, dimensions and orientations.

Macromolecular complexes within cells usually contain multiple protein species, whose precise arrangement is required for properly functioning molecular machines. Single particle analysis of electron microscopy (EM) images can build 3D reconstructions of such complexes, recently with near-atomic resolution1,2. To deduce the organization of specific proteins, computational methods have been used to dock structures from X-ray crystallography or NMR within 3D reconstructions1,3. Alternatively, immunogold or nanobody labelling can reveal the location of target proteins4,5, whereas electron density map differences can reveal the position of mutated or missing proteins5–7. Nevertheless, it remains challenging to locate native proteins within 3D reconstructions, which is essential

[#]Correspondence should be addressed to: Suliana Manley, suliana.manley@epfl.ch or Christian Sieben, christian.sieben@epfl.ch.
[*]authors contributed equally

Code availability
The version of SPARTAN used in this paper is available as Supplementary Software. Please see Supplementary Table 3 for a comparison of SPARTAN with other studies demonstrating single particle processing. The most recent version of SPARTAN, including the source code, can be downloaded from GitHub (https://github.com/christian-7/MultiColorSPR) following the instructions in the wiki.

Data availability. The data sets generated and analyzed in this study are available from the corresponding authors upon reasonable request. Sample datasets are available on Zenodo (https://doi.org/10.5281/zenodo.1288783).

Author Contributions
C.S, N.B, P.G, and S.M. conceived and designed the project. C.S, P.G. and S.M supervised the project. C.S. and N.B. performed all experiments and data analysis. C.S and K.M.D wrote analysis code. All authors wrote, revised and contributed to the final manuscript.

Competing Financial Interests Statement
The authors declare no competing interest.

Europe PMC Funders Author Manuscripts

for deciphering the assembly mechanisms and functional modules of macromolecular complexes.

Fluorescence-based single-molecule localization microscopy (SMLM) can help address this challenge, as demonstrated by 2D averaging of nuclear pore complexes8. Extending into 3D, single-particle reconstruction (SPR) demonstrated isotropic reconstruction from 2D SMLM images of DNA origami and simulated data9. However, multicolor particle reconstruction of actual macromolecular complexes requires generating large image libraries of multiple proteins and solving the problem of 3D multi-channel alignment. Here, we developed a systematic framework that addresses both of these challenges. We used a dedicated high-throughput SMLM setup10 to collect large multicolor particle datasets, which we processed using a semi-automated computational workflow to reconstruct and align multiple proteins onto a single 3D particle. Our reconstruction workflow comprises three successive steps (Supplementary Fig. 1): 1) SMLM imaging and particle extraction; 2) reconstructing separate 3D protein volumes (i.e. the reference and the protein(s) of interest); 3) mapping one or multiple proteins of interest onto the reference. To facilitate the analysis, we developed SPARTAN, an SMLM single-particle analysis software providing the major processing steps via a convenient graphical user interface (Supplementary Note 1, 2). We applied our method to the human centriole, reconstructing and aligning both on- and off-axis structures. Centrioles are evolutionarily conserved sub-diffraction limited organelles that seed the formation of cilia, flagella and centrosomes11. The mature human centriole comprises nine-fold symmetrically arranged microtubule triplets and contains >100 different proteins organized into distinct substructures12. For instance, distal appendages harbor the protein Cep164 and are key for cilium and flagellum formation13. A torus encircling the proximal part of the mature centriole and comprising the proteins Cep57/Cep63/Cep152 acts as a nucleation site for the new procentriole, whose assembly relies on the self-organization of the HsSAS-6 protein into a cartwheel14,15. The details of component dimensions within the Cep57/Cep63/Cep152 torus and the orientation of the emerging procentriole with respect to this torus remain unclear.

To demonstrate our multicolor 3D SMLM reconstruction workflow, we imaged proteins within centrioles and procentrioles. Centrosomes were isolated from human KE37 cells arrested in S phase, concentrated onto coverslips by centrifugation, then immunolabeled and stained (Supplementary Note 3). Next, we used dual-color high-throughput SMLM10 to image ~100-300 centrioles per field of view (Supplementary Fig. 2). Localizations belonging to centrioles and procentrioles (hereafter “particles”) were segmented using a mask generated through automated OTSU thresholding of the widefield images. A density-based filter (DBSCAN16) was then applied to separate adjacent particles (Supplementary Fig. 3, 4). Only densely labelled particles (typically 10-20% of the initial dataset) were rendered and used to populate the particle library.

Next, the reconstruction of a 3D volume from single particles of unknown orientation was performed using well-established EM routines (Supplementary Note 4). Particles from both fluorescence channels (Supplementary Fig. 5) were classified using Clustering 2D (CL2D)17 or template-free maximum-likelihood multi-reference refinement (ML2D)18. Due to the high degree of radial symmetry within centrioles11, a low number of classes

(typically 8-15) was chosen, thereby reducing computational complexity (Fig. 1a, Supplementary Fig. 6). Although we applied symmetry information as a last step to overcome non-uniform angular coverage (Supplementary Fig. 7), our workflow is capable of reconstructing and recovering the symmetry of unknown objects given a number of particles sufficient to fully sample the orientational space, as verified *in silico* (Supplementary Fig. 8). Information lost by underrepresented orientations would produce missing wedge artifacts similarly to electron tomography. The class averages best resembling the input particles (Supplementary Note 4) were then used to compute an initial 3D model followed by structural refinement based on matching its 2D projections to the input particles. In this manner, we reconstructed the torus protein Cep152, and measured its diameter to be ~270 nm (Fig. 1a), consistent with the ~242 nm measured for SNAP-Cep152 by STED microscopy when accounting for antibody size19. In addition, our 3D reconstruction revealed that the height of the torus is ~160 nm (Supplementary Fig. 9). Following the same procedure, we reconstructed the well-known bacteriophage T4 (Supplementary Fig. 10), demonstrating the generality of this 3D SMLM reconstruction workflow. In this case, particles aligned preferentially parallel to the coverslip, resulting in uneven angular sampling (Supplementary Fig. 7) which we compensated using the known phage symmetry.

