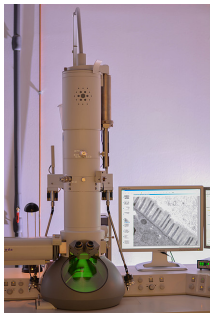
CSSB
Centre for Structural Systems Biology

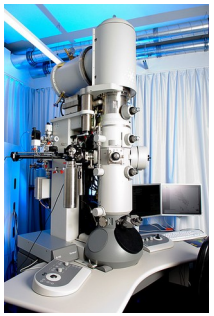
Cryo-Electron Microscopy Software

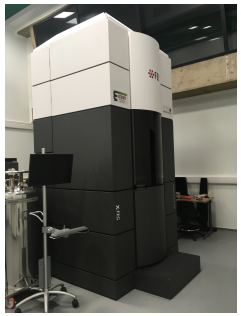
Wolfgang Lugmayr, UKE Hamburg

DESY GPU Computing & Machine Learning Workshop
6.11.2018

CSSB
Centre for Structural Systems Biology



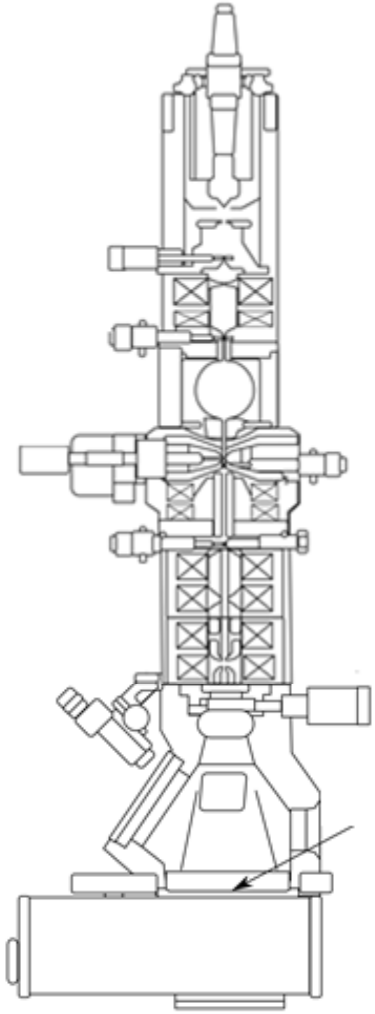




General considerations for CryoEM

- Illumination requirements
  - Beam diameter, Dose, Parallel beam
- Optical requirements
  - Astigmatism correction, Coma free alignment
- Image requirements
  - Pixel size (Nyquist theorem), Exposure time (dose fractionation), Camera type & mode

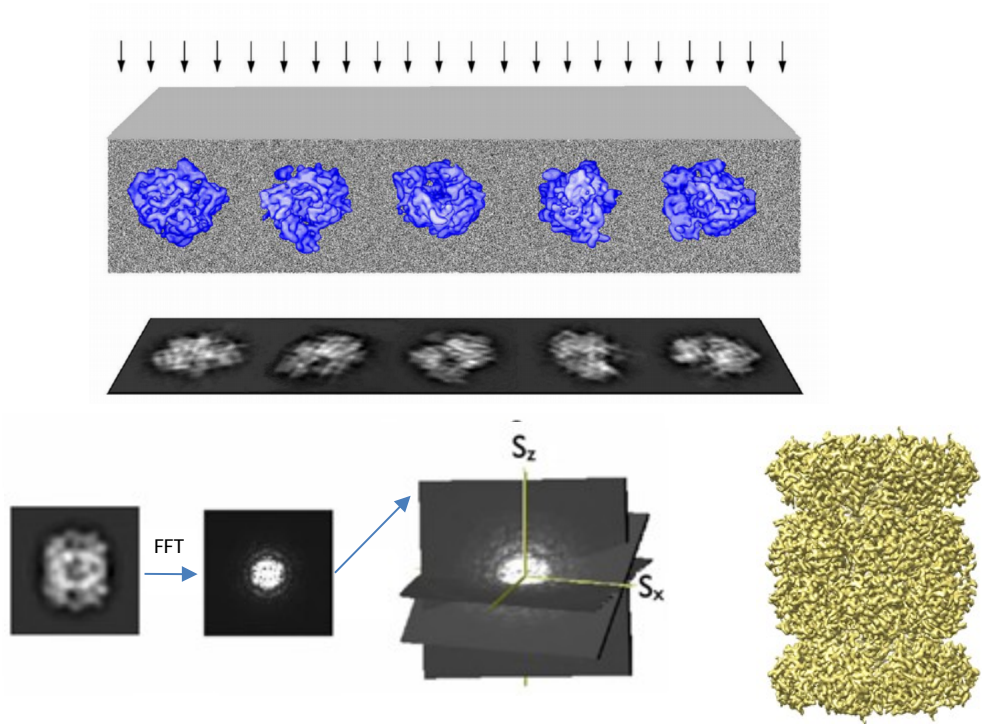
CSSB
Centre for Structural Systems Biology



