Towards the molecular architecture of the peroxisomal receptor docking complex

Supplementary Information

Strains and culture conditions

Saccharomyces cerevisiae strains used in this study are listed in Supplementary Table I. Strains in which the genomic copy of *PEX17aa1-167* contains epitope tags were generated as described by(1) using primers KU1008/KU1075 and KU1005/KU1006 for PCR reactions, respectively. Strains expressing the Pex14p mutants M51L and Δ1-50 were generated by genomic integration of PCR products obtained from the plasmids pAS415 (primers: O674/O675) and pIS29 (O674/O899), respectively, into the strain SC260. Transformants were selected for the appropriate marker and proper integration was confirmed by PCR. Yeast complete (YPD) and minimal media (SD) have been described previously(2). YNO medium contained 0.1% oleic acid, 0.05% Tween 40, 0.1% yeast extract and 0.67% yeast nitrogen base without amino acids, adjusted to pH 6.0. When necessary, auxotrophic requirements were added according to(3).

Oligonucleotides and Plasmids

Plasmids and oligonucleotides used in this study are listed in Supplementary Tables II and III. For two-hybrid analysis, DNA fragments of *PEX14* were fused to the DNA-binding domain of *GAL4* in the vector pPC86(4). To do so, corresponding *PEX14* fragments were amplified by PCR using pSK*PEX14*(5) as a template. The PCR fragments were digested with *EcoRI/BamHI* and cloned into *EcoRI/BglII*-prepared pPC86. Yeast expression vector with expression of Pex17p$_{1\text{-}167}$ was generated as follows. The cytochrome C oxidase termination (cyc) was amplified by PCR with primers KU171/KU288 and pUG35 as template. PCR fragments were digested with HincII and introduced into HincII-site directly behind the coding triplet coding for aa167 of plasmid pSK9/1.4 harbouring *PEX17*-coding region plus 5`- and 3`-non-coding segments(6). The fusion gene *PEX17*-Promotor- PEX17aa1-167-cyc-termination was cut out *SacI/KpnI* and introduced into same sites of yeast vector pRS416 resulting into pRS416-PEX17I. To generate yeast expression vector with expression of Pex17p$_{1\text{-}125}$, the *PEX17* promotor plus coding region for aa1-125 as well as the PEX17 termination were amplified by separated PCRs with primers KU133/KU172 and KU623/KU624, respectively. The PCR fragments were digested *EcoRI/SacI* and *SacI/HindIII*, respectively, and introduced both together in the *EcoRI/HindIII* sites od plasmid pRS416 leading to plasmid pRS416- PEX17II. pTSC19 which was created as follows. PEX17 was amplified by PCR with primers RE6399/RE6402 and plasmid pSK9/1.4(6) as template. The PCR-product was introduced into the *NdeI/XhoI* site of pCOLA-Duet leading to pCOLA-Duet-His$_6$-Pex17p. PEX14 coding region was amplified by PCR from plasmid pSK*PEX14* (5) with primers RE 5277/RE5278, which also introduce the Strep$_{II}$-tag. PCR product was introduced into *NcoI/EcoRI* site of pCOLA-Duet-His$_6$-Pex17p which gives plasmid Strep$_{II}$-Pex14p/His$_6$-Pex17p. Finally, coding triplet of methionine 51 was change to leucine by by QuikChange site-directed mutagenesis (RE6138/RE6139) according to the manufacturer's instructions (Stratagene, Amsterdam, Netherlands) leading to Strep$_{II}$-Pex14p(M51L)/His$_6$-Pex17p. Same was done for pCOLA-Duet-His$_6$-Pex14p/Pex17p to generate pCOLA-Duet-His$_6$-Pex14p(M51L)/Pex17p and for the yeast vector pRS316-PEX14 (pIS29) resulting in pRS316-PEX14(M51L) (pAS415).

Yeast cell extracts

Yeast cells were grown on 0.3% SD medium to late log phase and subsequently for 15 h in YNOD (0.1% dextrose, 0.1% oleic acid, 0.05% Tween 40, 0.1% Yeast extract and 0.67% yeast nitrogen base). Cells were harvested and aliquots of 30 mg of cells were resuspended in 300 µl potassium-phosphate buffer (pH 7.4) containing 20% trichloroacetic acid. The samples were frozen at –80°C for at least 30 min. Samples were sedimented, washed twice with ice-cold 50 % acetone and resuspended in 80 µl 10% SDS / 0.1 M NaOH and 20 µl SDS loading buffer (5% ß-mercaptoethanol, 15% glycerol, 0.01% bromphenol blue).

Two-hybrid analyses

The two-hybrid assay was based on the method of(7). Double transformants of PCY2*pex14*Δ cells expressing the GAL4-fusion proteins were selected, and β-galactosidase activity was determined by performing a filter assay using X-Gal as the substrate, as previously described(8). The quantification of ß-galactosidase activity was performed according to(9) .

Protein purification

Expression of recombinant proteins ($Strep_{II}$-Pex14p(M51L)/His_6-Pex17p, His_6-Pex14p(M51L)/Pex17p and Pex14p variants T7-Pex14p$_{1-95}$-His_6, T7-Pex14p$_{119-341}$-His_6, T7-Pex14p$_{213-341}$-His_6, Pex14p$_{WT}$-His_6, Pex14p$_{M51L}$-His_6) was performed in *E. coli* BL21-CodonPlus(DE3)-RIL cells (Agilent, Waldbronn, Germany). Cells were inoculated at an OD600nm of 0.1 and expression of the recombinant proteins was induced by adding 0.4 mM IPTG when reaching an OD600nm of 0.5 to 0.6. Time of expression was 4 h at 20 °C. Cells were harvested, resuspended in 10-fold-volume of binding buffer I (20 mM Tris, 500 mM NaCl, 1% (w/v) n-Dodecyl β-D-maltoside (DDM), pH 7.9). Buffers contained protease inhibitors (8 mM antipain, 0.3 mM aprotinin, 0.16 mg/ml benzamidin, 1 mM bestatin, 10 mM chymostatin, 5 mM leupeptin, 15 mM pepstatin, 1 mM PMSF, 0.21 mg/ml NaF) and 25 µg/ml DNAse I. Cells were broken using Emulsi Flex and obtained lysates were clarified by centrifugation at 14.000 rpm for 1 h (rotor SS-34, Thermo Scientific, Schwerte, Germany). The supernatant contained $Strep_{II}$-Pex14p(M51L) and His_6-Pex17p was loaded onto a Strep-Tactin-column (IBA-Lifescience, Goettingen, Germany). After washing with buffer II (20 mM Tris, 500 mM NaCl, 0.1% (w/v) DDM, pH 7.9), proteins were eluted with buffer III including the specific competitor desthiobiotin. The obtained eluate was subsequently loaded onto a Ni-NTA column (Protino Ni-NTA-Agarose, Macherey-Nagel, Düren, Germany). After washing with buffer II, proteins were eluted with buffer IV (500 mM Imidazole, 50 mM Tris, 150 mM NaCl, 0,1% (w/v) DDM, pH 8.0).

Gel filtration was performed on a Superose 6 10/300 GL column (GE Healthcare, Munich) equilibrated with TBS (50 mM Tris, 150 mM NaCl, 0,1% (w/v) DDM, pH 8.0) at 4 °C using the ÄKTApurifier system (GE Healthcare, Munich). 500 µl of Ni-NTA column eluate fractions were loaded and 500 µl fractions were collected. HMW-Gel Filtration Calibration Kit (GE Healthcare, Freiburg, Germany) contained thyreoglobulin (669 kDa), ferritin (440 kDa) and aldolase (158 kDa), ovalbumin (43 kDa) and carbonic anhydrase (29 kDa). Marker proteins were obtained from Pharmacia (HMW-marker for native electrophoresis).

For purification of Pex14p variants, cell pellets were resuspended in 10x volume of lysis buffer 1 by vortexing. 0.1% (v/v) Tween 20, 1 mM PMSF and 0.2 mg/ml lysozyme were added, followed by 1 h incubation on ice. To degrade nucleic acids, 50 units/10 ml cell suspension nuclease and 10 mM $MgCl_2$ were added, followed by further 15 min incubation. Cells were disrupted by sonication for 10 min at 90% and cycle 5x 10% (Bandelin electronic, Berlin, Germany). Subsequently, cell debris was pelleted by centrifugation (40 min, 4°C, 25.000 g) and the supernatant was filtered with a 0.22 µm bottle-top filter by vacuum for further affinity purification. To avoid unspecific binding to the Ni-NTA column (Histrap HP, 1 ml, GE Healthcare, Munich, Germany) on a Äkta start system (GE Healthcare, Munich, Germany) 20 mM imidazole were added to the lysate before the sample was applied to the column, which

had been equilibrated with IMAC buffer A (50 mM HEPES, 500 mM NaCl, 0.1% (v/v) Tween 20, 20 mM imidazole, pH 7.5). Unspecific binding was further reduced by washing the column with 6 column volumes (CVs) of IMAC buffer A. Elution of His_6 tagged proteins was performed by 5 CVs of IMAC buffer B (50 mM HEPES, 500 mM NaCl, 0.1% (v/v) Tween 20, 500 mM imidazole, pH 7.5). For further purification, peak fractions were pooled and loaded onto a Superdex 16/600 200pg size exclusion column (GE Healthcare, Munich, Germany) on a Äkta pure system (GE Healthcare, Munich, Germany) equilibrated with SEC buffer (50 mM HEPES, 200 mM NaCl, 10% (v/v) glycerine, pH 7.5). To identify fractions of interest, sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS PAGE) with subsequent colloidal Coomassie staining was performed using standard techniques. Fractions of interest were snap-frozen in liquid nitrogen and stored at -80°C for further analysis.

Antibodies and Immunoblotting

Immunoblot analysis was performed according to Harlow and Lane (Harlow and Lane, 1988) with polyclonal rabbit antibodies raised against, Pex5p(5), Pex7p(10), Pex10p(11), Pex12p(12), Pex13p(13) Pex14p(5), and Pex17p(6). Primary antibodies were detected with a IRDye 800CW goat anti-rabbit IgG secondary antibody (LI-COR Bioscience, Bad Homburg, Germany) followed by a detection using the “Infrared Imaging System“(LI-COR Bioscience, Bad Homburg, Germany). Semi-quantitative analyses of immunoblot signals were obtained using the “Infrared Imaging System Application Software Version 3.0” (LI-COR Bioscience, Bad Homburg, Germany).

Miscellaneous Methods

Immunopurification of native complexes using hsIgG-coupled sepharose was performed as described(11).

SEC-MALS of $Strep_{II}$-Pex14p(M51L)/His_6-Pex17p and $Strep_{II}$-Pex14p(M51L) in 0.1% DDM

For SEC-MALS analysis the complex was freshly purified and applied to a Superose 6 increase 10/300 column connected to an Agilent InfinityII HPLC system at a flow rate of 0.5 ml/min. Data collection was performed using a Wyatt System equipped with the 8 detector Dawn Heleos Neon, Optilab Neon RI detector, generic Agilent UV detector (280nm), Agilent Autosampler and an Agilent column oven at 8°C.Teh sample was measured in TBS buffer (50 mM Tris pH 8.0, 150 mM NaCl and 0.1% (w/v) DDM). 80 μl of 2.12 mg/ml protein were loaded to the SEC column after extensive equilibration with buffer. The data was analyzed using Astra 7.3.1.9. Detector normalization was done with BSA at 2 mg/ml in the same buffer and those values were applied to the actual measurement. Protein concentration was calculated using the dRI signal. dn/dc values for the protein were calculated with sedfit(14), dn/dc for DDM was taken from publication(15). Protein concentration and detergent concentration in the protein-detergent micelle were calculated using the “protein conjugate” analysis of Astra with calculated dn/dc and extinction coefficient for the protein and dn/dc=0.133 (extinction coefficient = 0) for DDM.

Negative Stain EM

4 μl of the various Pex14p/Pex17p complexes as well as the $Pex14p_{114-341}$ complex in TBS Buffer (50 mM Tris pH 8.0, 150 mM NaCl (0.1% (w/v) DDM)) was applied for 2 minutes on a freshly glow discharged carbon coated copper-grid (Agar scientific; G2400C) at room temperature. The sample was blotted with Whatman paper (No 5.), the grid was washed two times using 10 μl detergent free buffer containing 50mM Tris (pH 8) and NaCl 150mM. Subsequently after washing, a 10 μl drop of freshly prepared uranyl formate was applied on the grid for 30 seconds and afterwards air-dried. Images were recorded with a JEM-1400 (JEOL) or a FEI Tecnai G2 Spirit (FEI) both equipped with LaB_6 cathode and a 4K CMOS detector F416 (TVIPS). Single particles were boxed using cryolo(16). Reference-free stable class-averages were computed using the ISAC approach of the SPHIRE software package(17).

Reconstitution into POPC Liposomes

POPC Avanti lipids (Alabaster, #850457P) were solubilized by sonication in 20 mM Tris and 100 μM Cholate to a final concentration of 50 mM. 2 μM of Strep$_{II}$-Pex14p(M51L)/His$_6$-Pex17p complex were added to 300-fold molar excess of POPC lipids. Protein sample and lipids were incubated 2h in TBS buffer, containing 50mM Tris (pH 8),150mM NaCl and 0.1% DDM. The detergent was removed by absorption to Bio-Beads (Bio-Rad), 6 hours at 4°C. Samples were then adsorbed to EM grids and analyzed by negative stain EM and cryoEM (see below). Negative stain EM micrographs were recorded on a Tecnai Spirit G2 (FEI) at a nominal magnification of 42,000x and a corresponding pixel size of 2.6Å.