The general alignment of two volumes requires both translation and rotation of one of the volumes in 3D. The problem is less complex if the two proteins are symmetrically arranged and further simplified if they share a symmetry axis. To achieve multicolor reconstruction, we first considered the latter case of proteins sharing a principal symmetry axis, where the only alignment parameter is the displacement along the symmetry axis, Δz. We collected dual-color images of Cep152/Cep164, Cep152/Cep57, and Cep63/Cep57, using Cep152 and Cep57 as reference proteins. This reduced the problem of alignment to only two 3D volumes at a time. A consistent direction of the displacement was assumed, but a third marker could be used to determine particle orientation. We divided the alignment process into two steps: i) co-orient both particles and reconstruct their volumes; ii) translate one volume by the correct distance Δz along the symmetry axis (Fig. 1b). Since both proteins are integrated into the same structure, the corresponding particles share the same relative orientation. Therefore, it suffices to find the orientation of particles in one channel (i.e. the reference), then preserve and assign the orientation to the second channel. Given the challenge of imaging two or more proteins with superresolution due to often low protein abundance and/or labeling efficiency, this procedure offers the great advantage that only the reference protein images must contain enough information to be oriented. Following this procedure results in two co-oriented volumes.

To align the two protein volumes, it only remains to find the translational shift, Δz, between their side view (xz) projections (Fig. 1c-d, Supplementary Fig. 6). Thus, we performed orientational filtering by using supervised machine learning to identify top and side view projections of the reference protein from a combination of 12 calculated shape descriptors (Supplementary Fig. 11a, Supplementary Note 5). After models were trained on ~10% of particles, they successfully identified ~97 % (true positive rate) of side view projections (Supplementary Fig. 11b), typically yielding 50–100 side view particles per imaged protein pair. This method offers the advantage that after having been trained for a reference protein, it can directly be applied to other datasets using the same reference. Since individual two-

color particles suffer from heterogeneous labelling (Supplementary Fig. 12), we generated averaged side view projections via a 2D alignment comprising particle rotation in 3° increments followed by translational alignment and cross-correlation (Supplementary Fig. 13). These aligned averages permit a more precise estimate of Δz, and reveal the average particle dimensions and symmetry (Supplementary Fig. 9, 14, 15, 16).

Importantly, this workflow allowed us to reconstruct and co-align the toroidal complex Cep57, Cep152 and Cep63, with the distal appendage protein Cep164 in a four-color volumetric reconstruction of the mature human centriole (Fig. 1e, Supplementary Video 1). This revealed that whereas the Cep57 torus is aligned axially with Cep152 and Cep63 volumes, as expected from their known association in cells19, it has the smallest dimensions of the three (~230 nm in diameter and ~130 nm in height), placing it near the outer microtubule wall. We also discovered a nine-fold radially symmetric distribution of Cep57 and Cep152 (Supplementary Fig. 15), further suggesting association with the nine-fold symmetric outer microtubule wall of the centriole, perhaps via the microtubule binding domain of Cep5720. We confirmed Cep164's previously observed nine-fold symmetry, while locating its N-terminus more proximally and closer to the centriolar wall than previously reported21 (see also Supplementary Fig. 15).

The above approach works well for proteins sharing a principal symmetry axis, but there are important exceptions. We thus extended our method to the procentriole, marked by the protein HsSAS-6, which assembles from a single focus on the torus containing Cep57/Cep63/Cep152 22, and was suggested by EM to initially take on non-orthogonal orientations23. We collected dual-color images of Cep152/HsSAS-6, and generated average top and side views following the previous workflow (Fig. 2a). The orientation of Cep152 was insufficient to define that of HsSAS-6, since the two proteins do not share a symmetry axis (Fig. 2a). Therefore, we combined the images from both proteins into a single channel and performed class averaging and alignment on the resulting dataset. However, we found that when using a simple sum of the two channels, the signal from Cep152 dominated and prevented alignment of the smaller HsSAS-6 volume (Supplementary Fig. 17). To overcome this, we combined the two channels in a weighted sum (Supplementary Fig. 17, 18), and used the combined particles for structural refinement of the initial Cep152 volume with no symmetry constraint (Fig. 2b). Finally, we fit the individually reconstructed protein volumes into the asymmetric global structure to achieve a two-color volumetric reconstruction of the nascent procentriole in the context of the centriolar torus (Fig. 2c, Supplementary Video 2). The combined reconstruction has a lower resolution than the individual structures (Fig. 2c), likely reflecting a flexible relative positioning of the two entities. Indeed, we found the angle θ between the two proteins measured from individual side view particles (Fig. 2a) to be variable, with an average value of 15.4 ± 4.5 (SD, n = 75), in agreement with the angle obtained from our 3D reconstruction (θ = 13). Finally, in a three-color experiment, we used Centrin to mark the distal end of the centriole24 (Supplementary Fig. 19), revealing a preferential orientation of the procentriole toward the distal end. Together, these findings support a loosely defined orientation between the torus and the emerging procentriole, with a broken distal-proximal symmetry.