Gun	Filament	Current (filament)
	Wehnelt cylinder Accelerator stack	Bias (emission) Voltage (high tension)
Condenser lens system		Gun shift, tilt Spot size Intensity Condenser stigmatism, size, centering
Objective lens system	“beam” deflectors Specimen Objective lens, stigmator, aperture	Z-height, tilt Focus Objective stigmatism, size, centering
Projector lens system		Image shift, tilt Magnification
Camera	CCD, Direct electron detector	Integrative mode, Electron counting mode

CSSB
Centre for Structural
Systems Biology

- **Projection Theorem:** Each recorded 2D image is a projection of the 3D shape we want to reconstruct

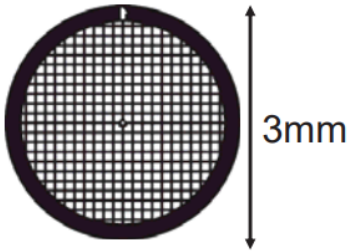


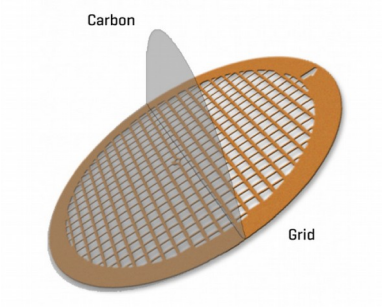
CSSB
Centre for Structural
Systems Biology

- CSSB Equipment
  - 1x FEI Talos L120C (120KV, CMOS Detector)
  - 1x FEI Talos Arctica (200KV, Direct Electron Detector)
  - 2x FEI Titan Krios (300KV, Direct Electron Detector)
    - FEI Falcon III Camera
    - Gatan K2 Camera
    - Phase Plate
- CSSB CryoEM Research Fields
  - Tomography
  - Single Particle Analysis

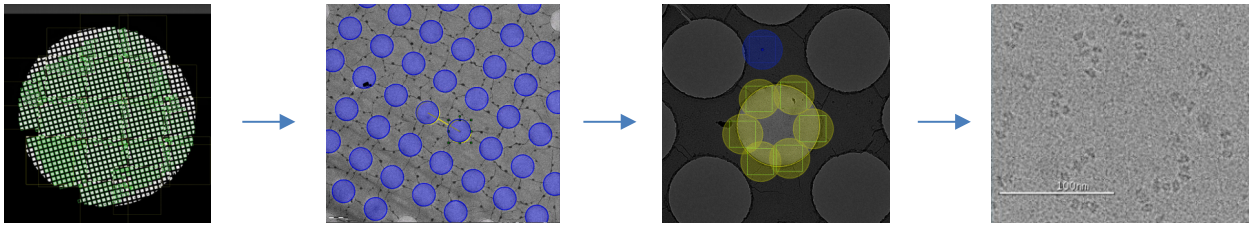
CSSB
Centre for Structural Systems Biology