Sample optimization for cryoEM; Reconstitution in amphipols and MSP1D1dH5, cMSP1D1dH4-6 nanodiscs

To optimize the sample for EM studies and increase contrast, we first replaced DDM with amphipol A8-35(18). The mixture containing the Strep$_{II}$-Pex14p(M51L)/His$_6$-Pex17p complex and amphipols (A8-35 (Anatrace#A835)) (five times molar excess of amphipols to the total protein mass) in TBS Buffer (50 mM Tris pH 8.0, 150 mM NaCl and 0.1% (w/v) DDM) was therefore incubated for 4 hours at 4°C. The detergent was removed by dialysis in 1L 50mM Tris (pH 8) and 150 mM NaCl, 16 hours at 4°C. The protein was afterwards concentrated using a 10 MWCO Amicon centrifugal filter unit (Millipore, #UFC5010) and further purified by size exclusion chromatography using a Superose 6 increase 5/150 GL column (GE Healthcare, #29091597) equilibrated in 50 mM Tris (pH 8.0),150 mM NaCl and 1 mM TECP and subsequently negatively stained at a concentration of 0.003mg/ml and visualized by EM, using a JEM-1400 (JEOL) at a nominal magnification of 60.000x and a pixel size of 1.84Å. The tight adsorption of amphipols to the hydrophobic transmembrane domains of Strep$_{II}$-Pex14p(M51L)/His$_6$-Pex17p surprisingly resulted in dissociation of the complex and separation of Pex14p and Pex17p (Supplementary Figure 14).

To overcome this issue, we exploited the possibility to reconstitute the Strep$_{II}$-Pex14p(M51L)/His$_6$-Pex17p complex into lipid nanodiscs. For nanodisc reconstitution, we first used the membrane scaffold protein (MSP)1D1dH5 with a rather narrow predicted scaffold diameter of 8-9nm. Strep$_{II}$-Pex14p(M51L)/His$_6$-Pex17p complex in TBS Buffer (50 mM Tris pH 8.0, 150 mM NaCl and 0.1% (w/v) DDM) was reconstituted in commercially available nanodiscs MSP1D1(CubeBiotech, #26132) and MSP1D1-His$_6$-Tag (CubeBiotech, #26112). First, the protein sample was concentrated using a 30 MWCO Amicon centrifugal unit (Millipore) to a final concentration of 3 mg/ml and afterwards incubated with nanodiscs in a ratio 1:50. The sample was then dialyzed against detergent-free buffer (50mM Tris (pH 8) and 150 mM NaCl) in order to remove the detergent and induce the reconstitution into the nanodiscs. Size-exclusion chromatography using a Superose 6 increase 5/150 GL (GE Healthcare, #29091597) column was performed to separate the reconstituted protein complex from aggregates and/or empty nanodiscs. Sample quality was screened by negative stain EM using a JEM-1400 (JEOL) at a nominal magnification of 60,000x and a corresponding pixel size of 1.84Å. Despite extensive optimization efforts, the resulting proteo-nanodiscs were heterogeneous and incorporated multiple copies of the Pex14p/Pex17p complex, usually from both sides (Supplementary Figure 3b).

We then repeated the reconstitution experiments using the recently reported circular MSP1D1Δ4-6 nanodiscs, which are according to our knowledge the smallest available nanodisc to date(19). Strep$_{II}$-Pex14p(M51L)/His$_6$-Pex17p complex in TBS Buffer (50 mM Tris pH 8.0, 150 mM NaCl and 0.1% (w/v) DDM) was reconstituted in commercially available nanodiscs MSP1D1(CubeBiotech, #26132) and MSP1D1-His$_6$-Tag (CubeBiotech, #26112). First, the protein sample was concentrated using a 30 MWCO Amicon centrifugal unit (Millipore) to a final concentration of 3 mg/ml and afterwards incubated with nanodiscs in a ratio 1:50. The sample was then dialyzed against detergent-free buffer (50 mM Tris (pH 8) and 150 mM NaCl) in order to remove the detergent and induce the reconstitution into the nanodiscs. Size-exclusion chromatography using a Superose 6 increase 5/150 GL (GE

Healthcare, #29091597) column was performed to separate the reconstituted protein complex from aggregates and/or empty nanodiscs. Sample quality was screened by negative stain EM using a JEM-1400 (JEOL) at a nominal magnification of 60,000x and a corresponding pixel size of 1.84Å. Indeed, the negative stain EM analysis revealed high-contrast monodispersed particles, with most nanodiscs trapping only a single copy of the complex (Supplementary Figure 3c).

We reconstituted Pex14p$_{M51L}$-His$_6$ with similar experimental set up into cMS46 nanodisc as described before.

Gold labeling

Purified Strep$_{II}$-Pex14p(M51L)/His$_6$-Pex17p and His$_6$-Pex14p(M51L)/Pex17p complex was reconstituted in commercially available MSP1D1(CubeBiotech, #26132) nanodiscs as described above. Afterwards 10nm Ni-NTA gold particles (Nanoprobes;2082-3 mL) were added in fivefold excess and incubated for 30 minutes on ice and subsequently visualized by negative stain EM. We analyzed a total of 500 His$_6$-Pex14p(M51L)/Pex17p and Strep$_{II}$-Pex14p(M51L)/His$_6$-Pex17p gold-labeled particles, respectively.

Sample Vitrification and Cryo-EM Data Acquisition

For cryo-EM, 4 μl of Strep$_{II}$-Pex14p(M51L)/His$_6$-Pex17p complex reconstituted in cMSP46 at an overall concentration of 0.15 mg/ml was applied to freshly glow discharged holey carbon Quantifoil 2/1 (#EMS300-Cu) grids. After 30 sec incubation on the grids, manual backside blotting was performed using Whatman paper (No. 5). 4 μl of the sample at the same concentration was then directly applied to the same grid, which was then immediately blotted and plunged into liquid ethane cooled by liquid nitrogen using a FEI Vitrobot Mark II system, 100% humidity, blotting time of 3.5 seconds at 4°C.

A dataset was collected with a Cs-corrected Titan Krios EM (Thermo Fisher) equipped with a K2 direct electron detector (Gatan), a quantum energy filter (Gatan) and a Volta Phase plate (Thermo Fisher). Movies were recorded in counting mode using the automated acquisition program EPU (Thermo Fisher) at a magnification of 130,000x corresponding to a pixel size of 1.09Å. 3282 movies were acquired in a defocus range of 0.7 to 1 μm. Each movie comprised 40 frames acquired over 15 seconds with a total dose of ~61e^-/Å^2.

The single CC dataset was collected with a Talos Arctica (Thermo Fisher) equipped with a Falcon III direct electron detector (Gatan). Movies were recorded in counting mode using the automated acquisition program EPU (Thermo Fisher) at a magnification of 120,000x corresponding to a pixel size of 1.21Å. 971 movies were acquired in a defocus range of 1.8 to 3.5 μm. Each movie comprised 40 frames acquired over 15 seconds with a total dose of ~56e^-/Å^2.

Image processing and reconstruction

Monitoring of image collection as well as image pre-processing was performed using the TranSPHIRE software package (https://github.com/MPI-Dortmund/transphire). Motion correction was thereby performed using MotionCor2(20) and the contrast transfer function was estimated using CTER(21). 10 representative images were then manually picked using the crYOLO-boxmanager tool, in order to train crYOLO(16) for further automated particle picking. A total of 361,862 particles were finally selected and extracted using a box size of 300 pixels. SPHIRE (1.2,1.3)(17) was used for further single particle analysis steps. Multiple reference-free 2D classification rounds were then using the stable alignment and cluster approach (ISAC)(22) in SPHIRE (Supplementary Figure 8). After each round of 2D classification, we carefully selected classes containing intact complexes in homogeneous nanodiscs. The resulting class averages were then sent to VIPER, to generate a low-resolution initial model. The respective class members (a set of~142.000 particles) were then subjected to 3D refinement with no symmetry imposed using the VIPER reconstruction as initial reference. The resulting projection parameters were used to re-center and re-extract the particles. The re-centered particles were again subjected to multiple rounds of 2D

classification and finally, ~82.600 particles with best-resolved features were selected for auto-refinement, which then produced a 10Å reconstruction. The resulting density map was used to create a soft-edge gaussian mask including the nanodisc density, which was used for signal subtraction of nanodisc projections from the experimental images. The resulting particle stack and a gaussian mask including only the rod (Supplementary Figure 8) were used as input for a local refinement, which further improved local resolution. Further 3D clustering did not result into distinct, reproducible 3D groups. We further subtracted projections of the rod from the experimental images and performed auto-refinements and local refinement focusing on the nanodisc alone, which however failed to converge. Resolution calculation, 3D masking, sharpening and low-pass filtering were automatically performed using the PostRefiner tool (SPHIRE).

CryoET

For cryo-ET, the $Strep_{II}$-Pex14p(M51L)/His_6-Pex17p protein complex was reconstituted in liposomes, as described above and mixed in a ratio of 20:1 with 5nm colloidal gold particles. 4 μl was applied on freshly glow discharged holey carbon Quantifoil 2/1 (#EMS300-Cu) grids and plunged into liquid ethane cooled by liquid nitrogen using a FEI Vitrobot Mark II system (3.5 seconds blotting time, 4°C and 100% humidity). Tomography was then performed on a Titan Krios G2 operated at 300kV, Volta Phase plate (Thermo Fisher), a GIF post-column energy filter (Gatan) and K2 Summit camera (Gatan).

Tilt-series were collected using SerialEM(23) under low dose conditions (total exposure of about 100$e^-/Å^2$), a defocus of -1 μM, covering the range of 40° in 3° increments. Images were reduced by 4 x 4-pixel averaging resulting in a pixel size of 7Å. Data were processed using the IMOD software package(24). Gold particles were tracked as fiducial markers to align the stack of the tilted images and tomograms were reconstructed by weighted back-projection. A selected sub-tomogram was manually segmented and rendering was performed in Chimera(25). The segmentation of the reconstruction was performed by manually tracing of distinct structural features though sequential slices of the tomogram.

Bioinformatic tools

Electron density maps and models were visualized using Chimera and ChimeraX(26). Map segmentation was performed using the module Segment Map of the Chimera software package. Secondary structure prediction analysis was carried out using the SABLE protein structure prediction server (http://sable.cchmc.org). Structure modeling of Pex14p was performed using the RaptorX web server (27). Coiled-coil domains in the sequences of Pex14p and Pex17p were predicted using Marcoil(28).

Chemical cross-linking

$Strep_{II}$-Pex14p/His_6-Pex17p complexes were dialyzed against 50 mM HEPES/150 mM NaCl/0.03% DDM (pH 8.0) and bis[sulfosuccinimidyl]suberate (BS3; Merck, Darmstadt, Germany) was added to a final concentration of 0.25 mM. The mixture was further incubated for 1 h at 16°C and under agitation at 1,100 rpm. The reaction was quenched by adding 1 M Tris buffer (pH 7.5) to a final concentration of 50 mM. For cross-linking of Pex14p variants, purified proteins were buffer exchanged in 20 mM HEPES/100 mM NaCl (pH 7.4) using Amicon Ultra 2 ml centrifugal filter units (Merck, Darmstadt, Germany) with a MWCO of 3 kDa and diluted to 0.2 mg/ml for cross-linking reactions. BS3 cross-linker was dissolved in ddH2O and directly added to the protein solution. The reaction was performed at 25°C, 300 rpm for 30 min and quenched by adding 50 mM Tris/HCl pH 7.8. Cross-linked proteins were separated by SDS-PAGE and stained with colloidal Coomassie Blue G-250.

Proteolytic digestion

In-solution tryptic digestion of proteins was performed as described previously(29). For in-gel protein digestion, gel slices of interest were excised and destained by three repetitions of incubation in 10 mM ammonium bicarbonate (AmBic) followed by incubation with 5 mM AmBic, 50% (v/v) ethanol), each for 10 min at room temperature. Prior to the digestion of cysteine-containing proteins, cysteine residues were reduced by incubation with 5 mM Tris(2-carboxy-ethyl)phosphine/10 mM AmBic for 10 min at 60°C and subsequently alkylated by incubation with 50 mM iodoacetamide/10 mM AmBic for 15 min at 37°C in the dark. For cysteine-free proteins, no reduction and alkylation step were conducted. Gel pieces were dehydrated and washed by incubation with 100% (v/v) ethanol (10 min at room temperature) and then dried in a vacuum centrifuge (Christ, Osterode, Germany). Proteins were digested in the gel matrix o/n at 37°C with a trypsin (Promega, Madison, USA) to protein mass ratio of 1:25 in 10 mM AmBic. Peptides were eluted twice by adding 0.05% TFA, 50% acetonitrile and sonication for 10 min. Combined eluates were dried in a vacuum centrifuge, resuspended in 0.1% TFA and desalted as described before(30). For subsequent liquid chromatography-tandem mass spectrometry (LC-MS/MS) analysis peptides were resuspended in 0.1% TFA by 4 min sonication and precipitates were removed by pelleting for 6 min at 12.000 rpm.