In conclusion, we developed a framework that generates multicolor 3D volumes from dual-color 2D SMLM datasets, and used it to reveal unknown features of human centriole and procentriole architecture. Our approach is directly applicable to any single particle dataset with sufficient angular coverage, although its extension to repetitive structures such as helices would require adaptation. Our flexible workflow is implemented in a software package that is suitable for other multiprotein complexes and imaging modalities. Combining information from 3D SMLM reconstructions with EM particle reconstructions will likely prove invaluable in the future, as will improvements in labelling, to permit higher fidelity of multicolor images to the underlying structure.

Online Methods

Centrosome and bacteriophage T4 preparation

Human centrioles were purified from KE37 cells incubated for 24h with thymidine following a standard protocol 25,26 and spun (10 min at 10,000 g in Corex tubes, JS13.1 Beckman swinging rotor) in 10 mM K-Pipes on gold-embedded fiducial cover slips (custom 18 mm, Hestzig) using a custom centrifuge concentrator, followed by methanol fixation (5 min at -20°C). Samples were then immunostained by overnight incubation at 4° C with primary antibodies (Supplementary Table 1), diluted 1:500 (in PBS supplemented with 1% BSA and 0.1% Tween 20), then washed three times 15 min in PBS and incubated with secondary antibodies coupled with Alexa 647 or DyLight 755 for 1h at room temperature. Finally, the samples were washed again three times for 15 min and stored in the dark at 4 °C until further use.

Bacteriophage T4 was grown and purified following established procedures27. To characterize the purified sample, phages were spotted on mica and imaged using atomic force microscopy (JPK Nanowizard). To achieve all-protein labelling, phages were incubated with Alexa 647 NHS-Ester (Life Technologies) (final concentration 10 μM in phosphate buffer, pH 8) overnight at 4° C. The labelled phages were separated from unbound dye using a NAP-5 size exclusion column (GE Healthcare) and stored in the dark at 4 °C until further use. Before SMLM imaging, phages were adsorbed on plasma-cleaned glass cover slips (1.5, Menzel, 25 mm) after coating with 0.1 % poly-L-lysine (Sigma) for 30 min.

Imaging sample preparation

Samples were imaged on gold-embedded fiducial cover slips (custom 18 mm, Hestzig). Imaging buffer components were purchased from Sigma. Additional gold fiducials were obtained from Corpuscular (C-Au-0.1) and diluted (1:5) in 0.1 % poly-L-lysine (Sigma) before application.

To create a bead sample for two-channel registration, glass cover slips (1.5, Menzel, 25 mm) were plasma cleaned, coated with 0.1 % poly-L-lysine (Sigma) for 30 min and incubated with FluoroSpheres (Dark Red, F8789, Life Technologies) diluted (1:50,000) in water for 10 min.

High-throughput SMLM

Two-color SMLM imaging was performed using a flat-field epi illumination microscope10. Briefly, two lasers with wavelengths of 642 nm (2RU-VFL-P-2000-642-B1R, MPB Communications) and 750 nm (2RU-VFL-P-500-750-B1R, MPB Communications) were used to switch off fluorophores within the sample, while a 405 nm laser (OBIS, Coherent) controlled the return rate of the fluorophores to the fluorescence-emitting state. A custom dichroic (ZT405/561/642/750/850rpc, Chroma) reflected the laser light and transmitted fluorescence emission before and after passing through the objective (CFI60 PlanApo Lambda x60/NA 1.4, Nikon). After passing the respective emission filter (ET700/75M, Chroma or ET810/90m, Chroma), emitted light from the sample was imaged onto an sCMOS camera (Prime, Photometrics, pixel size 106 nm). The sample was excited with laser output power of 1200 mW (642 nm) and 500 mW (750 nm), corresponding to 1000 mW (642 nm) and 350 mW (750 nm) at the objectives back focal plane. The 405 nm laser was operated with laser output power 1-10 mW. Axial sample position was controlled using the pgFocus open hardware autofocus module (http://big.umassmed.edu/wiki/index.php/PgFocus). Typically, 30-60k frames at 10 ms exposure time were recorded for each field of view using Micromanager28. Single- and dual-color SMLM imaging was performed using an optimized SMLM buffer as described previously29. See Supplementary Note 3 for more details on the choice of the fluorophores and buffer preparation.

Single-fluorophore localization, channel registration, drift correction

Image stacks were analyzed using a custom CMOS-adapted analysis routine (adapted from30). Correction of the Alexa 647 and DyLight 755 datasets for spherical and chromatic aberrations and lateral sample drift was carried out in three steps. The first step corrects for differences and aberrations (rotation, magnification) of the emission path between the two detection channels. We calculated a local weighted mean (LWM) transformation from images of fluorescent beads (see Materials and sample preparation section above) acquired in both channels, and applied it to the 755 channel to match the positions in the 647 channel. During the second step, both datasets were independently drift-corrected using gold fiducials visible in both channels. For each field of view, we selected 3–6 fiducial markers across the field of view and used their average trajectory for drift correction. As a third step, the drift-corrected fiducial centroid positions from both channels were matched by applying a final lateral translation again to the 755 channel. All processing steps were performed in MATLAB 2016a (Mathworks) and are available as part of the supplementary software package (SPARTAN>Image Registration).

