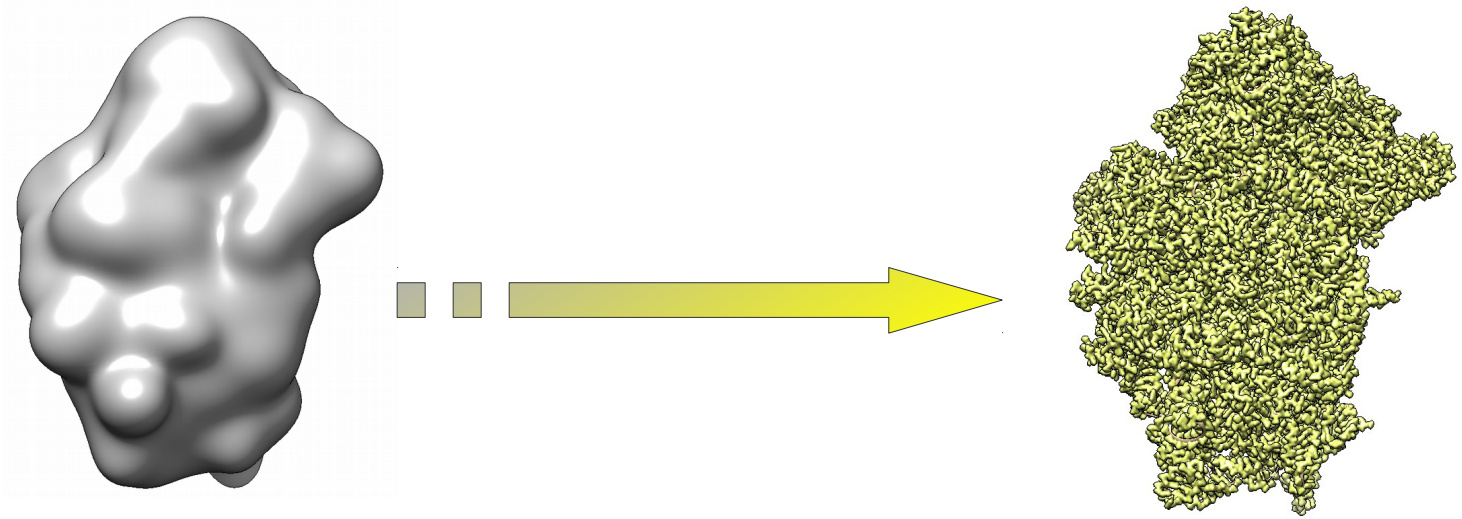
Cryo-EM, part. IV

From low-res blob... to atomic glory*(-ish)*

(and the many pittfalls to avoid)



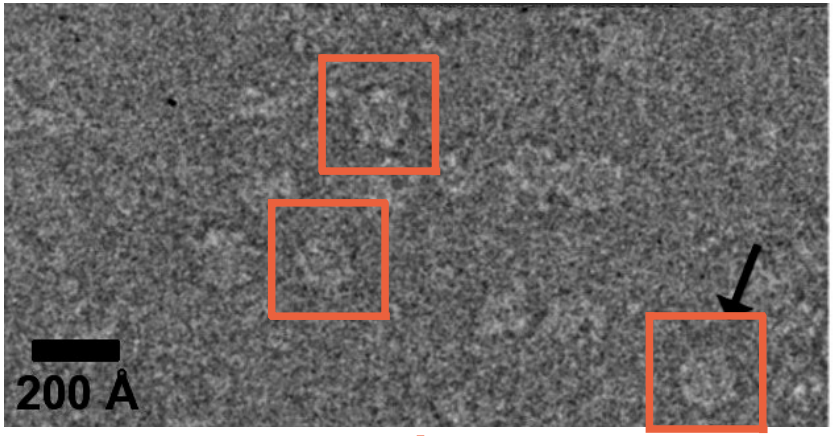
Single particle analysis :

- Workflow overview
- How to obtain an initial, low res volume
- Refinement, validation, interpretation of cryo-EM 3D maps

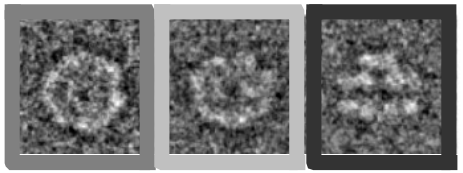
Célia Plisson-Chastang
celia.plisson@ibcg.biotoul.fr

Single particle analysis : General principles

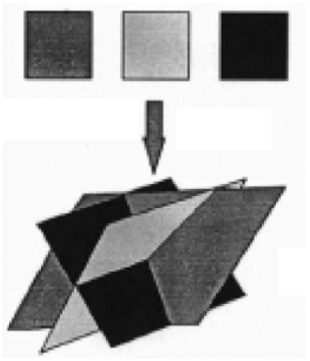
TEM images are 2D projections of randomly oriented 3D objects