LC-MS analyses

Peptide mixtures were separated on a UltiMate™ RSLCnano system (Thermo scientific, Dreieich, Germany) (precolumn: PepMap™ 0.3 mm / 5 mm, C18; analytical column: Acclaim PepMap™ 75 μm/500 mm, C18, 100 Å pore size) using binary solvent systems consisting of 0.1% (v/v) formic acid (solvent A) and 86% (v/v) acetonitrile, 0.1% (v/v) formic acid (solvent B) with a flow rate of 0.3 μl/min. Peptides derived from in-gel tryptic digestions of proteins were separated with a gradient of 4 to 42% B in 35 min, followed by 42 to 95% B in 3 min, and 95% B for 5 min. Cross-linked peptides derived from in-solution digestion of proteins were separated with a gradient of 4 to 40% B in 50 min, followed by 40 to 95% B in 5 min, and 95% B for 5 min. Eluted peptides were measured online on a Q Exactive Plus mass spectrometer (Thermo scientific, Dreieich, Germany) operated in a data-dependent acquisition mode. MS^1 scans were acquired in a scan range of *m/z* 375 to 1,700 at a resolution of 70,000 (at *m/z* 400) with an automatic gain control (AGC) of 3×10^6 ions and a maximum fill time of 60 ms. The 12 most intense multiple charged precursor ions (TOP12 method) were selected for higher-energy collisional dissociation (HCD). MS^2 data were acquired in a scan range of *m/z* 200 to 2000 at a resolution of 35,000, AGC of 1×10^5 ions and a maximum fill time of 120 ms. Peptides derived from $Pex14p_{WT}$-His_6 and $Pex14p_{M51L}$-His_6 (see Supplementary Figure 1e) were separated using a binary solvent systems consisting of 0.1% (v/v) formic acid 4% (v/v) DMSO (solvent A) and 50% (v/v) methanol/30% (v/v) acetonitrile/4% (v/v) DMSO/0,1% (v/v) formic acid, applying a gradient of 1 to 65% B in 30 min, 65 to 95% B in 5 min, and 95% B for 5 min, and analyzed on a LTQ-Orbitrap XL (Thermo Scientific, Dreieich, Germany) operated with the following parameters: scan range of *m/z* 370 to 1,700, resolution of 60,000 (at *m/z* 400), AGC of 5×10^5 maximum fill time of 500 ms, TOP5 method with collision-induced dissociation (CID) in the linear ion trap and acquisition of MS^2 spectra with an AGC of 1×10^4 ions and a maximum fill time of 100 ms.

Crosslinking MS data analysis

For identification of cross-linked peptides, the software pLink2 (version 2.3.5; (31)) was used to search against the forward and reversed amino acid sequences of the recombinant proteins. For data from $Pex14p_{119\text{-}341}$ variant, a mass tolerance of 20 ppm for MS^1 and MS^2, a maximum of 3 missed cleavages for trypsin, peptide length of 5-60, oxidation of methionine as variable and carbamidomethylation of cysteine as fixed modification, Lys and the protein N-terminus as cross-link sites were used as parameters. For filtering of search results, a MS^1 tolerance of 10 ppm and an initial false discovery rate (FDR) at peptide-spectrum match (PSM) level of 5% were applied. The resulting set of cross-linked peptide spectrum matches was filtered to remove possible false identifications. Matches having

E-values above 5×10^{-2} or comprising adjacent tryptic peptides of the same protein were rejected. After deriving intensity ratios as described below, only peptide matches with at least one identification at a q-value (FDR) of 0.01 and being identified in at least 2 out of 3 replicate experiments were retained. For data from Pex14p/Pex17p complexes, the same settings were used with the following modifications: minimum peptide length was restricted to 6 and Ser, Thr, Tyr were additionally considered as potential cross-link sites. As no cross-link pairs involving Tyr residues were identified it was omitted in the final search. FDR filter at PSM level was 1% and, for further filtering, all peptide spectrum matches were manually checked using pLabel (version 2.3.5(32)) and matches that did not follow a pattern of fragment ions including for both peptides at least a series of two consecutive fragment ions were rejected. Only cross-linked residue pairs being identified in at least 3 out of 4 biological replicates are reported. For graphical representation of cross-linked residue pairs, the software xiNet (version 2.0(33)) was used. To compare the abundances of cross-linked peptides recovered from gel bands of protein oligomers and monomers, a quantitative analysis was performed using the program MaxQuant (version 1.5.5.1)(34) . For each cross-linked peptide identified, intensities were retrieved from the list of peptide isotope patterns quantified by MaxQuant applying tolerances for precursor *m/z* values of 20 ppm and a retention time window of 1 min. To establish intensity ratios, missing values were imputed assuming peptide intensities below or near the detection limit. To this end, missing values were simulated based on a normal distribution with a mean and standard deviation derived from the measured $\log_{10}$ intensities distribution median and standard deviation of all peptides quantified by MaxQuant (35). We used a downshift of the mean by 2.0 and adjustment of the width by a factor of 0.2, respectively. Intensity ratios were normalized to the median of log ratios of non-cross-linked (regular) peptides. Cross-linked peptides selected for quantitative analysis were first manually verified by inspection of PSMs using pLabel (see Supplementary Figure 11a), and subsequently quantified using Skyline, version 4.2(36) by extracting ion chromatograms after import of linearized cross-linked peptides (see Supplementary Figure 11b). Linearization of cross-linked peptides was performed as described in(37).

Native MS analyses

Recombinantly expressed proteins were buffer-exchanged into 200 mM ammonium acetate pH 6.8 at 4°C using microcentrifuge gel filtration columns Bio-Spin 6 (Bio-Rad Laboratories). The protein concentration was determined spectrophotometrically using a Nanodrop 2000c (Thermo Scientific) by measuring absorption at 280nm. Spectra were obtained using a Synapt HDMS quadrupole time-of-flight mass spectrometer (Waters and MS Vision) equipped with a 32,000 m/z range quadrupole by introducing 3 μl sample at a concentration of 10 μM with an in-house manufactured gold-coated capillary needle (borosilicate thin wall with filament, OD 1.0 mm, ID 0.78 mm; Harvard Apparatus) by nano-electrospray ionization. The capillary voltage was set to 1.3kV and the cone voltage to 180V. Backing pressure was 7 mbar and the trap cell was filled with argon gas at a pressure of 7×10^{-3} mbar. CID was performed by applying acceleration voltages to the trap cells. The transfer cell was generally kept at a pressure of about 20% of that in the trap cell and a low acceleration voltage of 5 V. For detergent solubilized membrane proteins, concentrations of 4 μM for Pex14p/Pex17p complexes or 10 μM for Pex14p were used and ammonium acetate was supplemented with 0.015% DDM. Cone voltage of 200V, backing and trap cell pressures of 11 mbar and 2.5×10^{-2} mbar, respectively, were used. Spectra were calibrated externally using caesium iodide. For spectra analysis and processing, MassLynx v 4.1 (Waters) and UniDec v 2.7.3(38) were used.

Isolation of native membrane complexes from *S. cerevisiae*

Membrane-bound complexes with TPA-tagged bait proteins were purified by affinity purification using IgG-coupled Sepharose, and subsequent cleavage of bound protein complexes with TEV protease as described earlier(11)(39) except that 1% Triton X-100 was used as detergent during the isolation and 0.1% Triton X-100 during SEC. The sample was subsequently visualized by negative stain EM.

Fluorescence microscopy

Wide-field fluorescence microscopy of cells expressing GFP C-terminally fused to the PTS1 tripeptide SKL, including sample preparation and image processing, was carried out as described previously using a Zeiss Axio Observer. Z1 microscope equipped with an alpha Plan-Aprochromat 100x objective and an AxioCam MR R3. GFP signals were visualized with a Zeiss 38 HE filterset.

Supplementary Table I *S. cerevisiae* strains used

S. cerevisiae strain	Description	Source or Ref.
UTL-7A	*MATα leu2-3, 112 ura3-52 trp1*	(2)
pex17Δ	*MATα, 112 ura3-52 trp1, pex17::LEU2*	(6)
UTL-7A-Pex17pTPA	*MATα leu2-3, 112 ura3-52 trp1, PEX17-TEV-ProtA-kanMX6*	(40)
UTL-7A-Pex17p$_{1\text{-}167}$TPA	*MATα leu2-3, 112 ura3-52 trp1, PEX17(1-167)-TEV-ProtA-kanMX6*	This study
UTL-7A-Pex14pTPA	*MATα leu2-3, 112 ura3-52 trp1, PEX17-TEV-ProtA-kanMX6*	(11)
UTL-7A-Pex13pTPA	*MATα leu2-3, 112 ura3-52 trp1, PEX17-TEV-ProtA-kanMX6*	This study
PCY2*pex14Δ*	*MATα, gal4Δ, gal80Δ, URA3::GAL1-lacZ, lys2-801^{amber}, his3-Δ200, trp1-Δ63, leu2 ade2-101ochre, pex14::kanMX*	(6)
CB199 *pex14Δ* (SC260)	*MATa, ura3-52, leu2-1, trp1-63, his3-200, arg4::natMX4, lys1::natMX3, PEX14::URA3*	(41)
CB199-Pex14pΔ1-50 (SC369)	*MATa, ura3-52, leu2-1, trp1-63, his3-200, arg4::natMX4, lys1::natMX3, PEX14(51-341)-TEV-ProtA-kanMX*	This study
CB199-Pex14p-M51L (SC370)	*MATa, ura3-52, leu2-1, trp1-63, his3-200, arg4::natMX4, lys1::natMX3, PEX14-M51L-TEV-ProtA-kanMX*	This study

Supplementary Table II Plasmids used

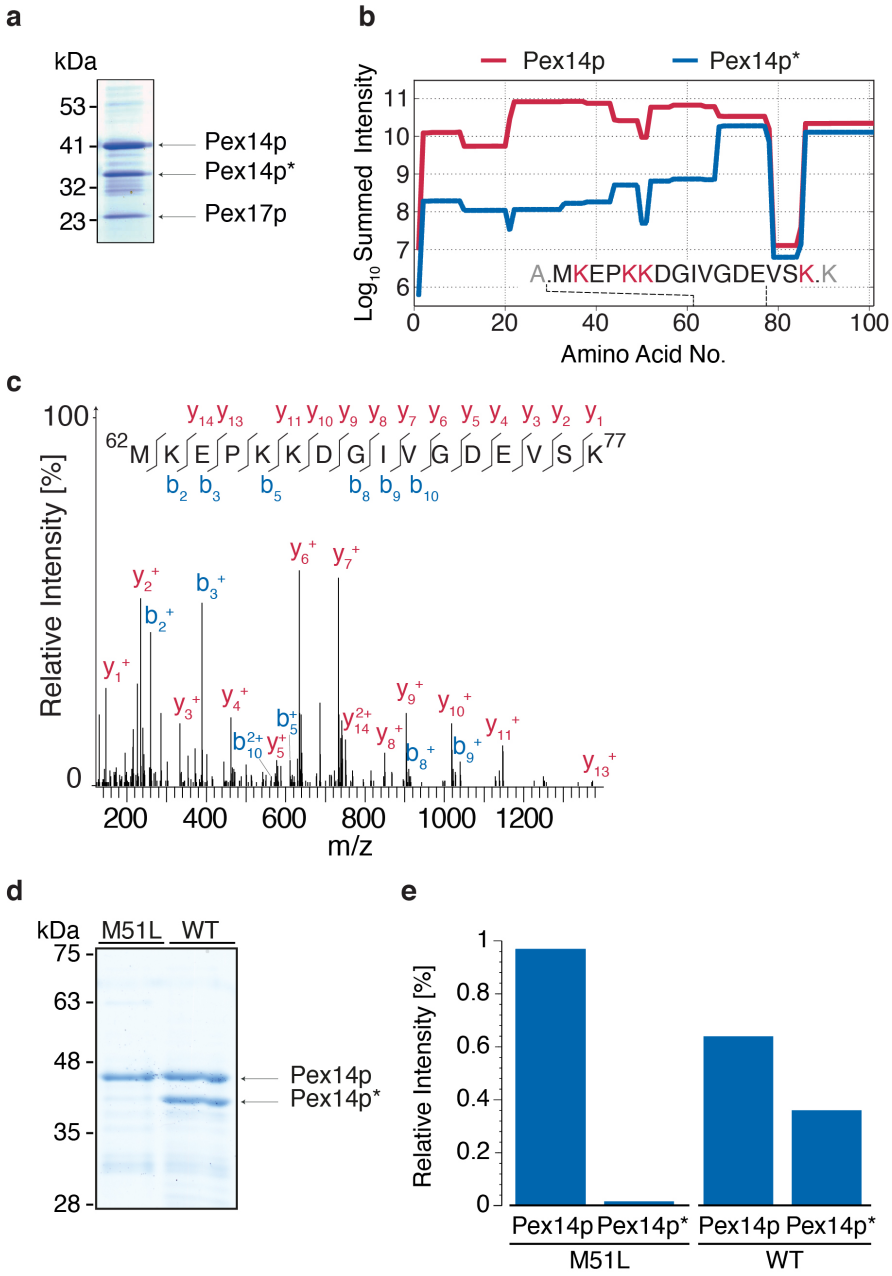
Plasmid	Description	Source or Reference	oligonucleotides
pPC97-*PEX17*	*GAL4-PEX17*aa1-199	(6)	
pPC97-*PEX17*XIII	*GAL4-PEX17*aa167-199	(40)	
pPC86-*PEX14*	*GAL4-PEX14*	(5)	
pPC86-*PEX14(aa1-270*	*GAL4-PEX14*	This study	KU107/KU176
pPC86-*PEX14(aa1-235)*	*GAL4-PEX14*	This study	KU107/KU515
pPC86-*PEX14(aa51-341)*	*GAL4-PEX14*	This study	KU173/KU177
pPC86-*PEX14(aa171-341)*	*GAL4-PEX14*	This study	KU174/KU177
pPC86-*PEX14(aa200-341)*	*GAL4-PEX14*	This study	KU XX/KU177
pPC86-*PEX14(aa171-270)*	*GAL4-PEX14*	This study	KU174/KU176
pPC86-*PEX14(aa171-197)*	*GAL4-PEX14*	This study	KU174/ RE XX
pSK9/1.4	PEX17	(6)	
pRS9/1.4	yeast vector, *PEX17*	(6)	
pRS416- *PEX17*I	yeast vector, *PEX17*aa1-167		KU170/KU288
pRS416- *PEX17*II	yeast vector, *PEX17*aa1-125		KU133/KU172/KU623/KU624
pTSC19	Strep$_{II}$-Pex14p/His$_6$-Pex17p	This study	RE5277/RE5278/RE6399/RE6402
pTH03	Strep$_{II}$-Pex14p(M51L)/ His$_6$-Pex17p	This study	RE6138/RE6139
pTH09	His$_6$-Pex14p(M51L)/Pex17p	This study	RE6138/RE6139
pTH30	Strep$_{II}$-Pex14p(M51L)	This study	
pET24a-PEX14(*aa119-341)*	*PEX14 (aa119-341)*	This study	O108/O109
pET24a-PEX14*(aa213-341)*	*PEX14 (aa213-341)*	This study	O108/O110
pET24a-PEX14*(aa1-95)*	*PEX14 (aa1-95)*	This study	O107/O111
pET24a-PEX14	*PEX14*	This study	O107/O108
pET24a-PEX14(M51L)	*PEX14 (M51L)*	This study	O889/O890
pRS316-PEX14 (pIS29)	yeast vector, *PEX14-TPA*	(41)	
pRS316-PEX14(M51L) (pAS415)	yeast vector, *PEX14-M51L-TPA*	This study	O889/O890
pGFP-SKL		(42)	