Cumulative error estimation

The LWM transformation resulted in an uncertainty of around ~10 nm. This remaining uncertainty is also referred to as the target registration error (TRE)31. After the second registration step, where the drift-corrected fiducial centroids are matched using a rigid translation, we find a final TRE between 10 - 20 nm. Note that the TRE increases after the second registration due to the added uncertainty of the drift correction.

Consequently, we estimate the total uncertainty σ_{total} as the sum of the individual error components $\sigma_{total} = \sigma_{loc} + \sigma_{Reg} + \sigma_{ab}$, where σ_{loc} is the localization precision

(Supplementary Figure 20), σ_{Reg} the final TRE after both registration steps and σ_{ab} represents the uncertainty added by a primary-secondary antibody conjugate. $\sigma_{total} = \sqrt{12^2 + 15^2 + 15^2} \sim 24$. This total uncertainty sets a lower boundary for the resolution of the obtained structure. The final resolution if the 3D model is further affected by the error introduced though non-homogeneus and/or sparse labeling. Considering this, the obtained value is in agreement with the obtained 3D resolution as measured by FSC (Supplementary Table 2).

Particle segmentation and 3D reconstruction

Following channel registration, the two localization datasets were ready for particle segmentation (SPARTAN>Particles>Particle Segmentation). The localization maps for each field of view were loaded into MATLAB together with the corresponding wide-field (WF) images taken prior to the SMLM stack acquisition. A WF image was used for automatic OTSU segmentation to identify the approximate location of individual particles within each field of view. Here we used the higher-contrast widefield image (typically from the reference protein). To accommodate small shifts between WF image and localization data, each identified region was expanded by up to five pixels on all sides.

Overlapping regions were removed and the localizations from both channels were extracted for each segmented particle. Particles were filtered for a minimum number of localizations (typically >100) to ensure good particle labelling. We also applied an upper cut-off to reject clusters of particles and misidentified gold fiducials. During the next step, labelling noise was removed and adjacent particles within the same region were separated using density-based clustering (DBSCAN16). An example is shown in Supplementary Fig. 3. If a low-density protein of interest was imaged (e.g. HsSAS-6), we used an additional filter selecting only the largest cluster (Supplementary Fig. 4). We then calculated a number of particle quality and shape descriptors (Supplementary Note 5), as well as the resolution (using Fourier ring correlation32,33) for each particle, which allowed for efficient particle filtering. Finally, particles from both channels were rendered into a pixel image using a 2D histogram function with a bin size of 10 nm and blurred using a Gaussian filter with σ corresponding to the measured localization precision. The final image approximates the probability density distribution of the fluorescent labels on the underlying structure and is a widely-used approach to visualize SMLM data. The particle images were stitched together using the Montage function in ImageJ (Miji for MATLAB) resulting in the final input image for the 3D reconstruction (example shown in Supplementary Fig. 5).

Single particle reconstruction was performed using Scipion, a freely available software package that integrates several widely-distributed and well-developed 3D EM particle reconstruction routines34. A brief tutorial of the required steps is provided in Supplementary Note 4. The particle montage images were imported into Scipion. Depending on the size of the dataset (400 – 6000, Supplementary Table 2), each montage contained ~500 particles; thus, each reconstruction required the generation and import of multiple montages. During particle extraction, we removed labelling noise around the densely labelled particles (Supplementary Figure 3), resulting in a high-contrast particle montage, which facilitates automatic particle picking (Xmipp3). The particles were then aligned using 2D clustering

(CL2D, Xmipp317) and classified using template-free multi reference maximum likelihood (ML2D, Xmipp318) or 2D clustering (CL2D, Xmipp317). Class averages were visually inspected. Some classes (typically 1-2) accumulate particle fragments which we removed at this stage. The remaining classes were used (see Supplementary Note 4) to generate the initial model. For symmetric centriolar reconstructions (Cep164, Cep57, Cep152, Cep63), we used between 8-15 classes. To allow reconstruction with a limited number of particles, we applied rotational or nine-fold symmetry at the final stage to fill in missing angular information (Supplementary Note 4). For bacteriophage T4, we used 6 class averages and calculated the initial model using rotational six-fold symmetry (c6). Initial models were calculated using either Xmipp235 or Eman236, with both classifications providing similar results. Finally, the initial model was refined using particle back projection (Xmipp3). Fourier shell correlation (FSC) was calculated during particle refinement (particle back projection, Xmipp3).

For the reconstruction of two proteins, we first reconstructed the reference protein using the steps described above, then applied the final alignment from each reference particle to the corresponding particle of the protein of interest (function alignment assign). The co-orientated particles of the protein of interest could then be reconstructed (function reconstruct in Relion37 or Eman36). See Supplementary Note 4 for a more detailed description. To generate a two-color volume of proteins sharing the same principal symmetry axis, it is also possible to use two independently reconstructed volumes. Here only the volume alignment (i.e. determination of Δz) requires a two-color SMLM dataset. The asymmetric reconstruction was performed using an adapted workflow in Scipion (Supplementary Figure 17). We first reconstructed each protein separately. The symmetric volume of the reference (Cep152) was then refined using the weighted sum of the input particle images (Cep152+2*HsSAS-6) without applying a symmetry constraint. Into the resulting asymmetric joint volume, we fit both individual protein volumes to obtain a high-resolution dual-color model. The volume fitting was performed with Chimera38 using the ‘Fit in Map’ tool (Tools > Volume Data > Fit in Map).