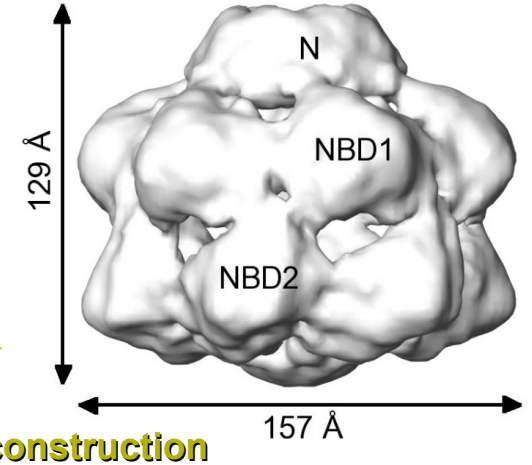
Particles Selection and Extraction



Alignment & Orientation Determination

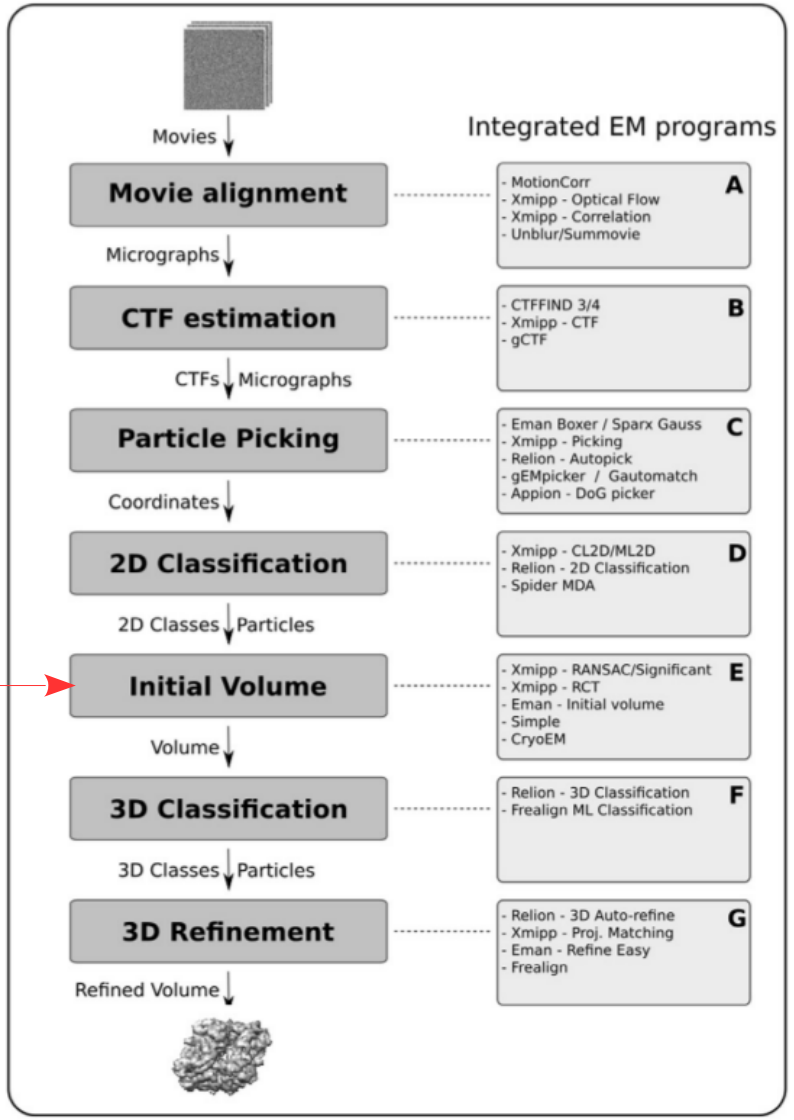


3D Reconstruction



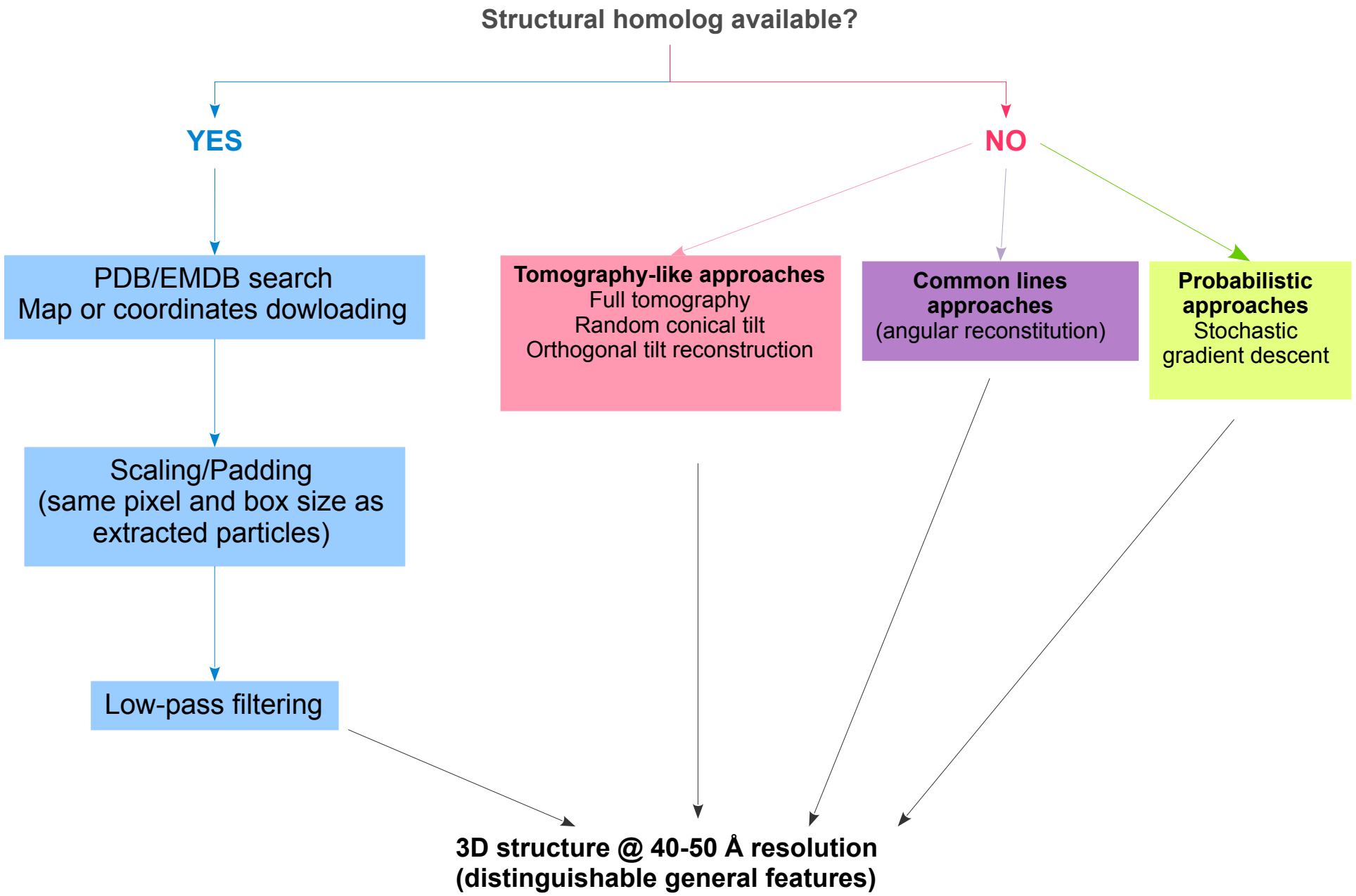
Single particle analysis : workflow

Structure 3D initiale (basse résolution) indispensable pour la suite des opérations



de la Rosa-Trevin et al., 2016, DOI: http://dx.doi.org/10.1016/j.jsb.2016.04.010

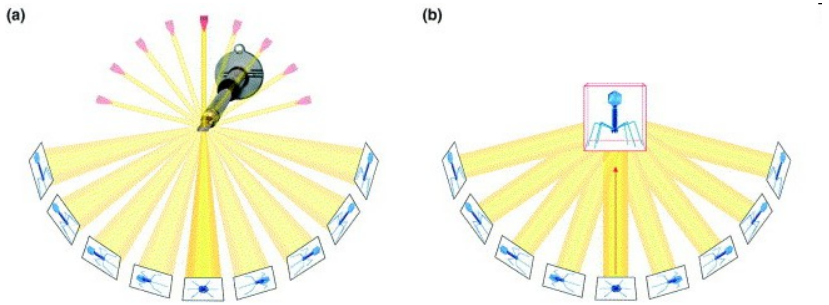
First 3D reconstruction : how do we get an initial, low res 3D volume?



Tomography-like approaches :

"Tomography allows to reconstruct essentially everything which generates some sort of contrast."

(Andy Hoenger, EMBO course on 3D cryo-EM, Heidelberg, August 2004)