Supplementary Table III Oligonucleotides used

Oligonucleotide

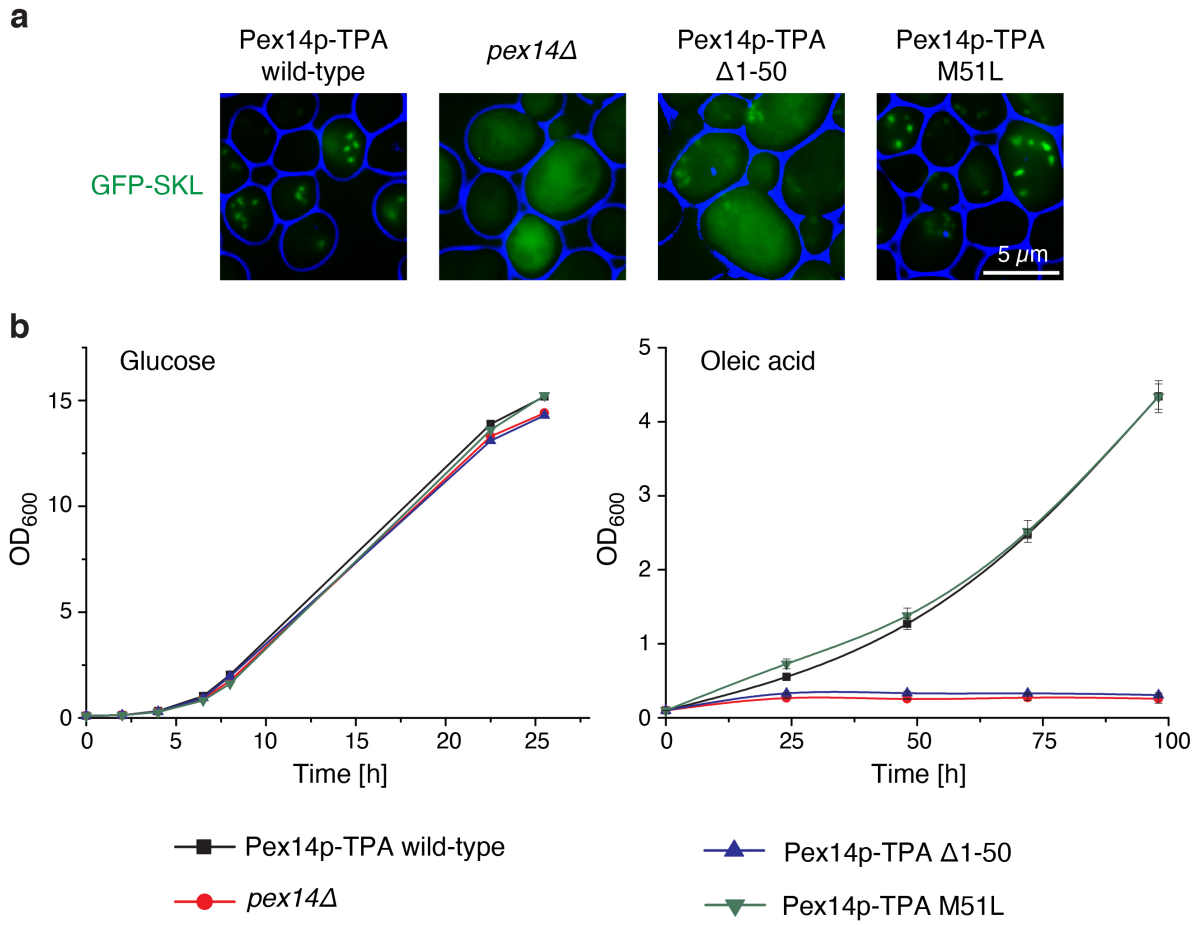
KU107 5´-TGTCTTTCAAATTAAGATGGTGACTGACCAGGTTGAGTTACGTACGCTGCAGGTCGAC-3´
KU170 5`-GGAATTCGATATCCCCCTTGGCACTTGGCCATTC-3`
KU172 5´-CCCAAGCTTGAATTCTGCGGAGATATGAAGAAG-3´
KU173 5`-GGAATTCGAGGCCTTATGAAAGAGCCCAAGAAA-3`
KU174 5`-GGAATTCGAGGCCTTATGAGTGACACGATTGCTGAA-3`
KU176 5`-GGACTAGTGGATCCCGGGAATCTCAGACGGACGCAGATGG-3`
KU177 5`-GGACTAGTCGGATCCCGGGGATGGAGTCTTCGA-3`
KU288 5`-CGGAATTCGTTAACATAATCTTGGATCCACTAG -3`
KU515 5´-TACTGAATTCAGTCGACCATGAGTGACGTGGTCA-3´
KU623 5´-GCGGATCCGAGCTCAATCAAAAATAATAGCATTG-3´
KU624 5´-CCCAAGCTTGATATCACAGGAACAAATG-3´
KU1005 5´-GAAGAAAATTGAGCATGTTGATGA TGAAACGCGTACACACCGTACGCTGCAGGTCGAC-3´
KU1006 5´-TATATATATATGCGAATATATGTGTGCAAATATTGATGCAATCGATGAATTCGAGCTCG-3´
KU1008 5´-CACTAGAGCGTTTTAAATTCAATGCTATTATTTTTGATTGATCGATGAATTCGAGCTCG-3´
KU1075 5´-TGTCTTTCAAATTAAGATGGTGACTGACCAGGTTGAGTTACGTACGCTGCAGGTCGAC-3´
REXX 5´-GGAATTCGAGGCCTTATGAGTGACGTGGTCAGT-3´
RE5277 5`- GATCCCATGGGGTGGTCTCACCCACAATTCGAGAAGAGTGCAAGTGACGTGGTCAGTAAAG-3`
RE5278 5`- GATCGAATTCCTATGGGATGGAGTCTTCG-3`
RE 6138 5'-AAAAGGAGATTGAAATAGCCCTGAAAGAGCCCAAGAAAGAC-3`
RE6139 5'-GTATTTCTTGGGCTCTTTCAGGGCTATTTCAATCTCCTTT-3'
RE6399 5‘ GATCCATATGATGCATCACCA 3‘
RE6402 5‘ GATCCTCGAGTTACCTTGGCA 3‘
O107 5’-GCATCGAGCTCATGAGTGACGTGGTCAGTAAAG- 3’
O108 5’-GTACTGCGGCCGCTGGGATGGAGTCTTCGACC- 3‘
O109 5’-GCATCGAGCTCGTGATACCAAATATTTTACCAGAAGC- 3’
O110 5’-GCATCGAGCTCGTTTCAGATAATGACGGCATGC- 3’
O111 5’-GTACTGCGGCCGCCTTCCAATCCCTGTGGGGC- 3’
O674 5'-CAATTACAATTTCCGTTAAAAAACTAATTACTTACATAGAATTGCGGTTTAAACTGGATGGCGGCG-3'
O675 5'-GTTAAATTCTCTTTCACTAAAAACTTCTTTAAATAGCTCCAACTCTTCCCTCCCGCCATAATTGCAC-3'
O889 5’-GAAATAGCCTTGAAAGAGCCC- 3’
O890 5’-GGGCTCTTTCAAGGCTATTTC- 3’
O899 5'-GAAAACTCAAGTAAAACAGAGAAGTTGTAAGGTGAATAAGGAATGAAAGAGCCCAAGAAAGACG-3'

Supplementary Figures

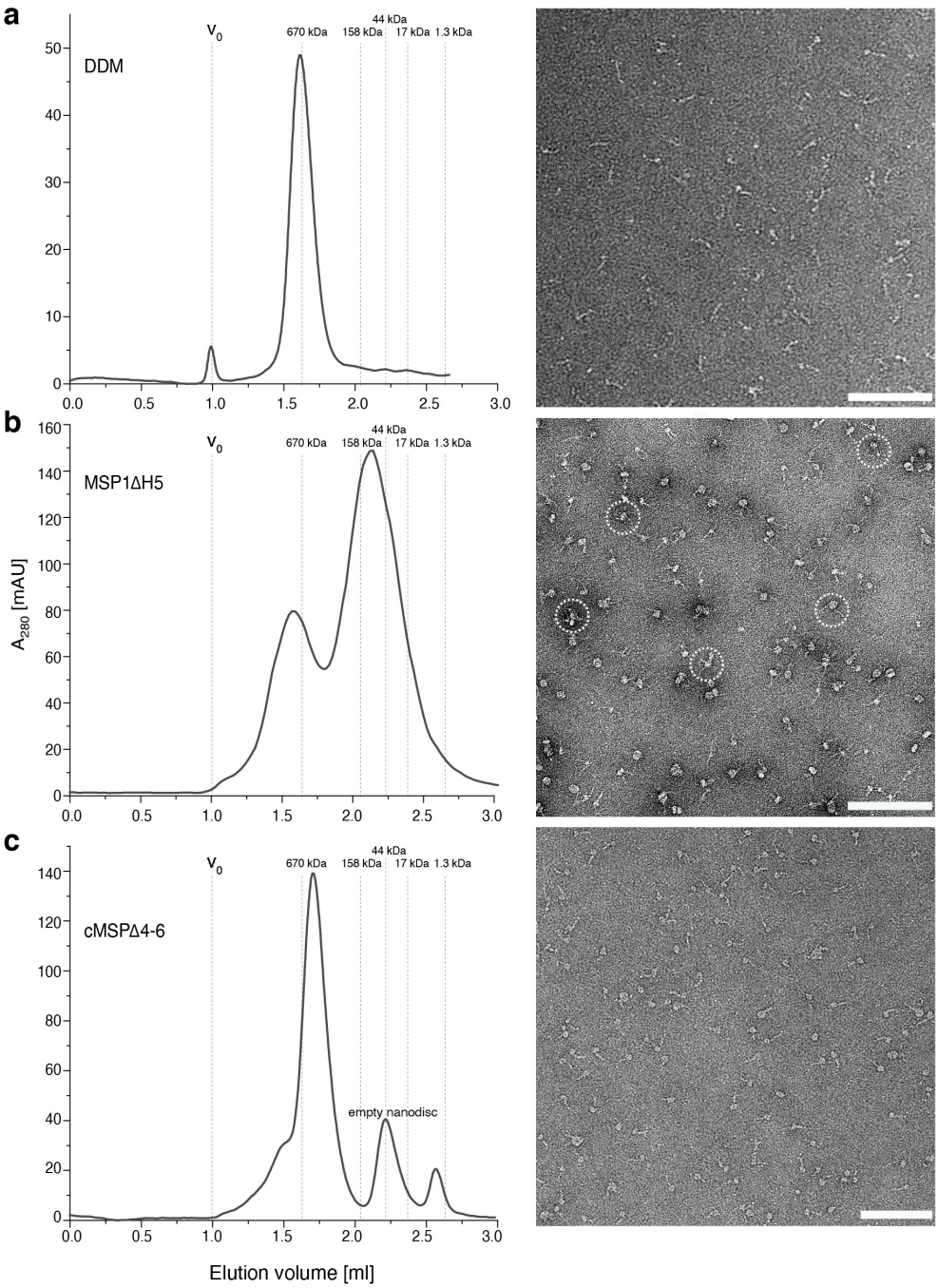


Supplementary Figure 1. Partial degradation of recombinant Pex14p by cleavage at methionine 51 was prevented by the mutation M51L. (a) $Strep_{II}$-Pex14p/His_6-Pex17pcomplex analyzed by SDS-PAGE revealed a dominant Pex14p-degradation product (Pex14p*) at an apparent molecular mass of 35 kDa. **(b)** Intensities measured by LC-MS of tryptic peptides recovered from individual gel bands of Pex14p (red line) and Pex14p* (blue line) indicate that the N-terminal part up to methionine 51 (M62 in $Strep_{II}$-Pex14p) was absent in the degradation product. The inset shows the identified peptide representing the putative N-terminus of Pex14p*. Predicted cleavage sites of trypsin are highlighted in red. **(c)** MS/MS spectrum identifying the peptide generated by cleavage N-terminal to methionine 51. Fragment ions of type y (red) and type b (blue) identifying the peptide sequence are marked. **(d)**