2D particle averaging and volume alignment

In general, the relative offset between the centers of mass of two distinct, three dimensional particles can be determined by triangulating between any two projected views. In the case of the centriole which has a principal rotational symmetry axis, only one projected orientation is needed to align two volumes. Any orientation that is not orthogonal to the symmetry axis (top view, xy), together with knowledge of the angle between the projection and the symmetry axis, is sufficient to determine the axial distance. In the simplest case, we can directly determine the axial distance between two volumes (Δz) from the projection into a plane parallel to the z axis (xz). We used particle projections of centriolar side (xz) views for this purpose (Supplementary Fig. 14). Below, we describe the orientational filtering of top (xy) and side (xz) views. While the volume alignment only requires the side view orientation, we use the top views to characterize the protein’s symmetry properties (Supplementary Figure 15). To efficiently identify particles with these orientations among a large number of individual particles with different orientations, we calculated 12 shape parameters whose values could be used as a characteristic signature for top (xy) and side

(xz) view projections (Supplementary Note 5). A similar approach was used recently to filter out centriole top view particles39. Next, a subset of 200 particles was selected and manually filtered into top, side or intermediate views (i.e. the response). The shape descriptors and the results of the manual sorting were copied into a data table that was used as a training dataset to generate models using supervised machine learning (Supplementary Figure 11). We used MATLAB's Classification Learner to identify the best model for predicting the classified outcome (response) based on the shape parameters. The best model was subsequently saved and could later be applied to other datasets. We found consistently good performance with support vector machine models, which are now also implemented into SPARTAN (Particles>Manual Classifier and Train SVM Classifier). The model's ability to identify a certain shape was in general more accurate for top (xy) orientations, requiring little manual selection/filtering. Importantly, only one of the two imaged centriolar proteins (i.e. the reference) needs to be classified onto top/side view.

All of the following operations are then performed on both channel datasets. Notably, while we used all particles for the determination of Δz, only a visually filtered subset was used for the investigation of the nine-fold symmetry. Many particles suffer from over- or underlabelling and were thus not considered. The identified side-view particles were registered to the center of mass of the reference protein and aligned using an extended version of efficient subpixel registration by cross-correlation40 as described previously41 (see also Supplementary Note 2). Specifically, during the first iteration, we rotated each image from 1 – 359° in 3° steps, resulting in 120 cross-correlations, from which we picked the orientation with the maximum root mean square (RMS) error, giving the optimal angle of rotation. The sum of all images was used as a reference for the first iteration. The alignment was refined over 3-10 iterations, using the sum of all aligned particles from the previous iteration as the reference (Supplementary Fig. 13). The translation between both channels along the z axis was determined using a line profile measurement of the two-color reconstruction (Fig. 1d, Supplementary Fig. 14. To generate a final multi-color volumetric representation, the co-oriented volumes were loaded into Chimera38 and centered on top of each other. The Δz axial transformation was applied using the transform coordinates tool (Tools > Movement > Transform coordinates). The volume obtained from the lower resolution SMLM channel (i.e. DyLight755 channel, Supplementary Fig. 5) was then replaced by a higher resolution volume of the same structure (taken in the Alexa647 channel). To this end, the high resolution volume was loaded into Chimera and aligned to the low resolution volume using the 'Fit in Map' tool (Tools > Volume Data > Fit in Map).

SMLM Simulations

To evaluate the contribution of labelling noise and efficiency as well as to test the particle processing workflow, we developed a particle simulator that generates localization maps from ground truth models. To define a ground truth model, we used the geometric dimensions of the complex as obtained from SMLM. The ground truth model was then randomly rotated and projected onto the XY plane. A random number of molecules were selected according to the labeling efficiency and a defined number of noise molecules placed at random positions around each particle. Localizations (single frame appearances) originating from each fluorophore were assigned parameters drawn from measured

distributions for photon count, localization precision, and on- and off-time. The distributions were obtained from single molecule measurements of Alexa 647 performed under experimental conditions. The resulting simulated particles were analyzed as described for experimental SMLM datasets. All simulations were performed using custom-written Matlab code supplied as part of the supplementary software package.

Statistics and reproducibility

Figures show representative data from ≥3 (Figure 2, Supplementary Figs. 2, 10, 13, 14, 16) or 2 (Figure 1, Supplementary Fig. 8, 9, 15, 19, 20) representative experiments, or from single high-throughput experiments (Supplementary Figs. 9, 12, 14, 15). Supplementary Figures 8, 13, 15, 16 show representative data from two similar independent simulations.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank Caroline Lehmann for growth, purification and AFM characterization of T4 Bacteriophages, Jose Miguel de la Rosa Trevin for support regarding the optimal use of Scipion and Ciaran G. Morrison for providing the Cep164 antibody (1F3G10). We thank Bo Huang and Xiaoyu Shi for providing their 2D particle alignment code. We also thank Romain Laine and Clemens Kaminski for providing a HSV SMLM dataset. N.B. was supported initially by a grant from the European Research Council (ERC) to P.G. (AdG 340227), and then by the EPFL Fellows postdoctoral fellowship program funded by the European Union's Horizon 2020 Framework Programme for Research and Innovation (Grant agreement 665667, MSCA-COFUND). Research in S.M.'s lab is supported by the National Centre for Competence in Research (NCCR) Chemical Biology. We thank Michal Daszykowski for providing a public DBSCAN implementation.