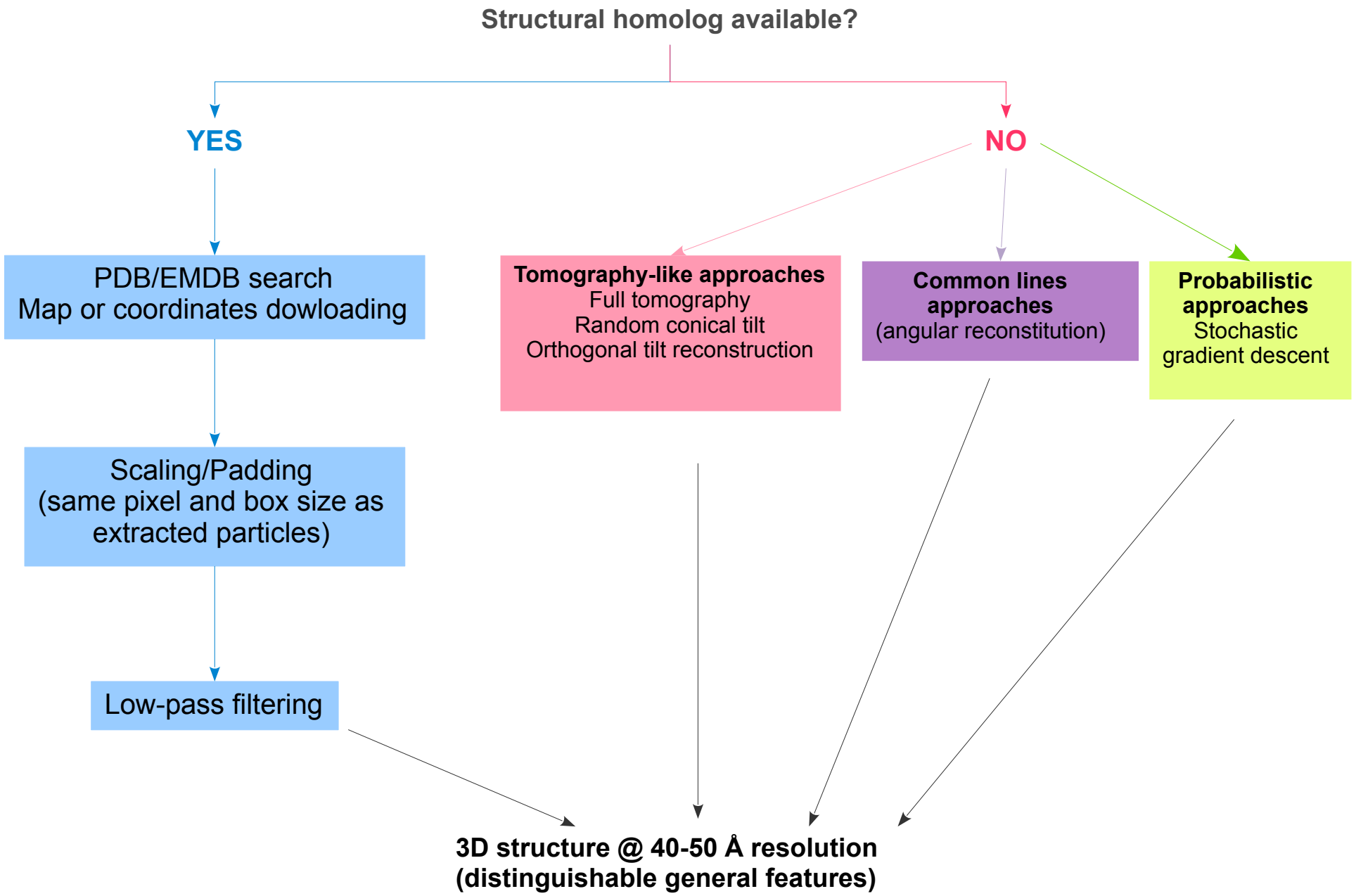
On the plus side…

No prior structural knowledge required

On the minus side…

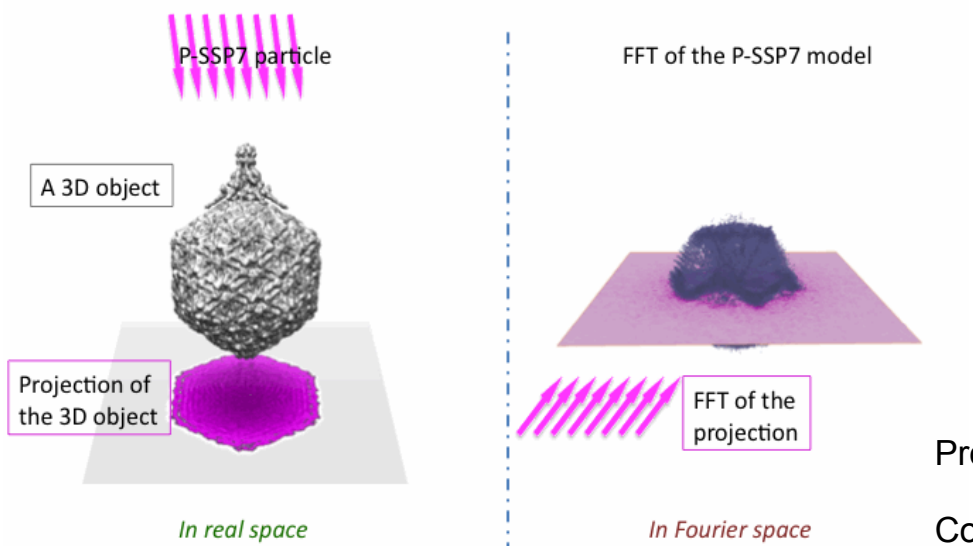
Technically challenging (the smaller the object, the more challenging)
RCT and ORT almost always realized on negatively stained preps
Missing wedge/missing cone => deformation can happen

First 3D reconstruction : how do we get an initial, low res 3D volume?

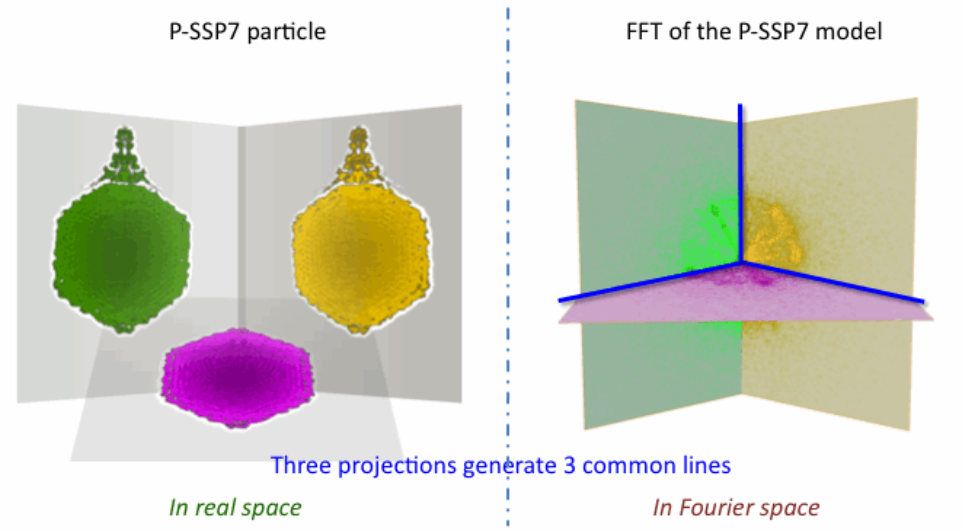


Common lines approaches (angular reconstitution)

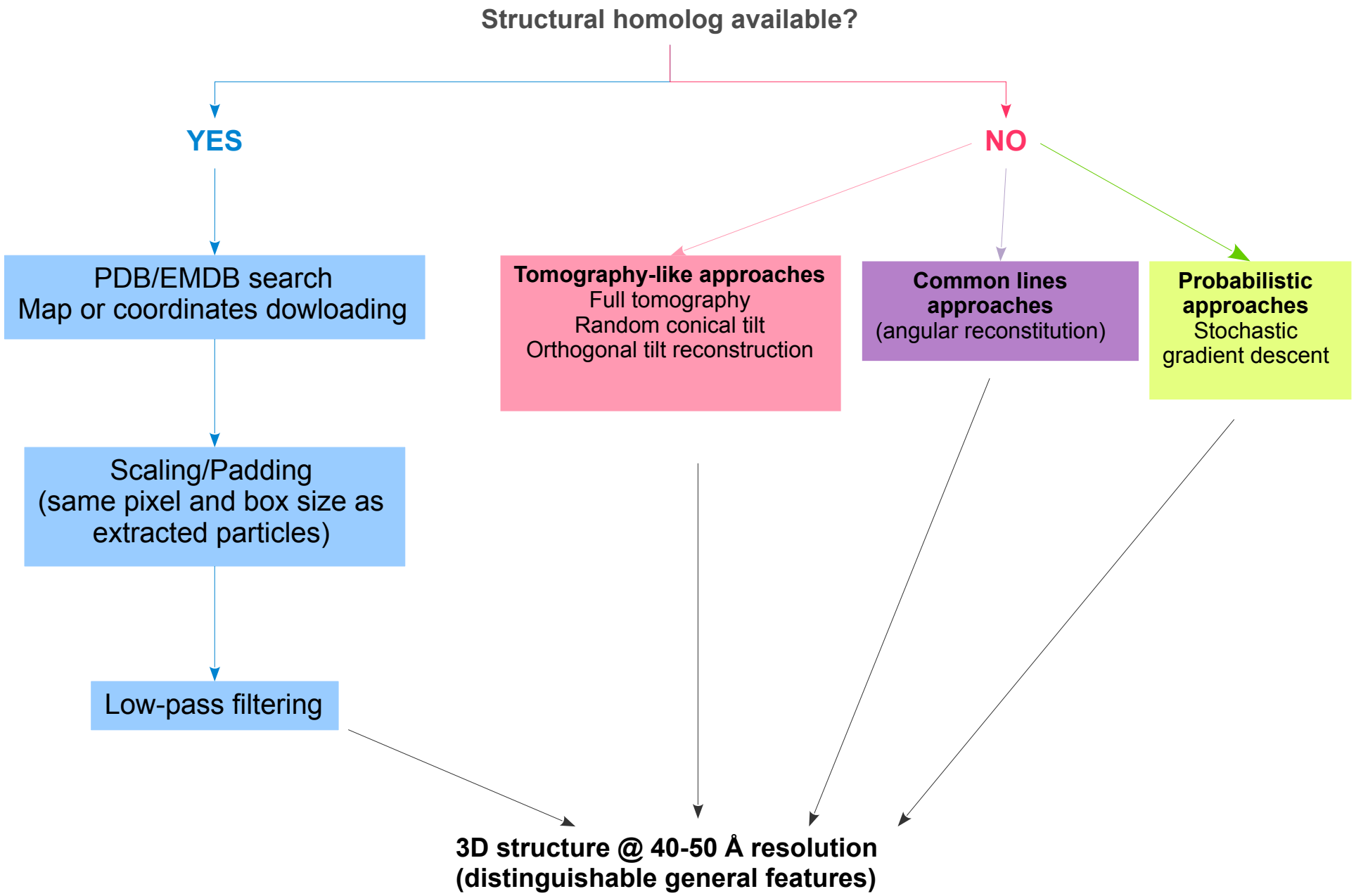
Central section theorem : The 2D Fourier Transform of the projection of a 3D object is a central section (a section passing through the origin) of the 3D Fourier transform of this object.



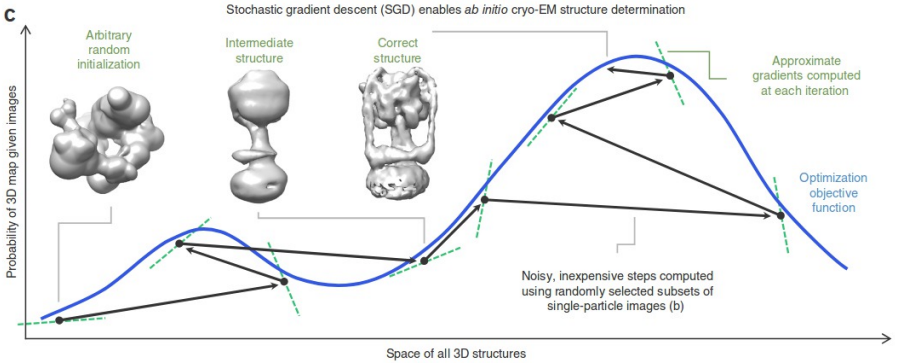
Pros : Works nicely for high symmetry objects (viruses)
Image acquisition straightforward (no tilted pairs)
Cons: For low symmetry objects, can be user-biased



First 3D reconstruction : how do we get an initial, low res 3D volume?