- Grid Preparation



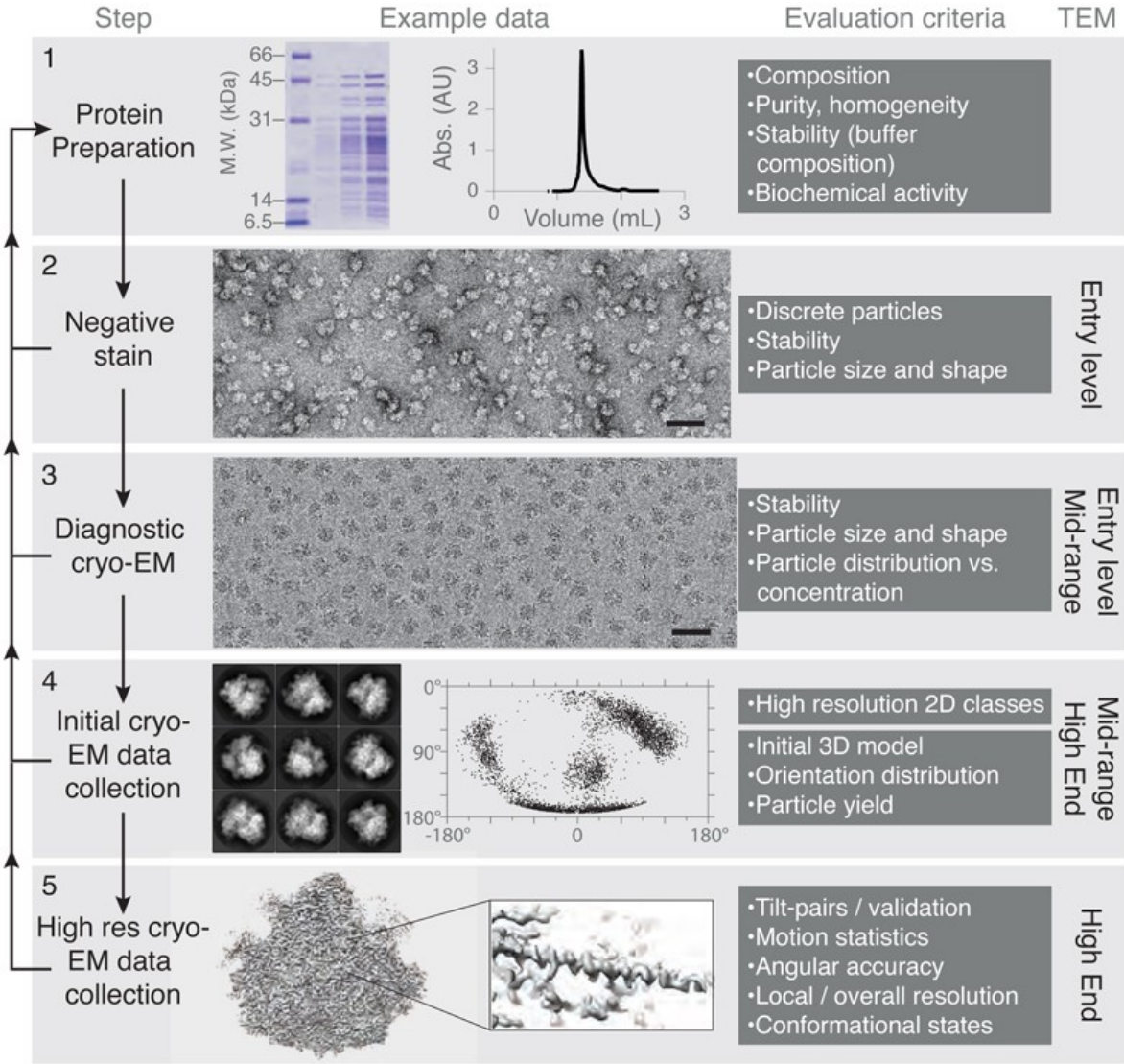


- Data Collection



CSSB
Centre for Structural Systems Biology

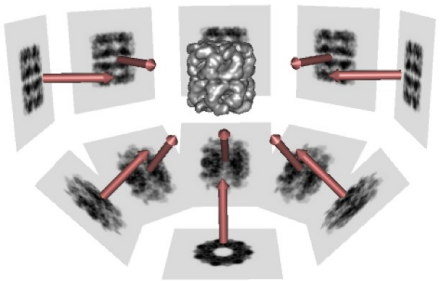
- Workflow



CSSB
Centre for Structural
Systems Biology

- Major Analysis Steps (GPU/CPU)
  - Motion correction
  - CTF estimation
  - Manual/Automated particle picking (Machine Learning)

  - 2D classification
  - 3D initial model
  - 3D classification
  - 3D refinement
  - Polishing

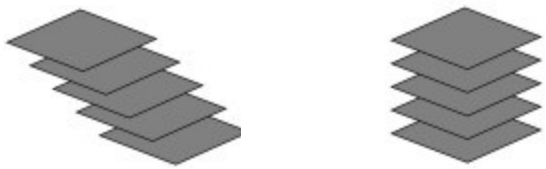


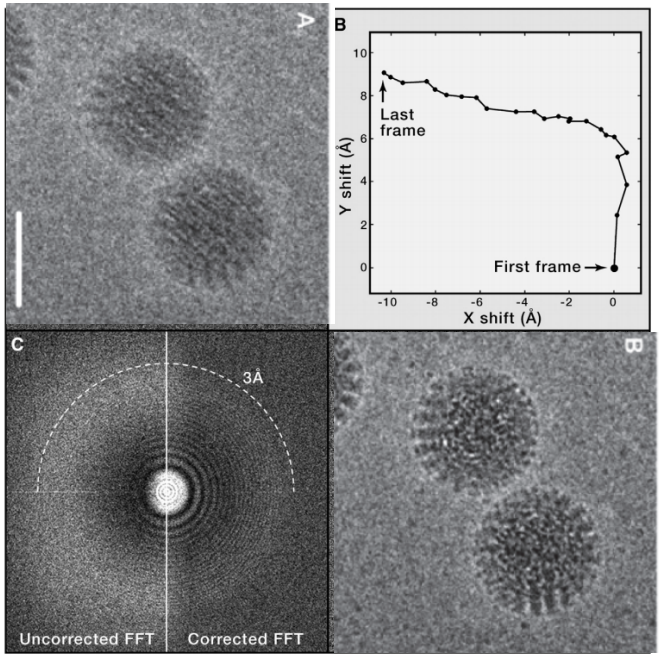
CSSB
Centre for Structural
Systems Biology

- Software suites for most analysis steps
  - Relion (CPU vectoring, GPU): Single Particle Analysis (SPA)
    - https://www2.mrc-lmb.cam.ac.uk/relion/
  - CryoSPARC2 (GPU): SPA, commercial
    - https://cryosparc.com/
  - EMAN2 (CPU): SPA, Tomography
    - https://blake.bcm.edu/emanwiki/EMAN2
  - SPHIRE (CPU): SPA, bundled with EMAN2, crYOLO
    - http://sphire.mpg.de/
  - Scipion (Wrapper): SPA, data conversion, uncommon external programs
    - http://scipion.i2pc.es/
  - IMOD (GPU, CPU): Tomography
    - http://bio3d.colorado.edu/imod/