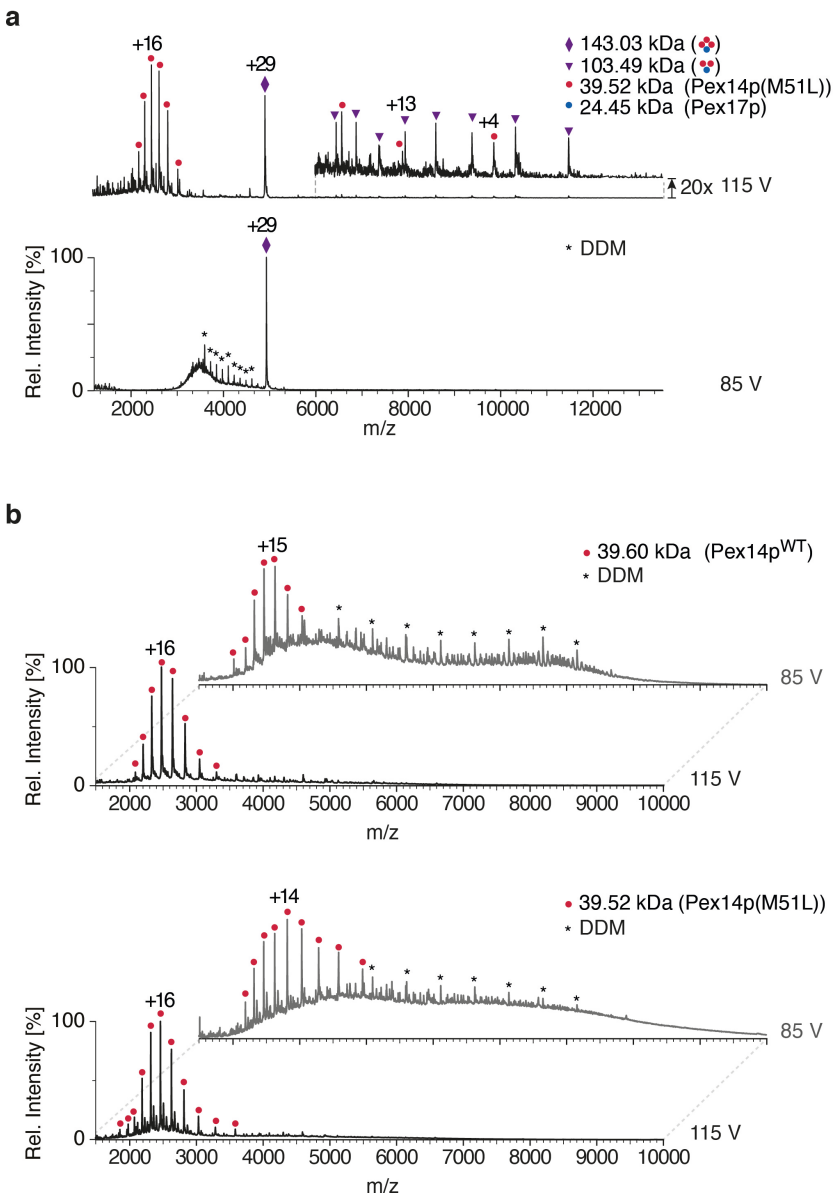
Recombinant Pex14p, wild-type and M51L were analyzed by SDS-PAGE and Coomassie staining. Proteins in individual gel slices were subjected to in-gel digestion using trypsin and quantitative LC-MS/MS analysis. The band assigned to the degradation product (Pex14p*) was significantly decreased in the M51L compared to wild-type. **(e)** LC-MS intensities of Pex14p peptides recovered from Pex14p* gel slices amounted to 1.6% and 36.1% of the total intensity for both gel slices for M51L and wild-type, respectively.



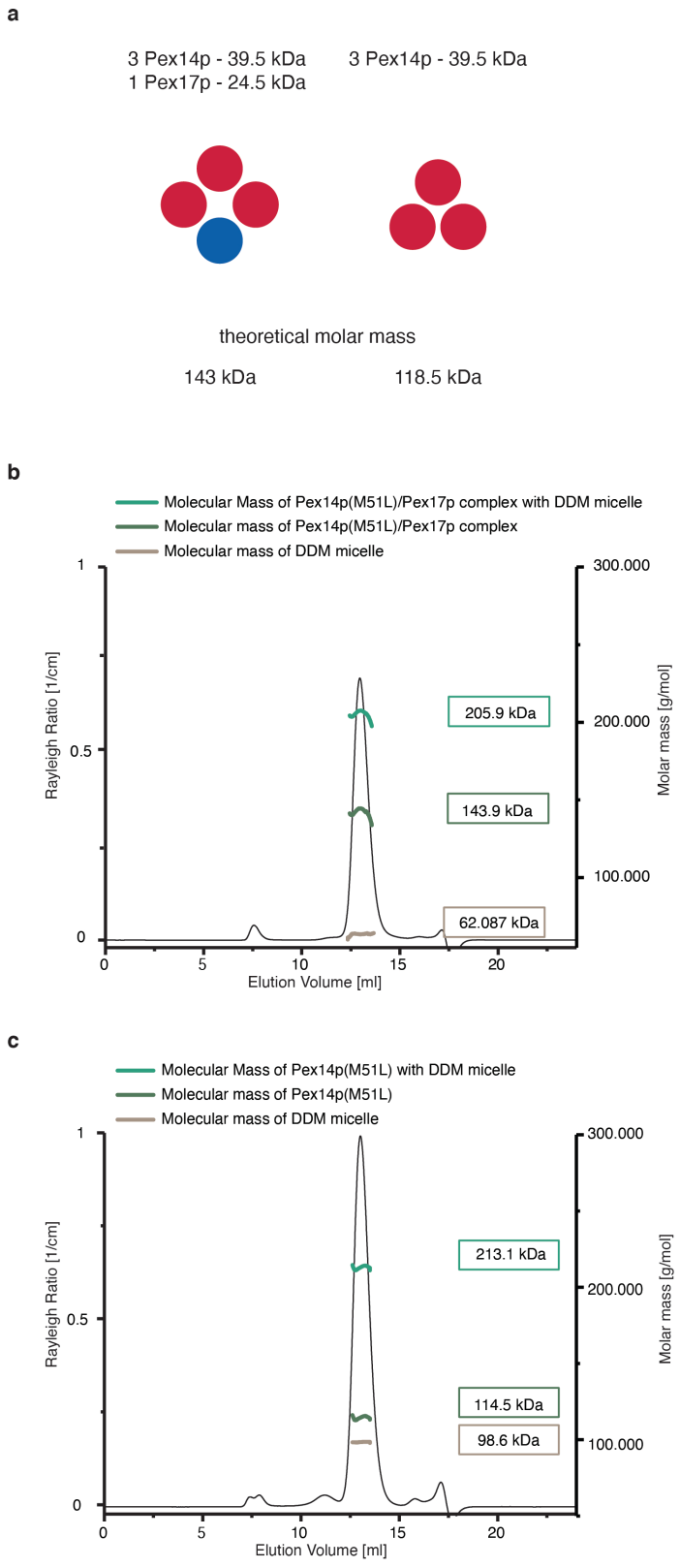
Supplementary Figure 2. Mutation M51L of Pex14p has no effect on growth and peroxisomal matrix protein import. (a) Representative fluorescence images of strains expressing native Pex14p (wild-type) and mutant forms (Δ1-50 or M51L) from the native *PEX14* locus as well as plasmid-encoded GFP-SKL, a peroxisomal reporter protein carrying a PTS1. Cells were pre-cultured in medium containing 0.3% glucose for 16 h at 30°C before they were shifted to oleic acid medium and grown for further 16 h. The *pex14*Δ strain serves as control for a defect in peroxisomal protein import. Scale bar, 5 µm. **(b)** Strains were pre-cultured in medium containing 0.3% glucose and grown for 16 h at 30°C. Cells were cultured as described in (a), but shifted to either glucose (2%) or oleic acid medium, samples were taken at indicated time points, and the OD_{600} was determined at indicated time points. Growth in oleic acid medium was analyzed in triplicates. Error bars indicate standard deviation.



Supplementary Figure 3. Size-exclusion chromatography and negative stain EM analysis of $Strep_{II}$-Pex14p(M51L)/His_6-Pex17p. Gel-filtration chromatogram (left panel) and typical EM micrograph (right panel) of Pex14p/Pex17p in 0.1% DDM **(a),** (MSP)1D1dH5 nanodiscs (ratio 1:30) **(b)** and MSP1D1Δ4-6 circular nanodiscs (ratio 1:30) **(c).** The chromatography experiments were performed with a Superose 6 increase 5/150 column with a total volume of 3 ml. Dashed lines represent the BioRad protein standard (#1511901). Dashed circles in **(b)** indicate proteo-nanodiscs including multiple copies of Pex14p/Pex17p. Scale bars 100nm.

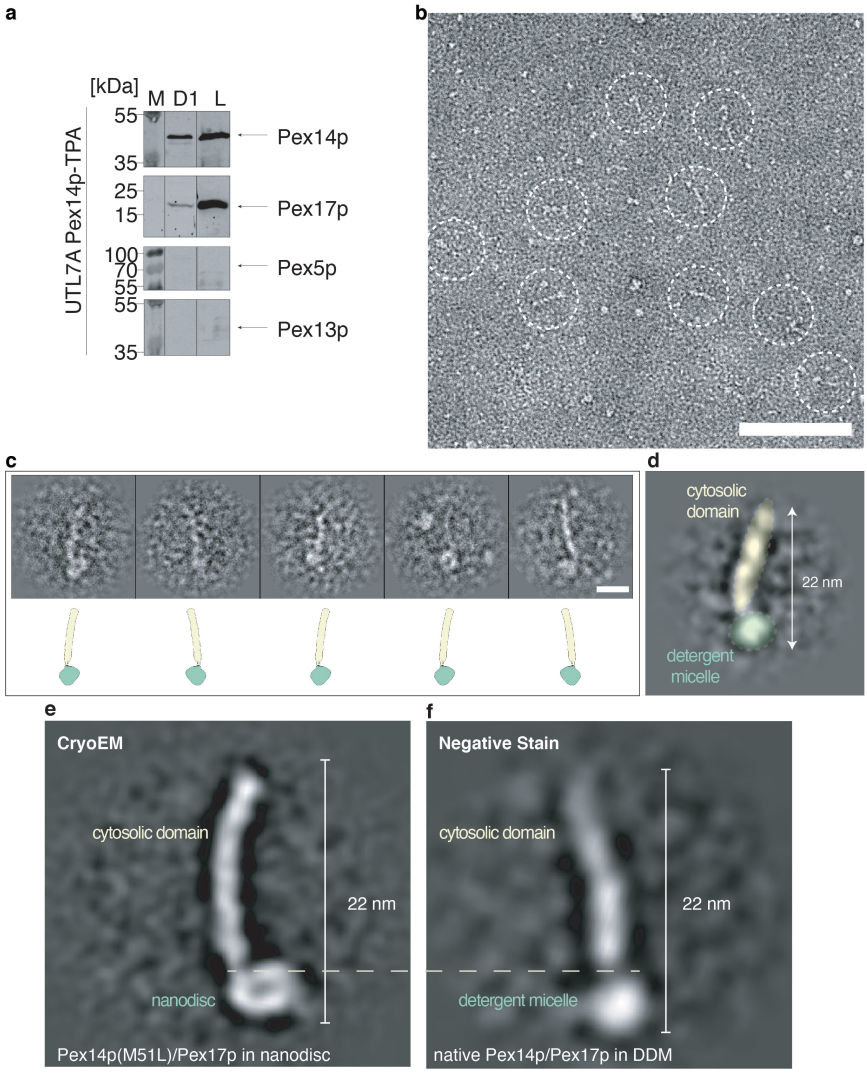


Supplementary Figure 4. Native MS analysis of recombinant Pex14p(M51L)/Pex17p, Pex14p(WT) and Pex14p(M51L). Proteins were transferred into 200 mM ammonium acetate, pH 6.8, 0.015% DDM. **(a)** MS^2 spectrum of the $Strep_{II}$-Pex14p(M51L)/His_6-Pex17p complex, generated by selection of the +29-charge state at *m/z* 4,940 in the quadrupole and subsequent collisional activation by acceleration voltage of 85 V (bottom) or 115 V (top). Main product ion series show highly charged monomeric Pex14p centered around charge state +16 and a charge-reduced oligomer with a mass of 103.49 kDa corresponding to a 2:1 subunit stoichiometry of Pex14p:Pex17p (inset; 20-fold magnified spectrum). **(b)** Native MS spectra of His_6-Pex14p (top) and $Strep_{II}$-Pex14p(M51L) (bottom) applying collisional activation by two different acceleration voltages as indicated. Signals are attributed to charge state series of the monomeric proteins as shown by the red filled circles. Signals attributed to clusters of DDM molecules are marked by asterisks.

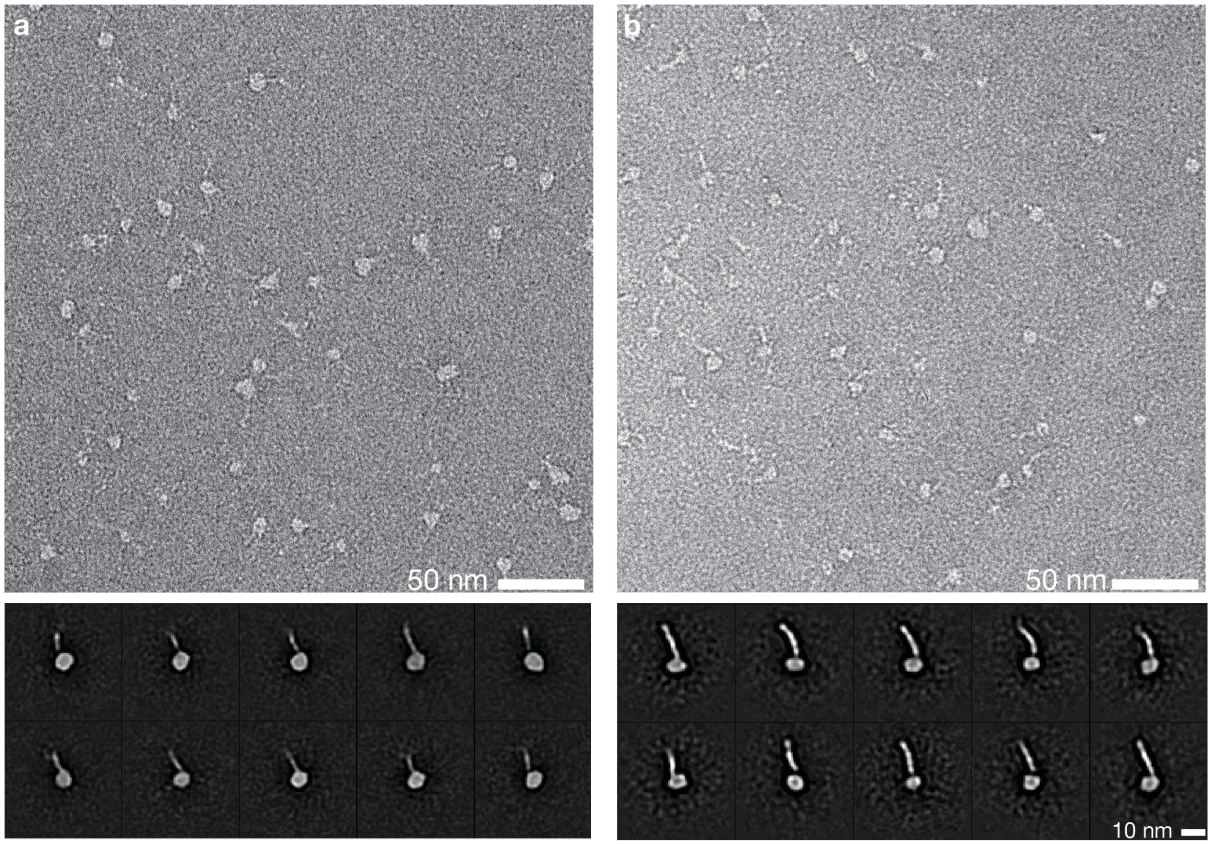


Supplementary Figure 5. Chromatogram of size exclusion chromatography with multi angle light scattering. SEC-MALS measurements of Pex14p(M51L) and Pex14p(M51L)/Pex17p complex in 0.1% DDM. In order to distinguish the protein from the molar mass of the DDM micelle a protein conjugate analysis using Astra 7.3.1 (Wyatt) has