Probabilistic approaches
Stochastic gradient descent (SGD) : Relion, cryoSPARC



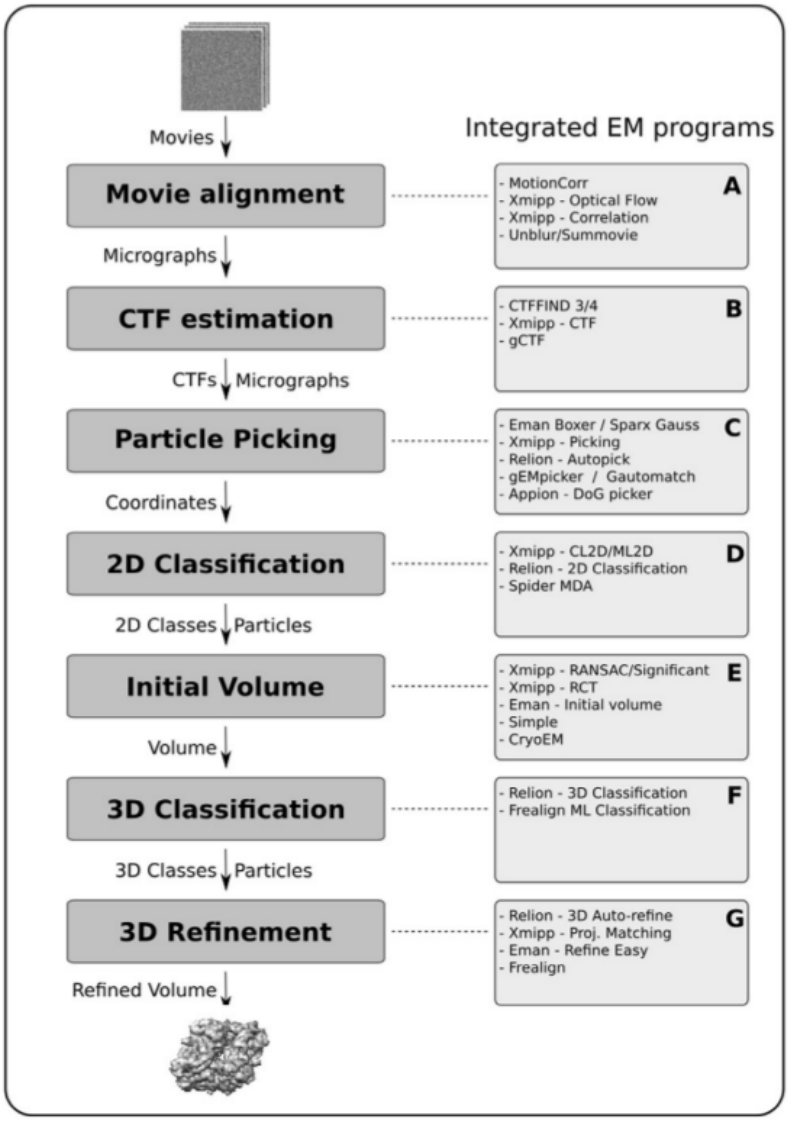
Punjani et al., Nature Methods2017

Pros : No prior structural knowledge required
Does not require user intervention/(in)experience
Very fast method
CryoSPARC can sort out 3D classes at this step

Cons : (Relion)
even distribution of viewing directions
data good enough to yield detailed 2D classes

Single particle analysis : workflow

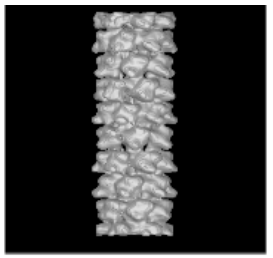
Projection-matching based techniques



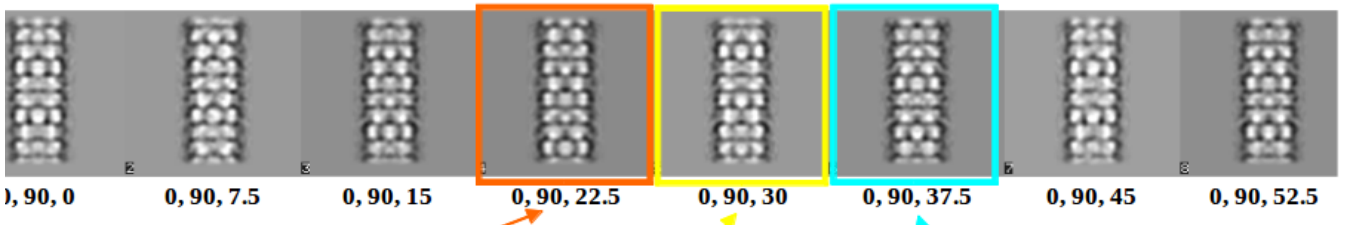
de la Rosa-Trevin et al., 2016, DOI: http://dx.doi.org/10.1016/j.jsb.2016.04.010

Orientation determination by projection matching

Particles are compared and aligned to all the 2D projections of an existing 3D model
Orientation is given by the reprojection giving the best correlation coefficient with the considered particle



3D Model



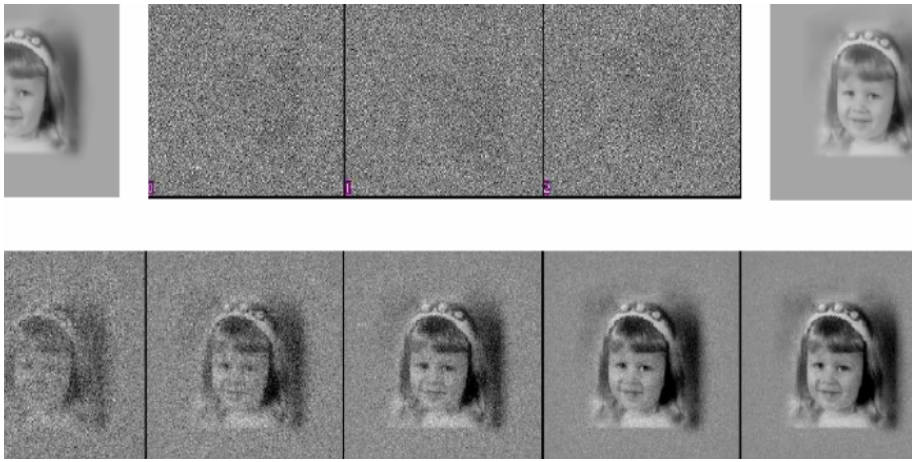
2D (re)projections
of the 3D model
known orientations

Euler angles

Raw images
Or 2D classes

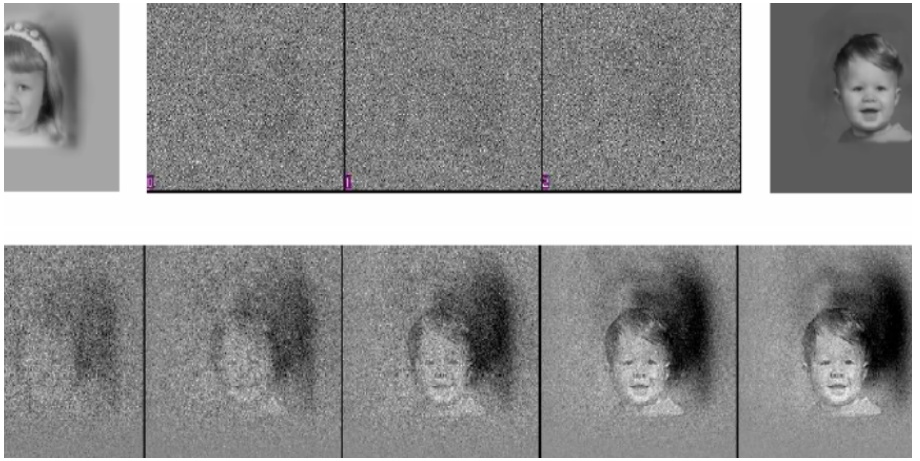
Data Overfitting / Model Bias (aka « Einstein from noise »)