References

1. Campbell MG, Veesler D, Cheng A, Potter CS, Carragher B. 2.8 Å resolution reconstruction of the Thermoplasma acidophilum 20S proteasome using cryo-electron microscopy. Elife. 2015; 4
2. Jiang J, Pentelute BL, Collier RJ, Zhou ZH. Atomic structure of anthrax protective antigen pore elucidates toxin translocation. Nature. 2015; 521:545–549. [PubMed: 25778700]
3. Byeon I-JL, et al. Structural convergence between Cryo-EM and NMR reveals intersubunit interactions critical for HIV-1 capsid function. Cell. 2009; 139:780–90. [PubMed: 19914170]
4. Beck M, Lučić V, Förster F, Baumeister W, Medalia O. Snapshots of nuclear pore complexes in action captured by cryo-electron tomography. Nature. 2007; 449:611–615. [PubMed: 17851530]
5. Strauss M, Schotte L, Karunatilaka KS, Filman DJ, Hogle JM, et al. Cryo-electron Microscopy Structures of Expanded Poliovirus with VHHs Sample the Conformational Repertoire of the Expanded State. J Virol. 2017; 91
6. Chang Y-W, et al. Architecture of the type IVa pilus machine. Science. 2016; 351:6278.
7. Chang Y-W, et al. Architecture of the Vibrio cholerae toxin-coregulated pilus machine revealed by electron cryotomography. Nat Microbiol. 2017; 2:16269. [PubMed: 28165453]
8. Szymborska A, et al. Nuclear pore scaffold structure analyzed by super-resolution microscopy and particle averaging. Science. 2013; 341:655–8. [PubMed: 23845946]
9. Salas D, et al. Angular reconstitution-based 3D reconstructions of nanomolecular structures from superresolution light-microscopy images. Proc Natl Acad Sci. 2017; 114:9273–9278. [PubMed: 28811371]
10. Douglass KM, Sieben C, Archetti A, Lambert A, Manley S. Super-resolution imaging of multiple cells by optimized flat-field epi-illumination. Nat Photonics. 2016; 10:705–708. [PubMed: 27818707]

11. Bornens M. The centrosome in cells and organisms. Science. 2012; 335:422–426. [PubMed: 22282802]
12. Bauer M, Cubizolles F, Schmidt A, Nigg Ea. Quantitative analysis of human centrosome architecture by targeted proteomics and fluorescence imaging. EMBO. 2016; 35:1–15.
13. Graser S, et al. Cep164, a novel centriole appendage protein required for primary cilium formation. J Cell Biol. 2007; 179:321–30. [PubMed: 17954613]
14. Kitagawa D, et al. Structural Basis of the 9-Fold Symmetry of Centrioles. Cell. 2011; 144:364–375. [PubMed: 21277013]
15. Gönczy P. Towards a molecular architecture of centriole assembly. Nat Rev Mol Cell Biol. 2012; 13:425–35. [PubMed: 22691849]
16. Ester M, Ester M, Kriegel H-P, Sander J, Xu X. A density-based algorithm for discovering clusters in large spatial databases with noise. Proc 2nd Internat Conf Knowl Discov Data Min; 1996. 226–231.
17. Sorzano COS, et al. A clustering approach to multireference alignment of single-particle projections in electron microscopy. J Struct Biol. 2010; 171:197–206. [PubMed: 20362059]
18. Scheres SHW, et al. Maximum-likelihood Multi-reference Refinement for Electron Microscopy Images. J Mol Biol. 2005; 348:139–149. [PubMed: 15808859]
19. Lukinavičius G, et al. Selective chemical crosslinking reveals a Cep57-Cep63-Cep152 centrosomal complex. Curr Biol. 2013; 23:265–270. [PubMed: 23333316]
20. Momotani K, Khromov AS, Miyake T, Stukenberg PT, Somlyo AV. Cep57, a multidomain protein with unique microtubule and centrosomal localization domains. Biochem J. 2008; 412:265–73. [PubMed: 18294141]
21. Sonnen KF, Schermelleh L, Leonhardt H, Nigg Ea. 3D-structured illumination microscopy provides novel insight into architecture of human centrosomes. Biol Open. 2012; 1:965–76. [PubMed: 23213374]
22. Banterle N, Gönczy P. Centriole Biogenesis: From Identifying the Characters to Understanding the Plot. Annu Rev Cell Dev Biol. 2017; 33:23–49. [PubMed: 28813178]
23. Loncarek J, Hergert P, Magidson V, Khodjakov A. Control of daughter centriole formation by the pericentriolar material. Nat Cell Biol. 2008; 10:322–328. [PubMed: 18297061]
24. Paoletti A, Moudjou M, Paintrand M, Salisbury JL, Bornens M. Most of centrin in animal cells is not centrosome-associated and centrosomal centrin is confined to the distal lumen of centrioles. J Cell Sci. 1996; 109:3089–102. [PubMed: 9004043]
25. Gogendeau D, Guichard P, Tassin A-M. Methods in cell biology. 2015; 129:171–189. [PubMed: 26175439]
26. Bornens M, Paintrand M, Berges J, Marty M-C, Karsenti E. Structural and chemical characterization of isolated centrosomes. Cell Motil Cytoskeleton. 1987; 8:238–249. [PubMed: 3690689]
27. Bourdin G, et al. Amplification and purification of T4-like escherichia coli phages for phage therapy: from laboratory to pilot scale. Appl Environ Microbiol. 2014; 80:1469–76. [PubMed: 24362424]
28. Edelstein A, , et al. Current Protocols in Molecular Biology. John Wiley & Sons, Inc.; 2010. 14.20.1–14.20.17.
29. Olivier N, Keller D, Gönczy P, Manley S. Resolution doubling in 3D-STORM imaging through improved buffers. PLoS One. 2013; 8:e69004. [PubMed: 23874848]
30. Huang F, et al. Video-rate nanoscopy using sCMOS camera–specific single-molecule localization algorithms. Nat Methods. 2013; 10:653–658. [PubMed: 23708387]
31. Churchman LS, Spudich JA. Colocalization of fluorescent probes: accurate and precise registration with nanometer resolution. Cold Spring Harb Protoc. 2012; 2012:141–9. [PubMed: 22301660]
32. Nieuwenhuizen RPJ, et al. Measuring image resolution in optical nanoscopy. Nat Methods. 2013; 10:557–62. [PubMed: 23624665]
33. Banterle N, Bui KH, Lemke Ea, Beck M. Fourier ring correlation as a resolution criterion for super-resolution microscopy. J Struct Biol. 2013; 183:363–367. [PubMed: 23684965]