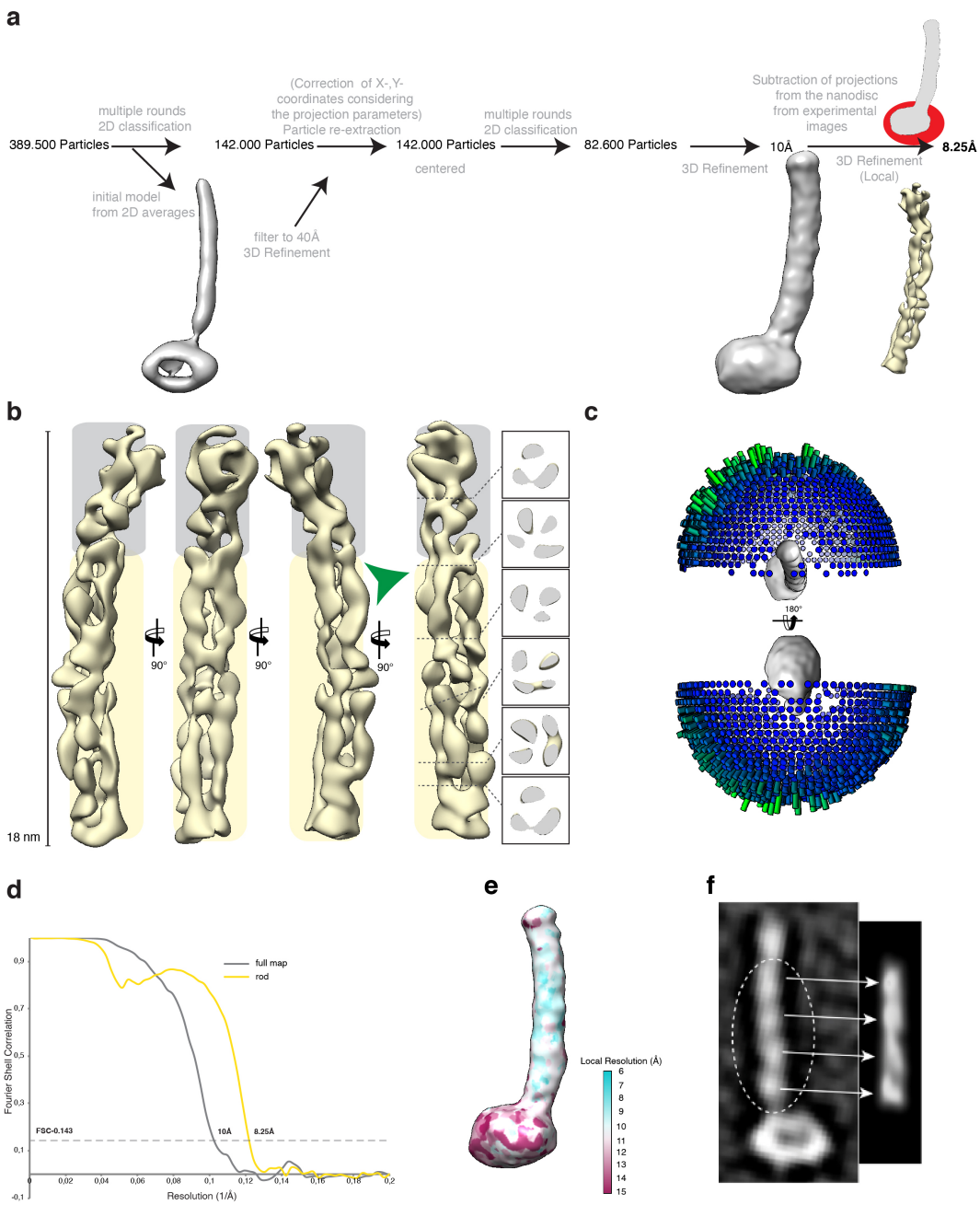
been performed. **(a)** Schematic view of Pex14p and Pex14p/Pex17p complex representing the calculated molar mass for a 3:1 and 3:0 complex, respectively. **(b)** SEC-MALS protein conjugate analysis of the Pex14p(M51L)/Pex17p complex in DDM. The calculated molecular masses for total (205.9 kDa), protein (143.9 kDa) and modifier (DDM micelle 62 kDa) are indicated. **(c)** Measurement of Pex14p(M51L). The measurement resulted in a molecular weight of 114.5 kDa with a corresponding DDM micelle of 98.6 kDa. All measurements were performed in TBS buffer with 0.1% DDM on a Superose 6 10/300 column at 6°C



Supplementary Figure 6. Negative stain EM analysis of the native Pex14p/Pex17p complex (a) Native complexes were isolated from Triton X-100 solubilized whole cell membranes from oleic acid induced cells expressing a Pex14p-TPA. Complexes were eluted from IgG-Sepharose by TEV protease cleavage. The obtained eluate was further analyzed by size-exclusion chromatography and subjected to immunoblot analysis with antibodies as indicated. The isolated native Pex14p/Pex17p complex in fraction D1 was subsequently analyzed with negative stain EM. (M=Marker; D1=Fraction 1 of the gel filtration (approximately 700 kDa); L= Load) **(b)** Representative negative stain micrograph of fraction D1. White circles indicate typical Pex14p/Pex17p rod-like particles. **(c)** Representative single particles. Scale bar 10nm **d)** Representative reference-free class average of the native Pex14p/Pex17p complex. Densities of the detergent micelle (green) and the cytosolic domains (yellow) are highlighted. **(e, f)** Side-by-side comparison between reference-free class averages of the recombinantly expressed $Strep_{II}$-Pex14p(M51L)/His_6-Pex17p in nanodisc (cryoEM) and the native Pex14p/Pex17p complex in Triton X-100 (negative stain).

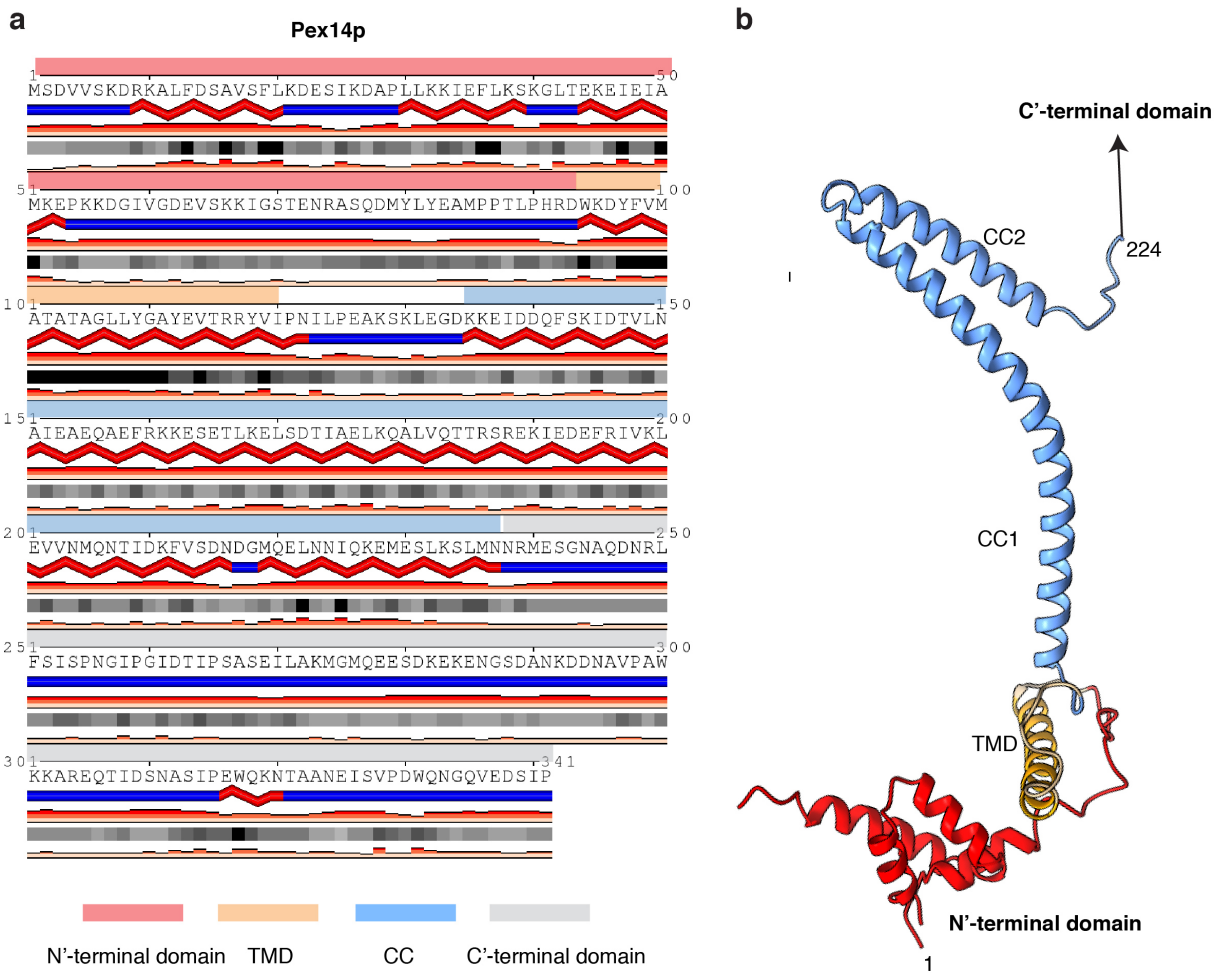


Supplementary Figure 7 Negative stain comparison of Pex14p(M51L) and Pex14p(M51L)/Pex17p reconstituted in cMSP46 nanodiscs. (a) Micrograph of Pex14p(M51L) showing various populations of the particles. Bottom panel: Reference free two-dimensional class averages of Pex14p(M51L). The classes represent a high variability (10-17nm) in length at the part of the elongated rod. **(b)** Negative stain micrograph of reconstituted Pex14p(M51L)/Pex17p. Bottom panel: Reference free class averages of Pex14p(M51L)/Pex17p. Classes display a homogenous size distribution with an overall size of 20-22nm

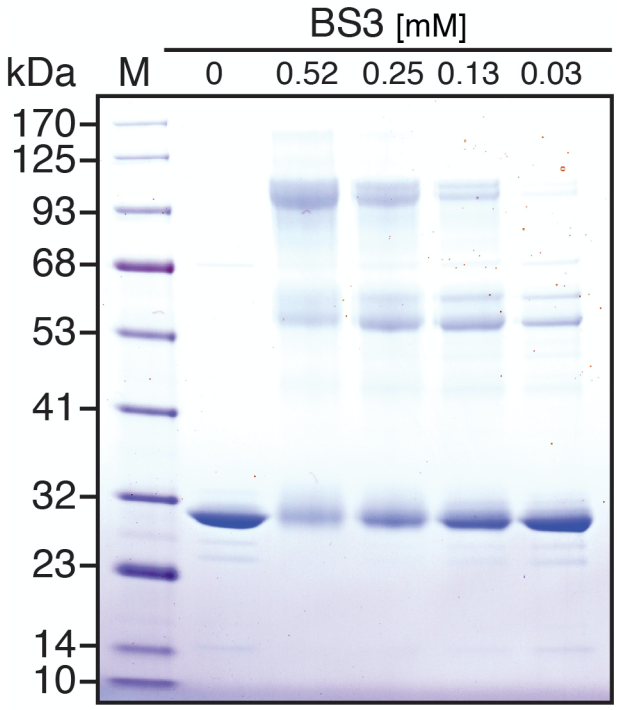


Supplementary Figure 8. CryoEM analysis of Pex14p(M51L)/Pex17p. (a) Single particle processing workflow for Pex14p(M51L)/Pex17p structure determination. **(b)** Resulting map from focused 3D local refinement on the rod region after signal subtraction of the nanodisc. Insets show slices through the rod density at different heights. The two structural regions of the rod are highlighted in gray and yellow. The green arrowhead indicates the kinked connection between these two structural regions. **(c)** Orientation distribution of the particles used in the final refinement round prior signal subtraction. **(d)** Fourier Shell Correlation (FSC) curves between two independently refined half maps derived after standard refinement of the complete dataset and after signal subtraction of the nanodisc projections from the experimental images and subsequent local refinement (see, a). **(e)** Local resolution map of the