« Ideal » case

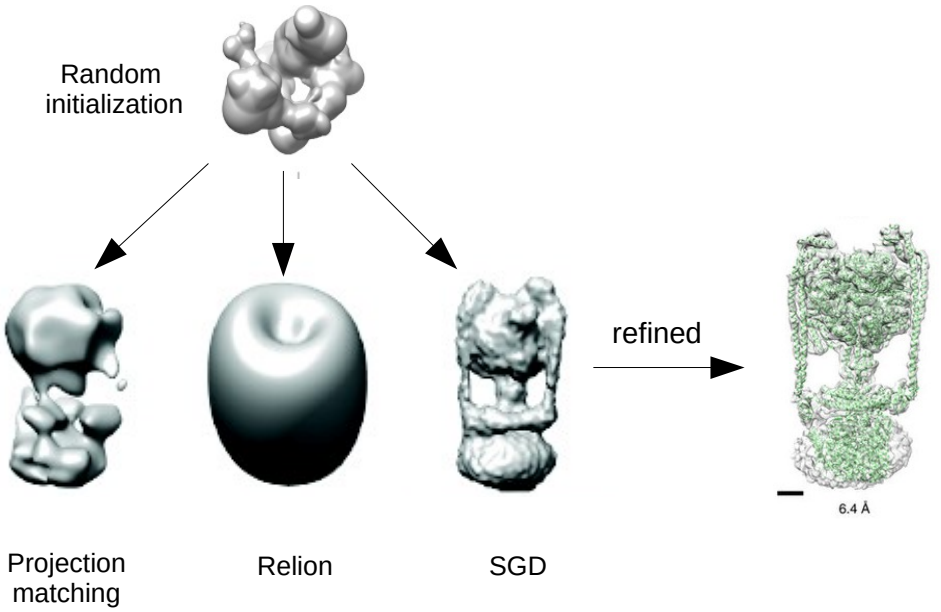


White noise

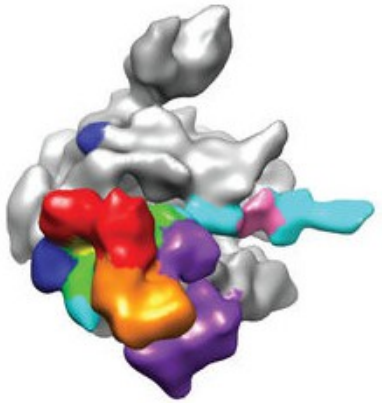
Wrong model



Bad model in => bad model out

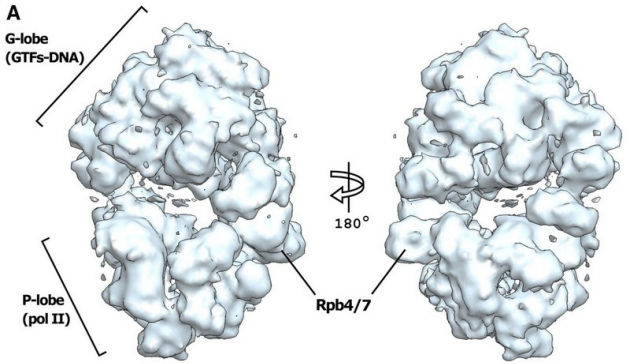


Human RNA Pol II PIC



He et al. & Nogales, Nature 2013

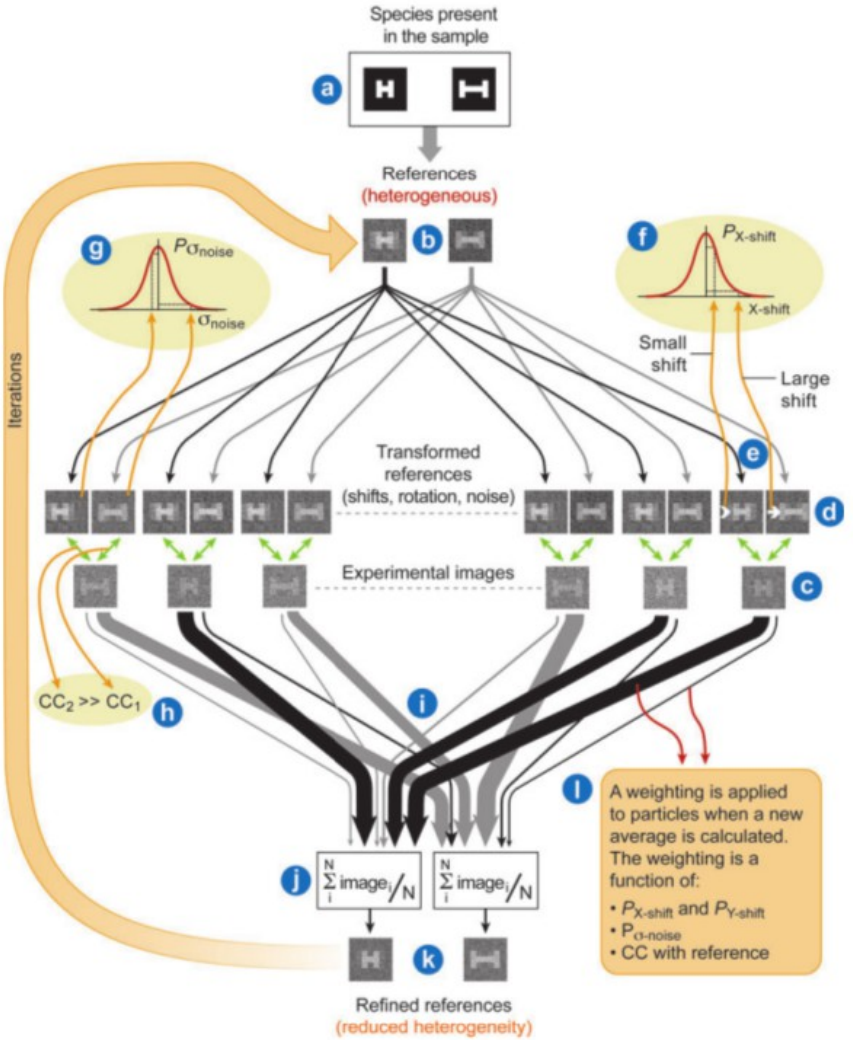
Yeast RNA Pol II PIC



Murakami et al. & Kornberg, Science 2013

Orientation determination: PM + probabilistic approaches (maximum likelihood)

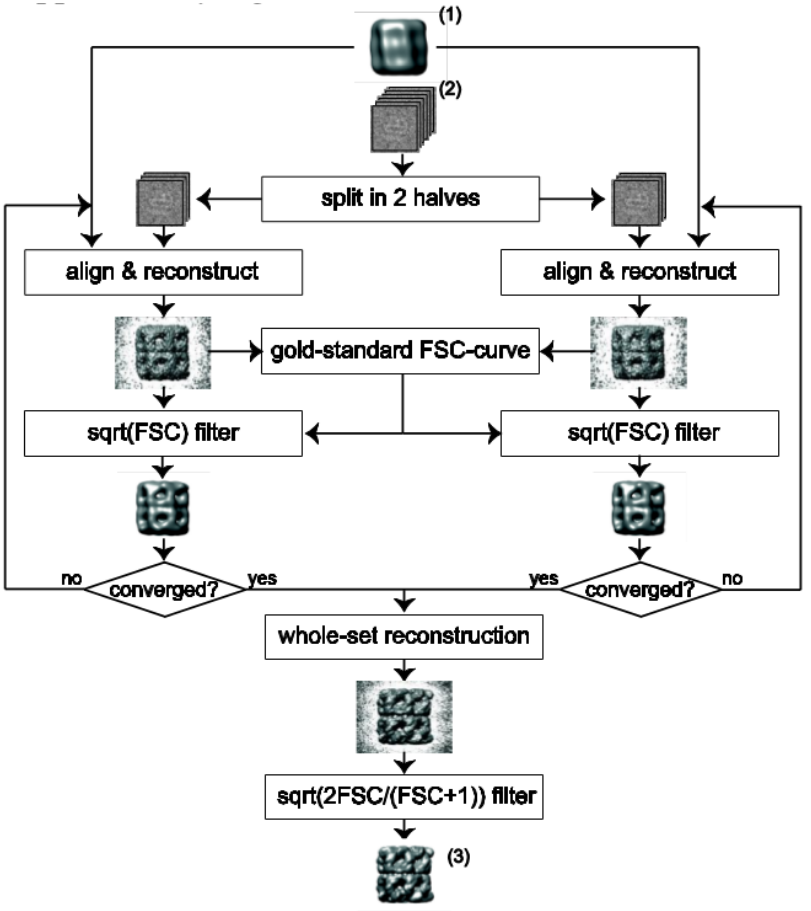
Particles orientation are given by projection matching, but a weighting is applied to each particle before 3D reconstruction. “Good” particles have more weight than “bad” ones. This can at least partly prevent overfitting.