34. de la Rosa-Trevin JM, et al. Scipion: A software framework toward integration, reproducibility and validation in 3D electron microscopy. J Struct Biol. 2016; 195:93–99. [PubMed: 27108186]
35. Sorzano COS, et al. XMIPP: a new generation of an open-source image processing package for electron microscopy. J Struct Biol. 2004; 148:194–204. [PubMed: 15477099]
36. Ludtke SJ, Baldwin PR, Chiu W. EMAN: Semiautomated Software for High-Resolution Single-Particle Reconstructions. J Struct Biol. 1999; 128:82–97. [PubMed: 10600563]
37. Scheres SHW. RELION: Implementation of a Bayesian approach to cryo-EM structure determination. J Struct Biol. 2012; 180:519–530. [PubMed: 23000701]
38. Pettersen EF, et al. UCSF Chimera - A visualization system for exploratory research and analysis. J Comput Chem. 2004; 25:1605–1612. [PubMed: 15264254]
39. Gartenmann L, et al. A combined 3D-SIM/SMLM approach allows centriole proteins to be localized with a precision of ~4–5 nm. Curr Biol. 2017; 27:R1054–R1055. [PubMed: 29017036]
40. Guizar-Sicairos M, Thurman ST, Fienup JR. Efficient subpixel image registration algorithms. Opt Lett. 2008; 33:156. [PubMed: 18197224]
41. Shi X, et al. Super-resolution microscopy reveals that disruption of ciliary transition-zone architecture causes Joubert syndrome. Nat Cell Biol. 2017; 19:1178–1188. [PubMed: 28846093]

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

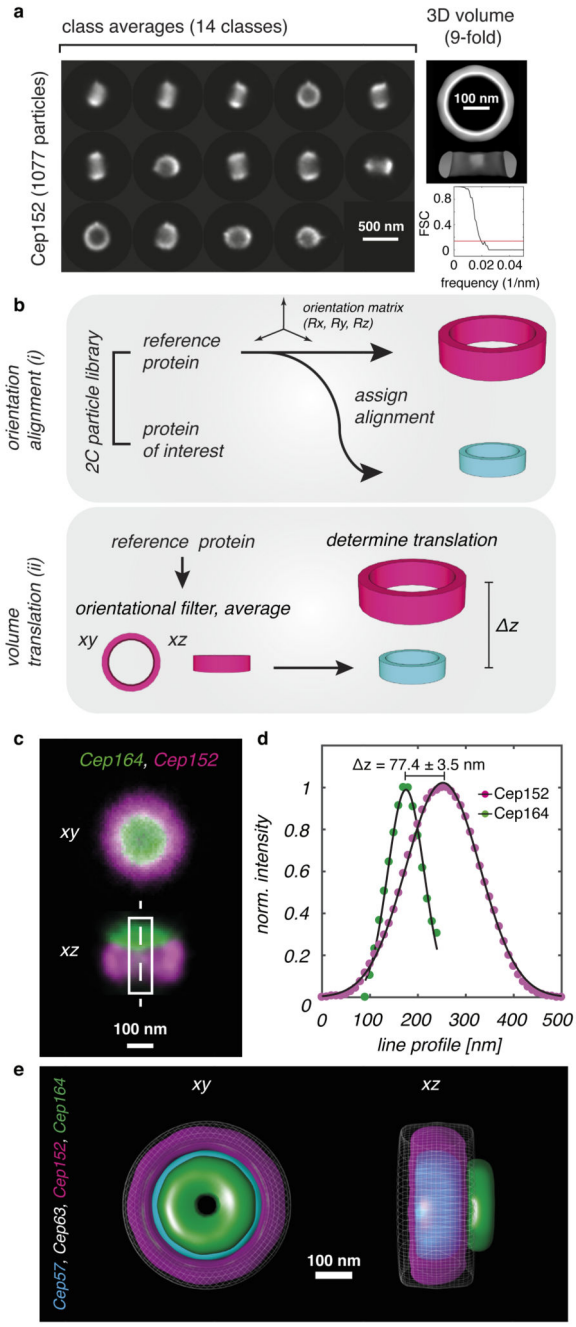

Figure 1. Multi-color single particle reconstruction.

(**a**) Purified human centrosomes immunolabelled against Cep152 were imaged using high-throughput SMLM. After filtering, 1077 particles were classified and used to reconstruct a 3D model (upper right) with FSC resolution of 52 nm (frequency = 0.019 nm^{-1}, lower right). (**b**) Multi-volume alignment from two-color (2C) particles: (i) orientational alignment by reconstructing the reference protein (magenta) and assigning single particle orientations to the protein of interest (cyan) and (ii) volume translation from side views. (**c**) Shown for Cep152/Cep164, top (xy) and side (xz) views are identified; (**d**) and fit to Gaussian

distributions to determine Δz. (**e**) Reconstruction of the protein pairs Cep152/Cep164, Cep152/Cep57 and Cep57/Cep63 yields a four-color map of the human centriole. Data in (**c**, **e**) are from two independent experiments showing similar results. Data in **d** show the mean and SD from the Gaussian fitting of averaged 2C projections (**c**, n = 102).