CSSB
Centre for Structural
Systems Biology

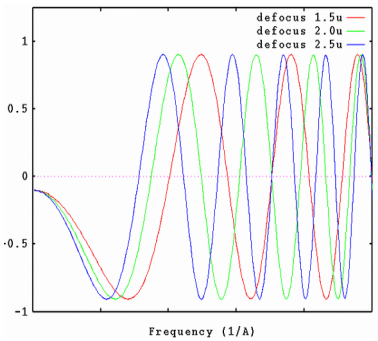
- # Motion Correction
- Beam induce movement correction
  - MotionCor2 (GPU, K40+)
    - Max. 1 min / movie stack
    - http://msg.ucsf.edu/em/software/
  - Unblur (CPU)
    - Max. 3 min / movie stack
    - http://grigoriefflab.janelia.org/unblur
  - Relion 3.0 beta (CPU)
    - No experience yet



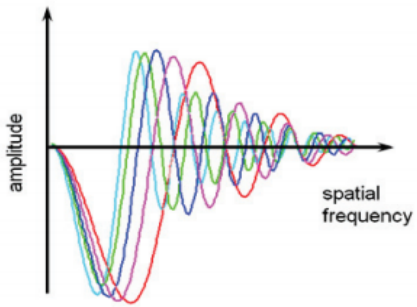


CSSB
Centre for Structural Systems Biology

- CTF Estimation
- Contrast Transfer Function
- All analysis <1min / mic

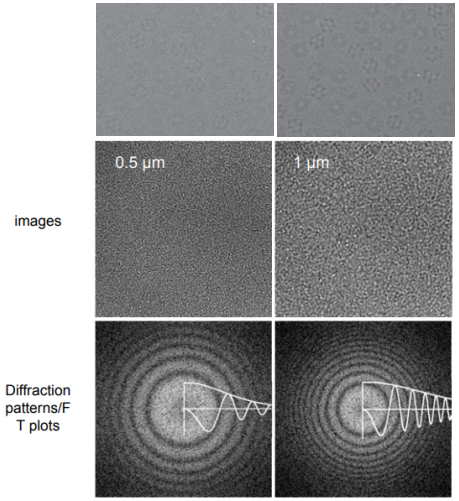


Perfect microscope



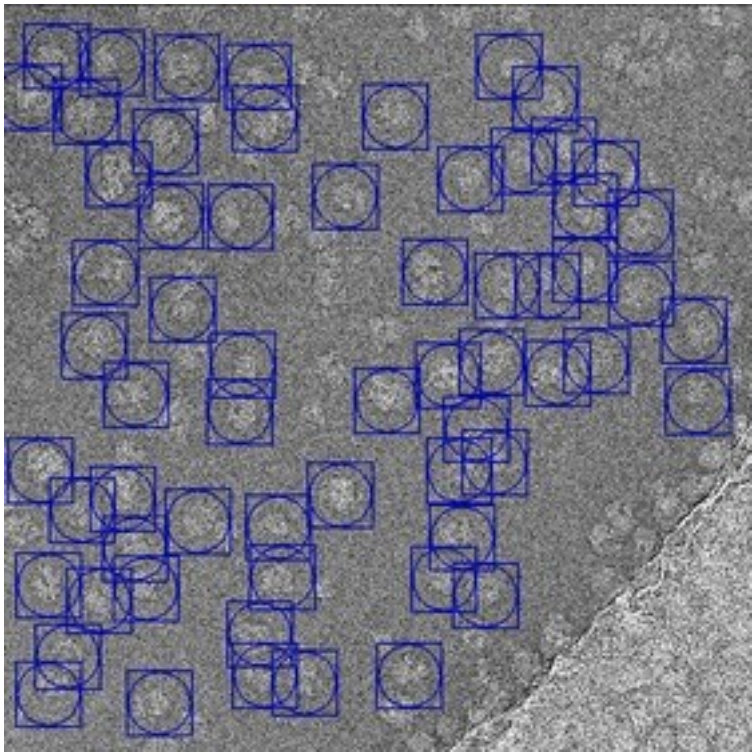
Real microscope

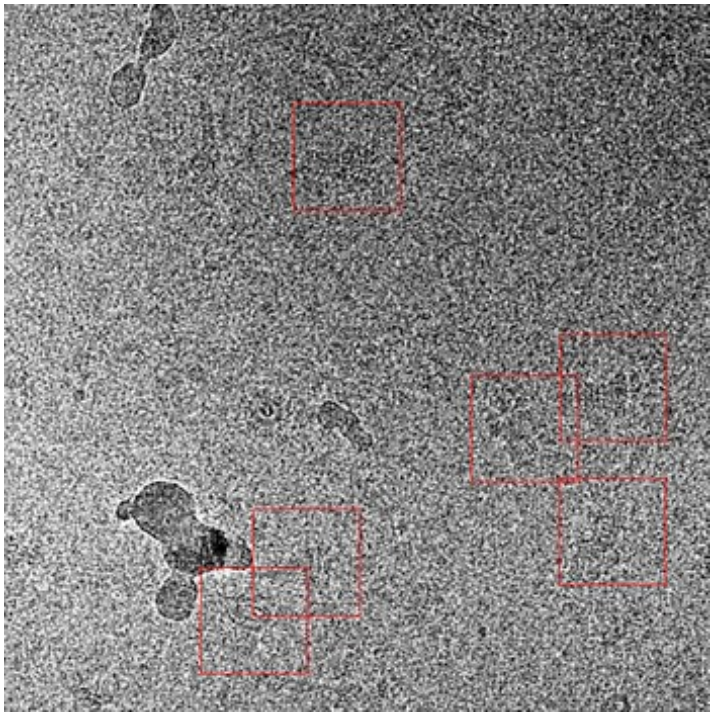
- CTFFind4 (CPU)
  - http://grigoriefflab.janelia.org/ctf
- Gctf (GPU, K20+)
  - https://www.mrc-lmb.cam.ac.uk/kzhang/Gctf/
- NovaCTF (CPU): 3D tomograms
  - https://www.embl.de/research/units/scb/briggs/NovaCTF/



CSSB
Centre for Structural
Systems Biology

- Manual Particle Picking





CSSB
Centre for Structural
Systems Biology

- Machine learning-supported picking
  - CrYOLO (http://sphire.mpg.de/, 4GB+ GPU)
    - Python-based: uses keras, tensorflow-gpu, cuDNN packages
    - Convolutional neural network (CNN)
    - Training needs about 10 hand-picked micrographs
    - Picks 6 micrographs per second, no sliding box/window approach
    - Authors collect and provide trained model database for various specimens
  - EMAN2 (version 2.22+, Aug. 2018)