Leschziner and Nogales, 2007. DOI: 10.1146/annurev.biophys.36.040306.132742

3D Refinement

Iterative ML projection matching from an initial, low resolution model
Goal : align and determine orientation of particles with the highest possible precision degree
At the beginning, the dataset is randomly divided into two groups, refined separately from each other
The 3D reconstruction resulting from a cycle of MLPM will be used as reference for the next one

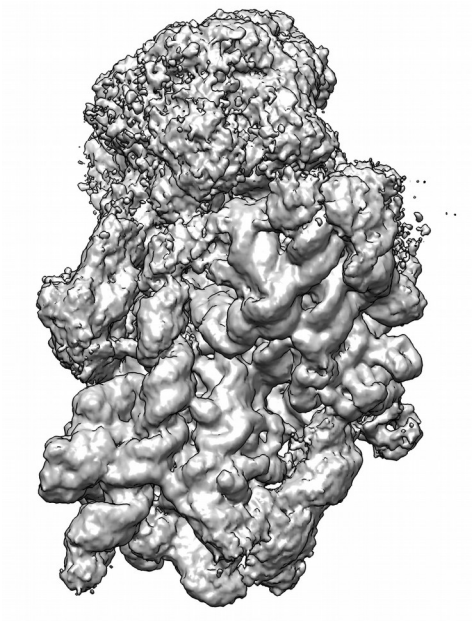


At the end of 3D refinement

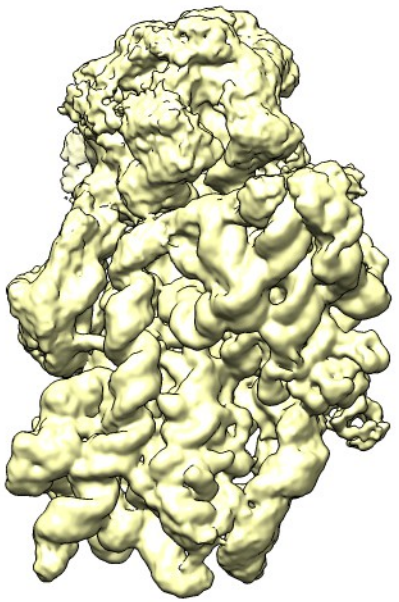
Calculate a full (combined half datasets) map

Estimate map resolution

Low-pass filter the full map according to resolution



Unfiltered half map1



Full filtered map

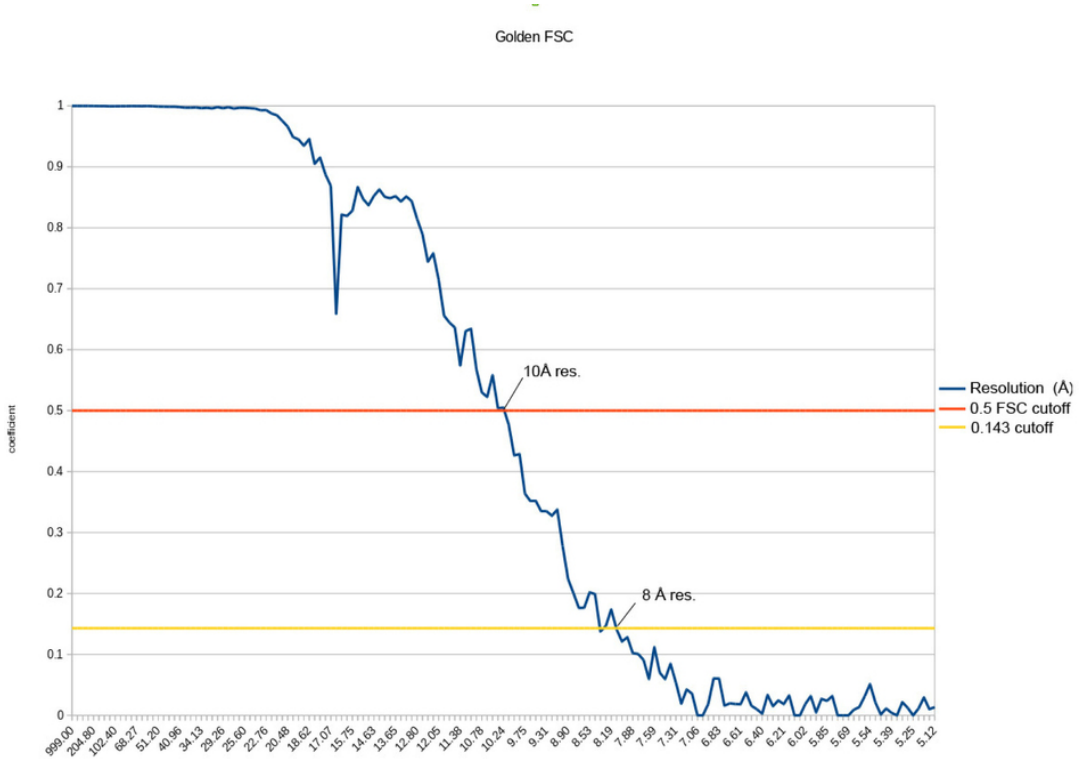
Resolution Estimation:
Fourier Shell Correlation

The aligned & oriented particles dataset is randomly divided in two, and two 3D reconstructions are calculated

Both 3D independent “half maps” are compared in Fourier space

The resolution of the whole 3D reconstruction corresponds to the spatial frequency where correlation coefficient between half maps drops below a value X

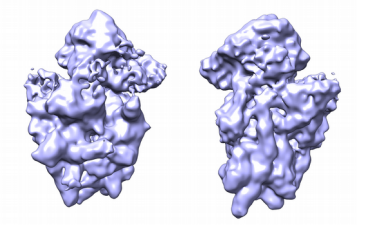
X was 0.5 for a long time, and has been lowered to 0.143, according to Rosenthal et Henderson, *j mol biol* 2003



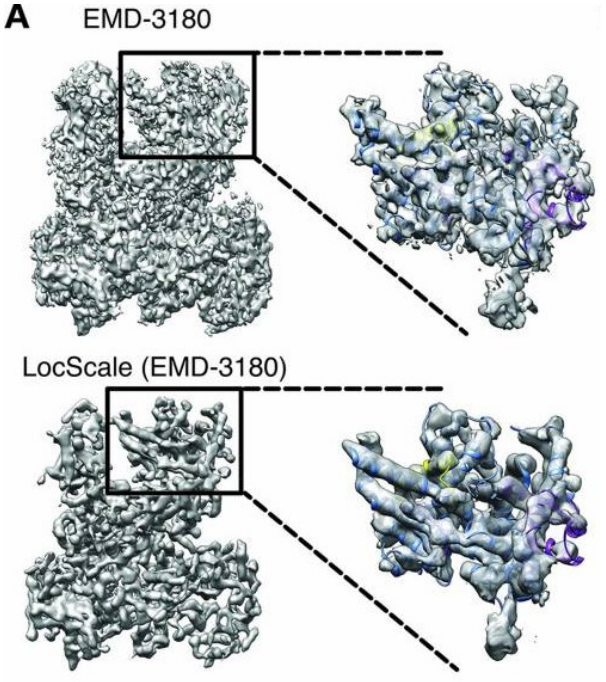
C_{Ref} of 0.5 corresponds to an FSC of 0.143

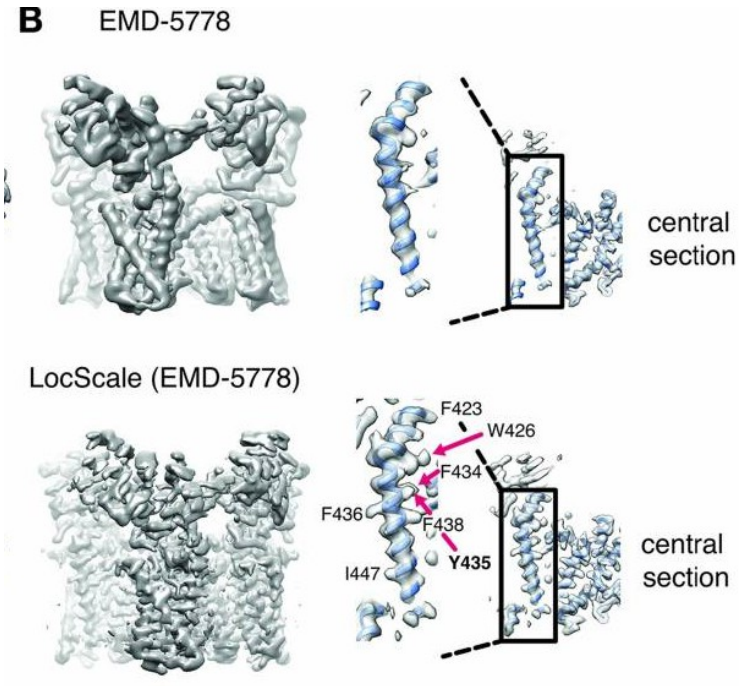
C_{Ref} : coef. between a perfect map without noise and an experimental map reconstructed from the full dataset

Refined 3D Map : what now?