Europe PMC Funders Author Manuscripts

Europe PMC Funders Author Manuscripts

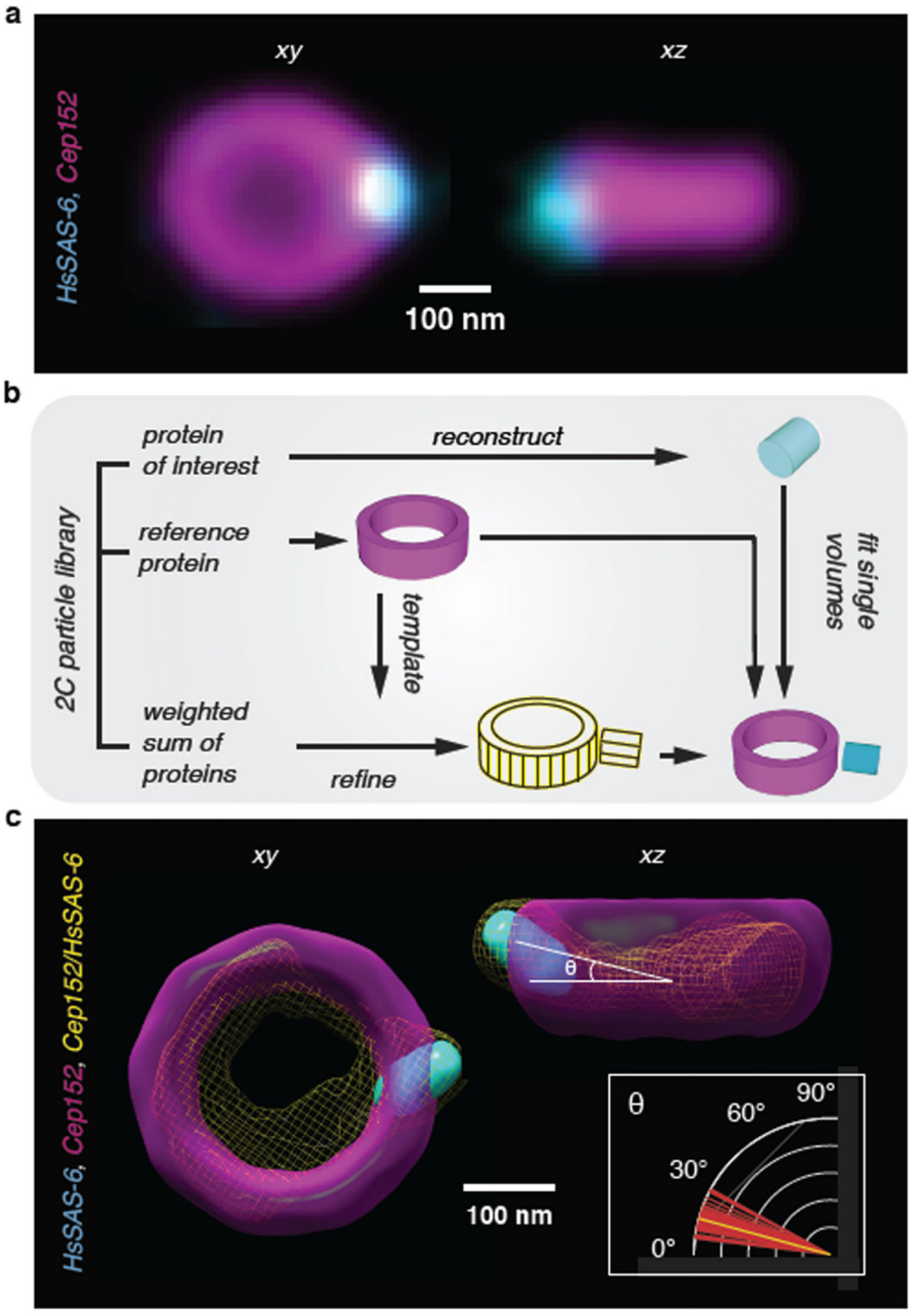

Figure 2. Multi-color single particle reconstruction of an asymmetric protein complex.

(**a**) Top (xy, n = 61) and side (xz, n = 75) view averages for HsSAS-6 (cyan) and Cep152 (magenta) display a protruding, off-axis HsSAS-6 density. For xz, to align the HsSAS-6 densities to a single focus, we selected centriole side views with the HsSAS-6 density at the same position relative to the Cep152 torus (i.e. lower left side). (**b**) For 3D reconstruction of an asymmetric assembly, the reference protein (magenta) and the protein of interest (cyan) are reconstructed separately, then the joint density (yellow mesh) is reconstructed using the weighted sum of the individual channels. The individual volumes are fit into the joint

density map to obtain the two-color reconstruction. (**c**) Reconstruction of Cep152/HsSAS-6. The individual volumes of Cep152 (magenta) and HsSAS-6 (cyan) were fit into their joint density map (yellow mesh). Inset: the average orientation of HsSAS-6 arises from a broad distribution of individual particle orientations (red, individual particles; yellow, average angle (θ = 15.4 ± 4.5° (mean ± SD), n = 75)).

Europe PMC Funders Author Manuscripts