    - New eboxer supports model training incl. tomograms
    - Currently buggy and no documentation on this topic

CSSB
Centre for Structural
Systems Biology

Incoming processing pipeline (what we want)

- Motion correction: MotionCor2 (15-30 min)
- CTF estimation: Gctf (15 min)
- Particle picking: crYOLO (15 min – no days/weeks of handpicking!)
- Particle extraction: Relion (5 min+ with 1 node)
- Preview ab-initio reconstruction: CryoSPARC2
  - Without initial model (2-4h, 800 micrographs 8-12 Angström)
  - With initial model (1-2h, 800 mics 6-7A)

- Real analysis (first cycle: 3-10d, permanently 2000 cores or 6-10 P100 GPUs, 800 mics 4.8A, cleaned 115.000 particles)

CSSB
Centre for Structural
Systems Biology

- Other Technical Considerations (DESY Maxwell)
  - Relion
    - 3D refinement needs 2 GPUs (2 independent half-map calculation)
    - P100 with 16GB up to 520 pixel boxsize, bigger boxes must use CPUs
    - Hand-choosen continue after “all” partition kick-out
  - CryoSPARC2
    - Webinterface & Master, Maxwell GPU client submissions
    - MongoDB as central database, files in user directories
    - Extremly fast, compiles special CUDA source during runtime
    - Commercial & closed source, free academic user license
  - More GPUs!!! Local SSDs for longer GPU runs!!!

CSSB
Centre for Structural
Systems Biology

- Thank you
  - Jiri Wald
  - The UKE / Marlovits group members
  - All CSSB people
  - All participants