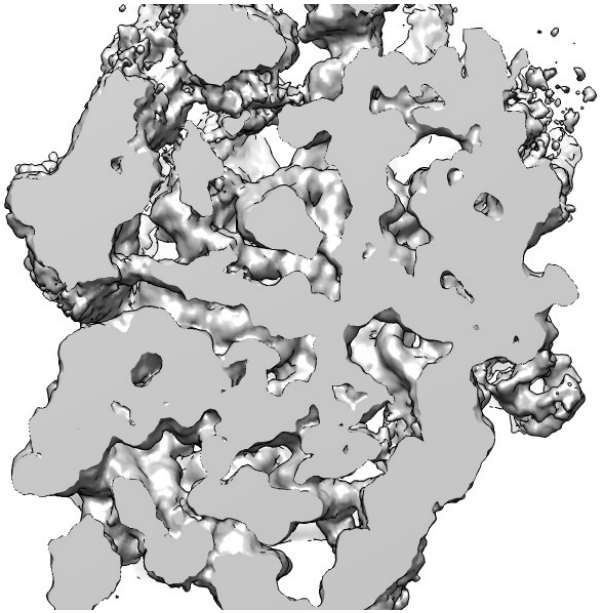


- Local 3D Classification / Refinement
- Map sharpening (global/local) : amplitudes correction
  so that low-res & high-res terms have correct relative scaling

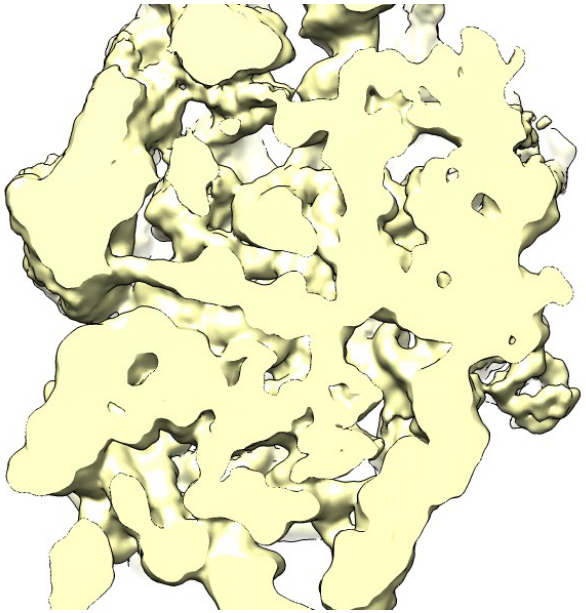




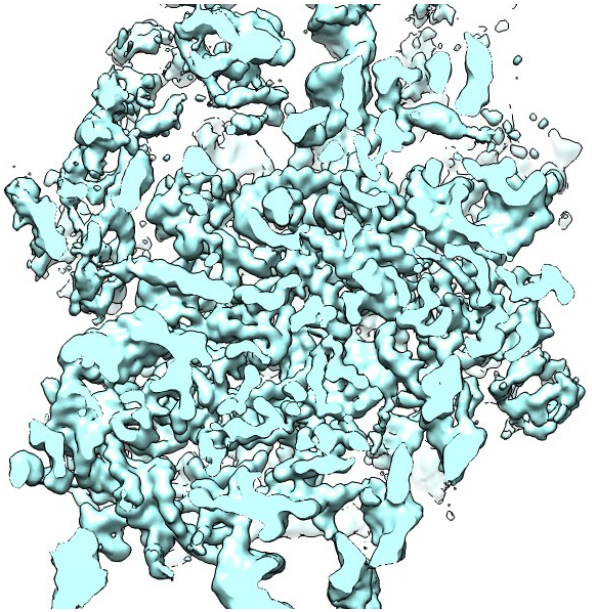
Filtering vs. sharpening



Unfiltered half map1



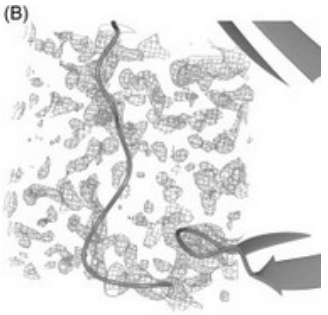
Full filtered map



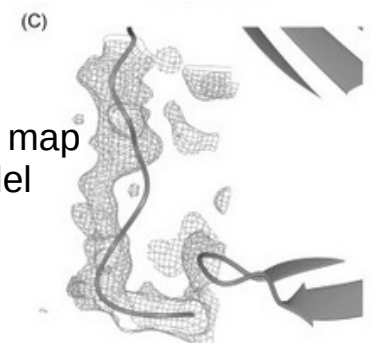
Full filtered+sharpened map

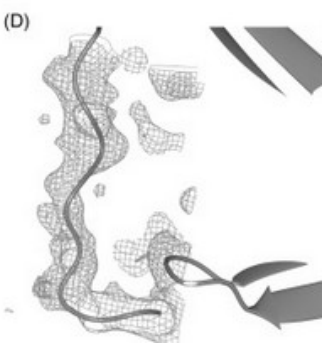
Map oversharpening effects

Originally published
map +model



Re-sharpened map
+ (wrong) model

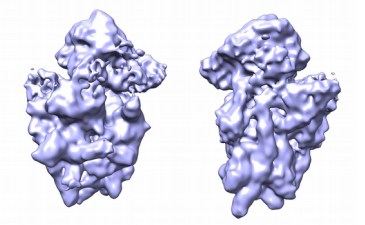




Re-sharpened map
+ corrected model

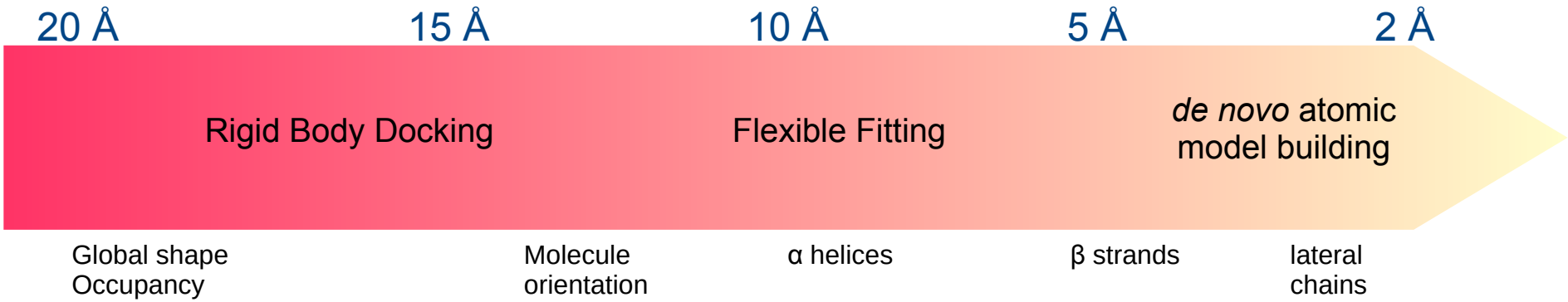
Murshudov, 2016

Refined 3D Map : what now?