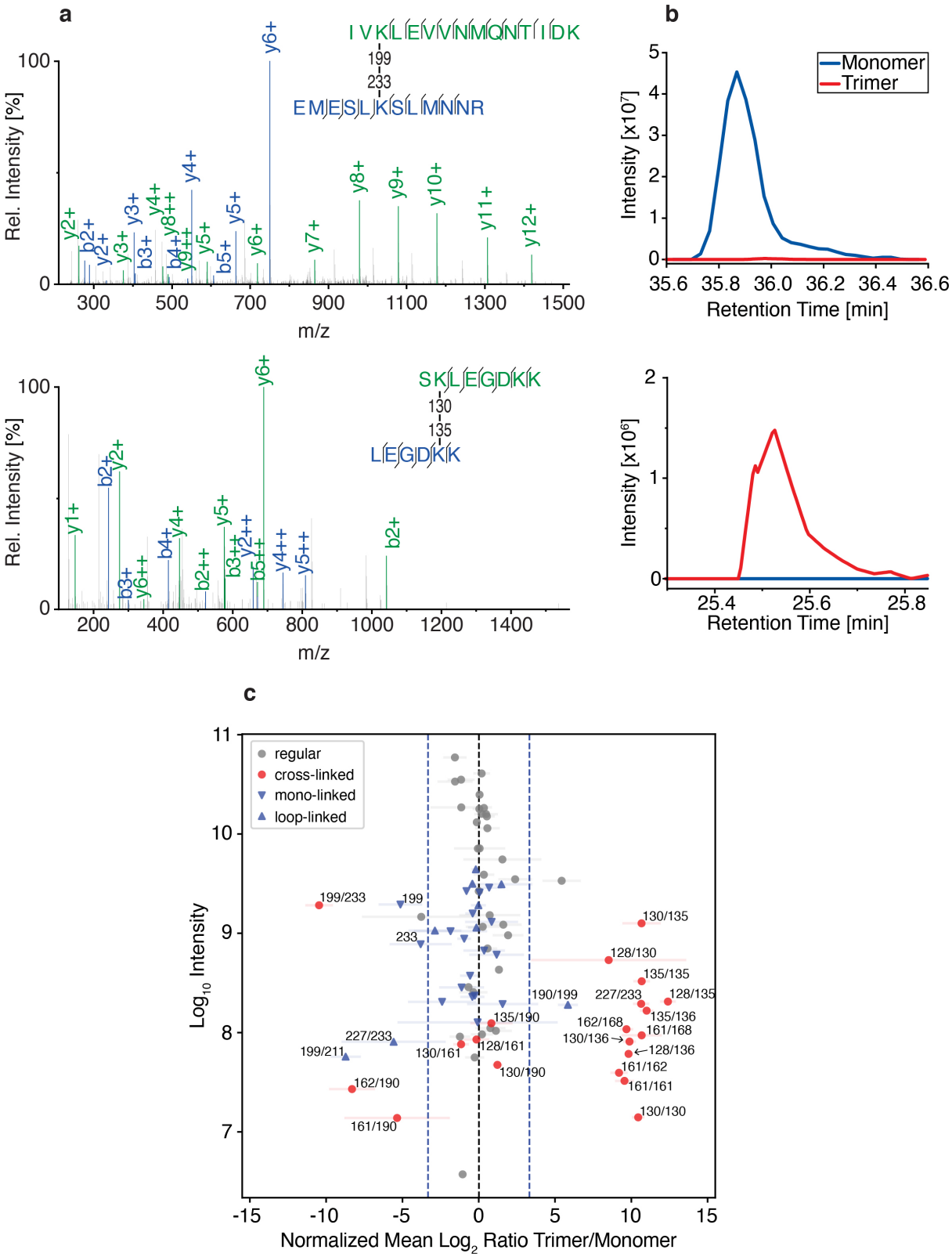
Pex14p/Pex17p complex prior signal subtraction. **(f)** Side-by-side comparison between a characteristic 2D cryoEM class-average of Pex14p/Pex17p (left image) and a 2D back projection (right image) from a density simulated from the poly-alanine model of the lower rod domain at 20Å (dashed white circle). Note the characteristic elongated shape with four densities arranged in a chain-like manner, matching well the respective region of the 2D average (arrows). This region is followed by the predicted second short coiled-coil domain of Pex14p and the C-terminal domain of Pex17p (fifth density in the average). This region was not well resolved in our cryoEM density and therefore not included in the poly-alanine model.



Supplementary Figure 9. Structure prediction for the Pex14p subunit. (a) Sequence-based secondary structure prediction for Pex14p. Red bands represent α-helices. The subunit can be roughly divided into four domains (see also Figure 1a): N-terminal α-helical domain (highlighted in red), transmembrane domain (TMD; highlighted in brown), coiled-coil domain (highlighted in light blue) and the unstructured C-terminal domain (highlighted in gray) **(b)** Three-dimensional structure of Pex14p as predicted by RaptorX (27) (coloring similar to (a)).

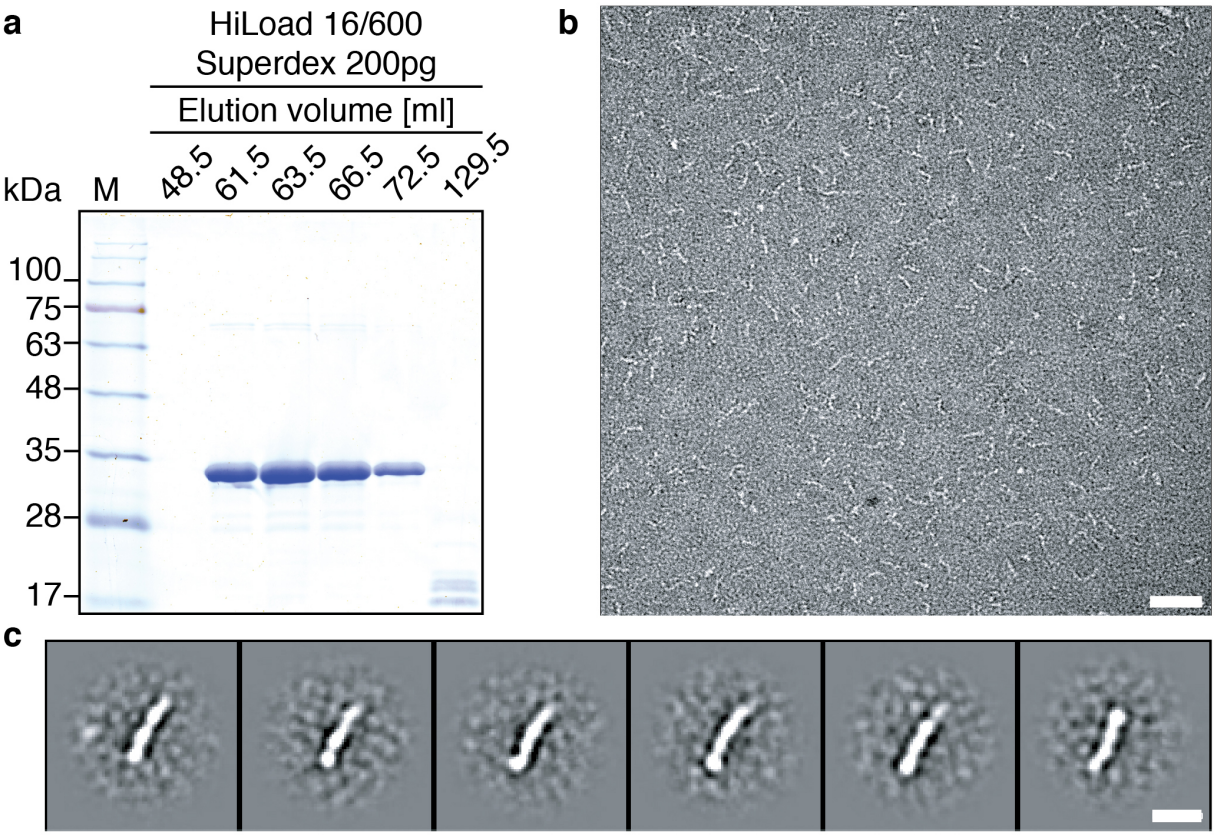


Supplementary Figure 10. *In vitro* cross-linking leads to trimers of Pex14p$_{119-341}$. Recombinant Pex14p$_{119-341}$ was cross-linked using BS3 at increasing concentration as indicated and separated by SDS-PAGE followed by Coomassie staining.

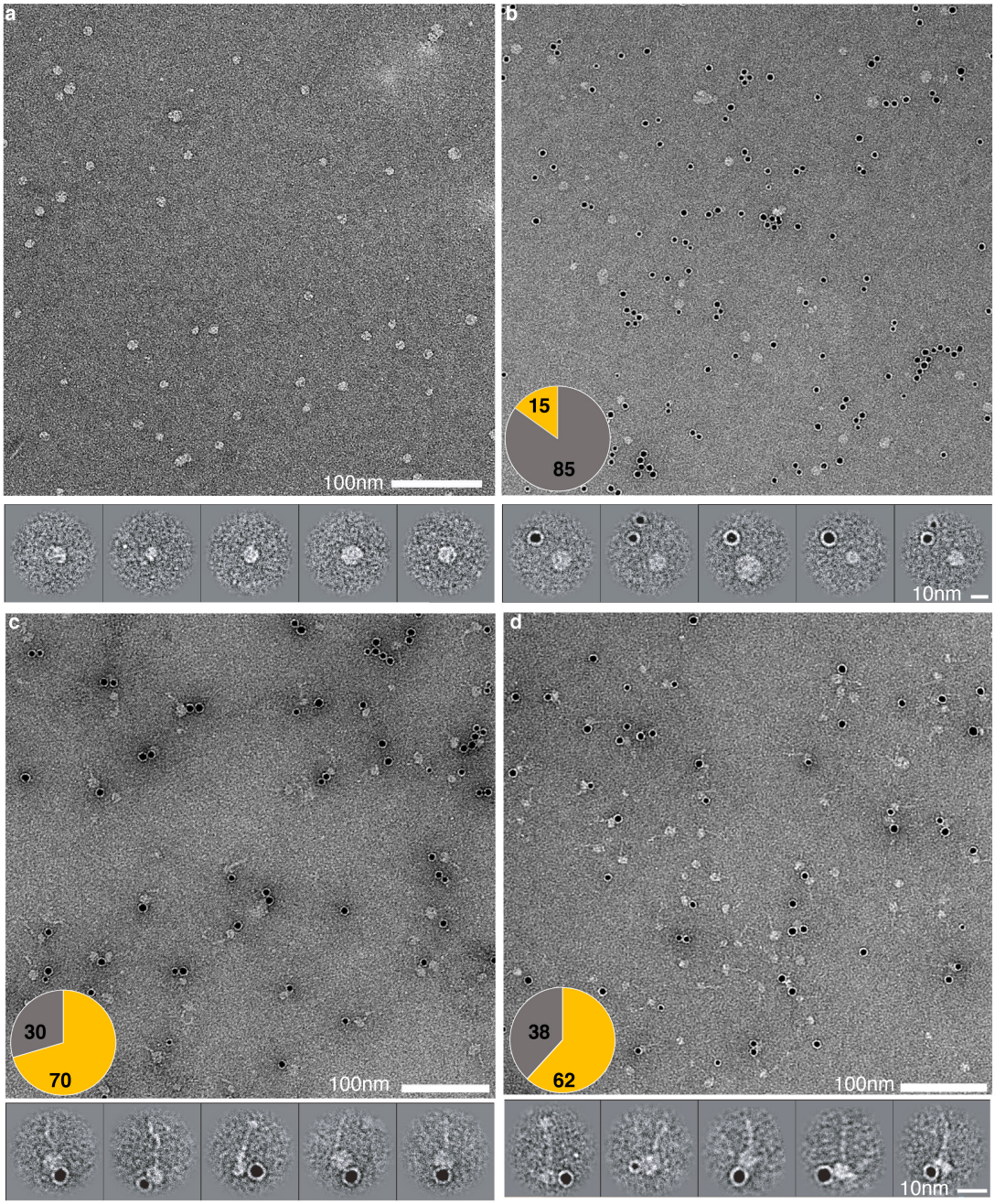


Supplementary Figure 11. Intra- and intermolecular cross-links of Pex14p homooligomers. Following chemical cross-linking with BS3 and SDS-PAGE analysis, monomeric, dimeric and trimeric Pex14p$_{119-341}$ forms (see Figure 4a) were subjected to in-gel digestion with trypsin and analyzed by LC-MS/MS. **(a)** MS/MS spectrum identifying a monomer-specific Pex14p intraprotein cross-linked peptide pair (top) and an oligomer-specific

interprotein cross-linked peptide pair (bottom). Ions matched to masses of b- and y-ions (see inset) of the longer (green) and shorter (blue) peptides are marked. **(b)** Representative ion chromatograms showing the elution profiles of the corresponding monomer- (top) and trimer-specific (bottom) cross-linked peptide pairs. **(c)** Quantitative LC-MS analysis of cross-linked amino acid pairs (red dots), regular peptides (gray dots) and mono- (blue triangles down) or loop-linked amino acids (blue triangles up). Normalized $\log_2$ trimer-to-monomer intensity ratios were computed and plotted against their $\log_{10}$ intensities. Identified linkages are annotated with their sequence positions. Blue vertical lines indicate a 10-fold enrichment either in the trimer or monomer sample. Error bars indicate standard deviation.

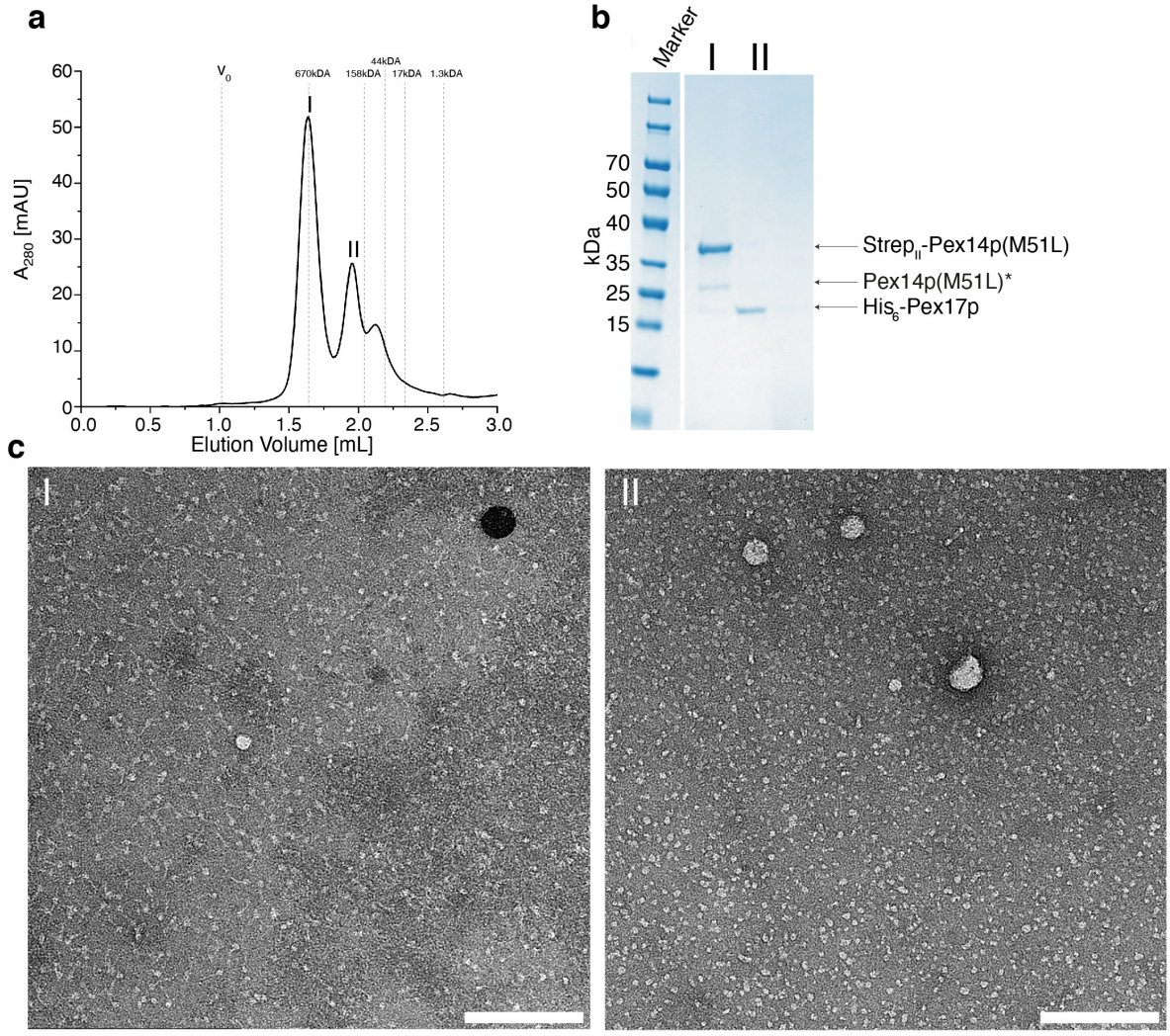


Supplementary Figure 12. Negative stain EM analysis of the homotrimeric Pex14p$_{119-341}$ variant. (a) Coomassie stained SDS-PAGE of the Pex14p$_{119-341}$ variant after purification using a HiLoad 16/600 Superdex 200. **(b)** Representative EM micrograph with **(c)** corresponding classes**,** scale bars 50nm and 10nm, respectively.



Supplementary Figure 13. Immunogold labeling of His_6-Pex14(M51L)p/Pex17p and Pex14(M51L)p/His_6-Pex17 (a) Representative negative stain EM micrograph of untagged MSP1D1 nanodiscs. Extracted single particles are shown. **(b)** Representative micrograph and extracted raw particles of the untagged MSP1S1 nanodiscs upon 30 min incubation with Ni-NTA gold. Only few nanodiscs are in close proximity with gold-particles, indicating rather random colocalization. **(c)** Representative micrograph of the immunogold labeling experiment for the localization of the N-terminal His_6-tag of Pex14p and extracted raw particles. Ni-NTA gold is binding the majority of the particles close to the nanodisc density at the opposite side of the rod. **(d)** Representative micrograph of the immunogold labeling experiment for the localization of the N-terminal His_6-tag of Pex17p and extracted raw particles. Ni-NTA gold is binding also in this case in close proximity to the nanodisc at the opposite side of the rod.

Insets show the statistics of gold labeled particles for the respective sample. Given numbers represent the ratio of labelled (orange) and unlabeled (grey) particles (%) from datasets of 500 particles of each Pex14Pex17p sample. The dataset of the control (b) contained 200 particles.



Supplementary Figure 14. Strep$_{II}$-Pex14p(M51L)/His$_6$-Pex17p dissociates after exchange of DDM with amphipols A8-35. (a) Size exclusion chromatography profile of Pex14p/Pex17p in amphipols using a Superose 6 increase 5/150 GL column with a total volume of 3ml. Peak I and II correspond to Pex14p and Pex17p, respectively. **(b)** SDS-PAGE of the peak fractions I and II after size exclusion chromatography. **(c)** Representative negative stain electron micrographs of the Pex14p (left panel) and Pex17p (right panel) fractions. Scale bar, 100nm.

Video 1. Flexibility of Pex14p/Pex17p in 2D. The arrow indicates flexible densities below the nanodisc.

Video 2. CryoEM map of the cytosolic rod. The cryoEM map of the cytosolic domains obtained after signal subtraction of the nanodisc density from the experimental images and subsequent focused 3D local refinement is rotated in order to show the overall structure. α-helixes are fitted within rod-like densities. α-helices of the coiled-coil domain 1 of Pex14p are shown in blue (different shade of blue for each Pex14p subunit) and the helices of coiled-coil domain 1 and 2 of Pex17p are shown in yellow.

Video 3. 3D segmentation of a Pex14p/Pex17p proteoliposome.

REFERENCES

1. Knop M, et al. (1999) Epitope tagging of yeast genes using a PCR-based strategy: more tags and improved practical routines. *Yeast* 15(10B):963–972.

2. Erdmann R, Veenhuis M, Mertens D, Kunau WH (1989) Isolation of peroxisome-deficient mutants of Saccharomyces cerevisiae. *Proc Natl Acad Sci USA* 86(14):5419–5423.

3. FJ A, et al. (1992) *Current Protocols in Molecular Biology* (Greene Publishing Associates, New York) Available at: http://www.aun.edu.eg/molecular_biology/PCR(1)/Current%20Protocols%20in%20Mol.%20Biol..pdf.

4. Chevray PM, Nathans D (1992) Protein interaction cloning in yeast: identification of mammalian proteins that react with the leucine zipper of Jun. *Proc Natl Acad Sci USA* 89(13):5789–5793.

5. Albertini M, et al. (1997) Pex14p, a peroxisomal membrane protein binding both receptors of the two PTS-dependent import pathways. *Cell* 89(1):83–92.

6. Huhse B, et al. (1998) Pex17p of Saccharomyces cerevisiae is a novel peroxin and component of the peroxisomal protein translocation machinery. *J Cell Biol* 140(1):49–60.

7. Fields S, Song O (1989) A novel genetic system to detect protein-protein interactions. *Nature* 340(6230):245–246.

8. Rehling P, et al. (1996) The import receptor for the peroxisomal targeting signal 2 (PTS2) in Saccharomyces cerevisiae is encoded by the PAS7 gene. *EMBO J* 15(12):2901–2913.

9. Schneider S, Buchert M, Hovens CM (1996) An in vitro assay of beta-galactosidase from yeast. *BioTechniques* 20(6):960–962.

10. Stein K, Schell-Steven A, Erdmann R, Rottensteiner H (2002) Interactions of Pex7p and Pex18p/Pex21p with the peroxisomal docking machinery: implications for the first steps in PTS2 protein import. *Mol Cell Biol* 22(17):6056–6069.

11. Agne B, et al. (2003) Pex8p: an intraperoxisomal organizer of the peroxisomal import machinery. *Mol Cell* 11(3):635–646.

12. Albertini M, Girzalsky W, Veenhuis M, Kunau WH (2001) Pex12p of Saccharomyces cerevisiae is a component of a multi-protein complex essential for peroxisomal matrix protein import. *Eur J Cell Biol* 80(4):257–270.

13. Girzalsky W, et al. (1999) Involvement of Pex13p in Pex14p localization and peroxisomal targeting signal 2-dependent protein import into peroxisomes. *J Cell Biol* 144(6):1151–1162.

14. Zhao H, Brown PH, Schuck P (2011) On the distribution of protein refractive index increments. *Biophys J* 100(9):2309–2317.

15. Strop P, Brunger AT (2005) Refractive index-based determination of detergent concentration and its application to the study of membrane proteins. *Protein Sci* 14(8):2207–2211.

16. Wagner T, et al. (2019) SPHIRE-crYOLO is a fast and accurate fully automated particle picker for cryo-EM. *Communications Biology 2019 2:1* 2(1):218.

17. Moriya T, et al. (2017) High-resolution Single Particle Analysis from Electron Cryo-microscopy Images Using SPHIRE. *J Vis Exp* (123). doi:10.3791/55448.

18. Tribet C, Audebert R, Popot JL (1996) Amphipols: polymers that keep membrane proteins soluble in aqueous solutions. *Proc Natl Acad Sci USA* 93(26):15047–15050.

19. Miehling J, Goricanec D, Hagn F (2018) A Split-Intein-Based Method for the Efficient Production of Circularized Nanodiscs for Structural Studies of Membrane Proteins. *Chembiochem* 19(18):1927–1933.

20. Zheng SQ, et al. (2017) MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nature Methods* 14(4):331–332.

21. Penczek PA, et al. (2014) CTER-rapid estimation of CTF parameters with error assessment. *Ultramicroscopy* 140:9–19.

22. Yang Z, Fang J, Chittuluru J, Asturias FJ, Penczek PA (2012) Iterative Stable Alignment and Clustering of 2D Transmission Electron Microscope Images. *Structure/Folding and Design* 20(2):237–247.

23. Schorb M, Haberbosch I, Hagen WJH, Schwab Y, Mastronarde DN (2019) Software tools for automated transmission electron microscopy. *Nature Methods* 16(6):471–477.

24. Kremer JR, Mastronarde DN, McIntosh JR (1996) Computer visualization of three-dimensional image data using IMOD. *Journal of Structural Biology* 116(1):71–76.

25. Pettersen EF, et al. (2004) UCSF Chimera?A visualization system for exploratory research and analysis. *J Comput Chem* 25(13):1605–1612.

26. Goddard TD, et al. (2018) UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci* 27(1):14–25.

27. Källberg M, et al. (2012) Template-based protein structure modeling using the RaptorX web server. *Nature Protocols* 7(8):1511–1522.

28. Delorenzi M, Speed T (2002) An HMM model for coiled-coil domains and a comparison with PSSM-based predictions. *Bioinformatics* 18(4):617–625.

29. Morgenstern M, et al. (2017) Definition of a High-Confidence Mitochondrial Proteome at Quantitative Scale. *Cell Rep* 19(13):2836–2852.

30. Rappsilber J, Mann M, Ishihama Y (2007) Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nature Protocols* 2(8):1896–1906.

31. Chen Z-L, et al. (2019) A high-speed search engine pLink 2 with systematic evaluation for proteome-scale identification of cross-linked peptides. *Nat Commun* 10(1):3404–12.

32. Wang L-H, et al. (2007) pFind 2.0: a software package for peptide and protein identification via tandem mass spectrometry. *Rapid Commun Mass Spectrom* 21(18):2985–2991.

33. Combe CW, Fischer L, Rappsilber J (2015) xiNET: cross-link network maps with residue resolution. *Mol Cell Proteomics* 14(4):1137–1147.

34. Tyanova S, Temu T, Cox J (2016) The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nature Protocols* 11(12):2301–2319.

35. Deeb SJ, D'Souza RCJ, Cox J, Schmidt-Supprian M, Mann M (2012) Super-SILAC allows classification of diffuse large B-cell lymphoma subtypes by their protein expression profiles. *Mol Cell Proteomics* 11(5):77–89.

36. Schilling B, et al. (2012) Platform-independent and label-free quantitation of proteomic data using MS1 extracted ion chromatograms in skyline: application to protein acetylation and phosphorylation. *Mol Cell Proteomics* 11(5):202–214.

37. Müller F, Fischer L, Chen ZA, Auchynnikava T, Rappsilber J (2018) On the Reproducibility of Label-Free Quantitative Cross-Linking/Mass Spectrometry. *J Am Soc Mass Spectrom* 29(2):405–412.

38. Marty MT, et al. (2015) Bayesian deconvolution of mass and ion mobility spectra: from binary interactions to polydisperse ensembles. *Anal Chem* 87(8):4370–4376.

39. Hansen T, et al. (2017) Isolation of Native Soluble and Membrane-Bound Protein Complexes from Yeast Saccharomyces cerevisiae. *Methods Mol Biol* 1595(10B):37–44.

40. Girzalsky W, et al. (2006) Pex19p-dependent targeting of Pex17p, a peripheral component of the peroxisomal protein import machinery. *Journal of Biological Chemistry* 281(28):19417–19425.

41. Schummer A, et al. (2020) Pex14p Phosphorylation Modulates Import of Citrate Synthase 2 Into Peroxisomes in Saccharomyces cerevisiae. *Front Cell Dev Biol* 8:635.

42. Schäfer A, Kerssen D, Veenhuis M, Kunau WH, Schliebs W (2004) Functional similarity between the peroxisomal PTS2 receptor binding protein Pex18p and the N-terminal half of the PTS1 receptor Pex5p. *Mol Cell Biol* 24(20):8895–8906.