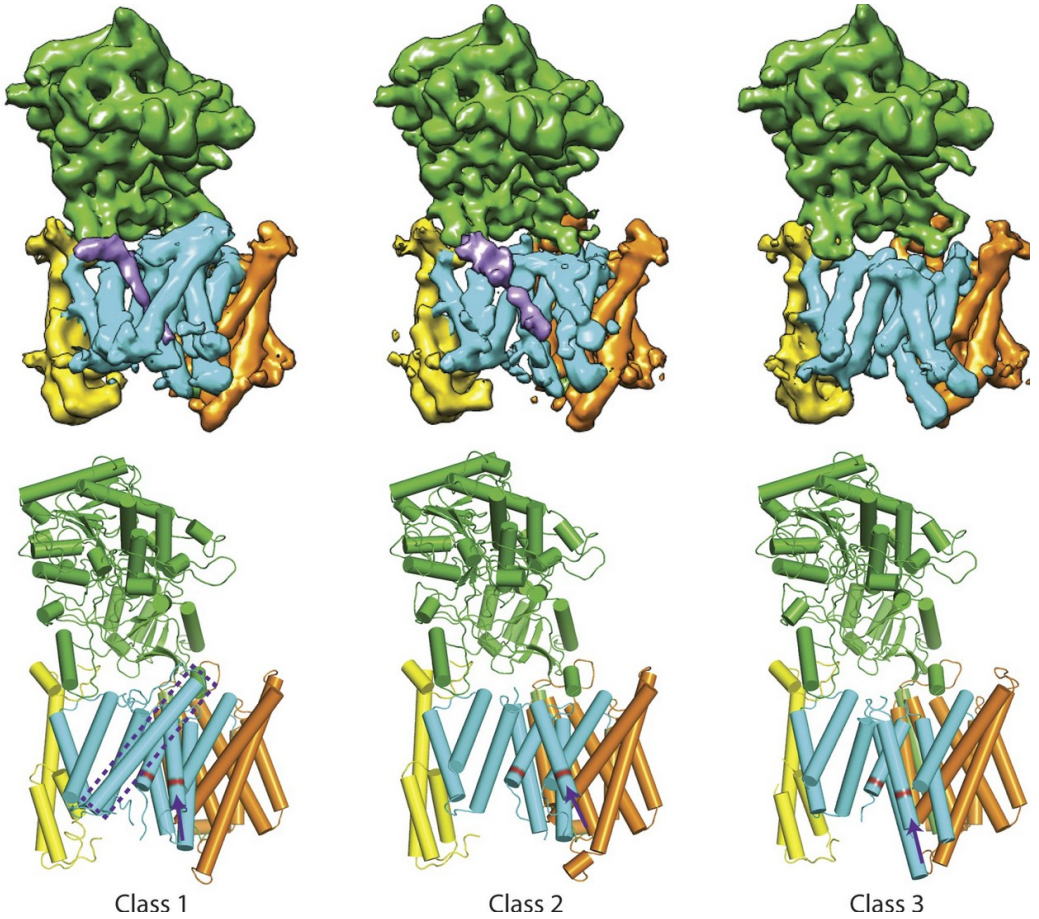
- Local 3D Classification / Refinement
- Map sharpening (global/local) : amplitudes correction
  so that low-res & high-res terms have correct relative scaling
- Map Interpretation : fitting techniques depends on map resolution

True for maps @10 Å resolution and better !



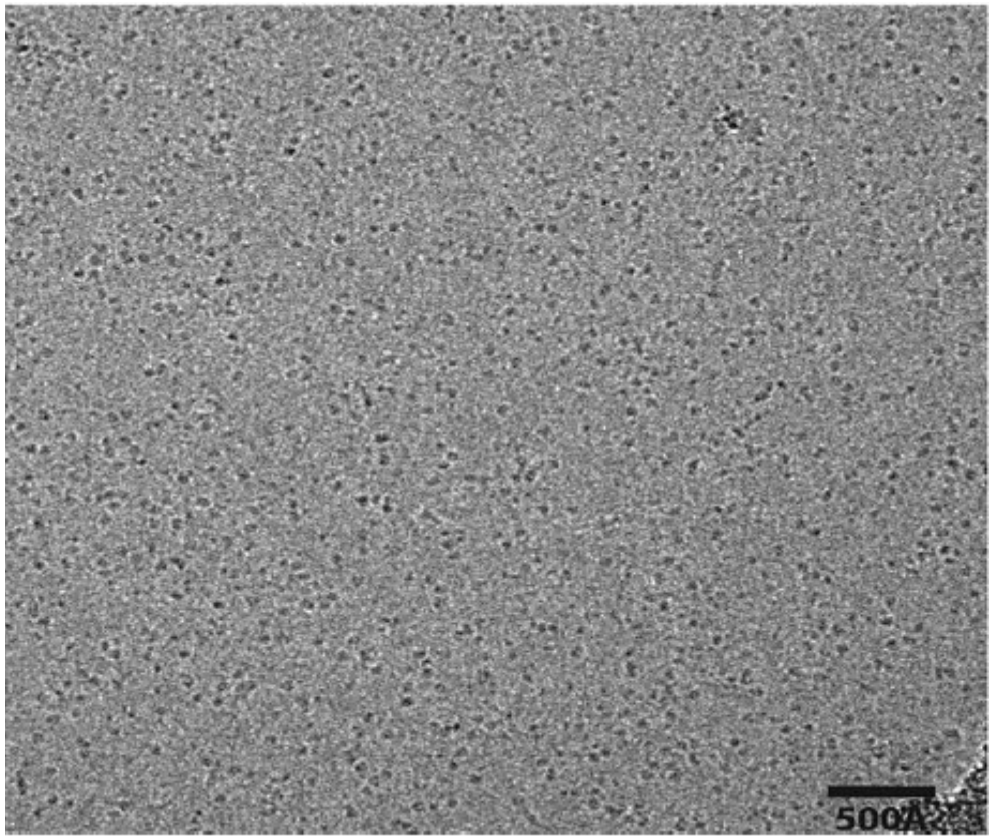
Hétérogénéité structurale / états conformationnels

Gamma secretase : disorder in the catalytic site of presenilin

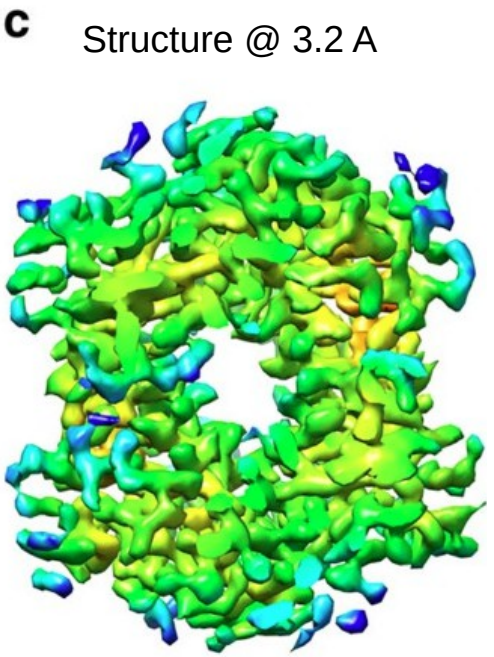


Bai et al., 2015
DOI: http://dx.doi.org/10.7554/eLife.11182

Single particle analysis : breaking the 100kDa



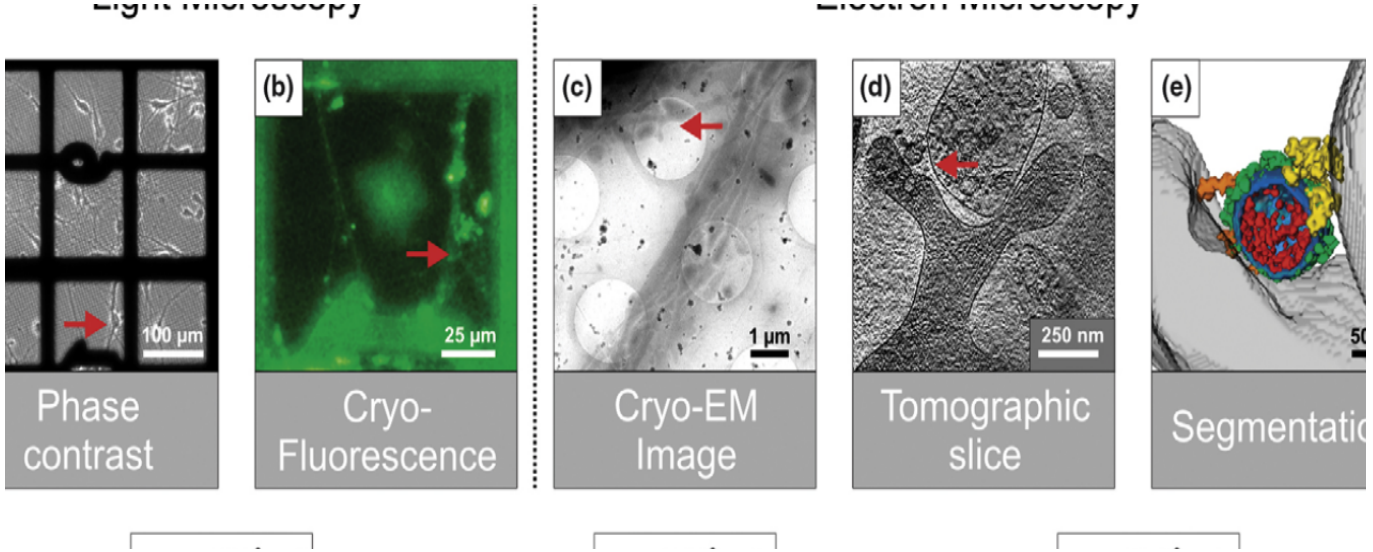
Hemoglobin (64 kDa)



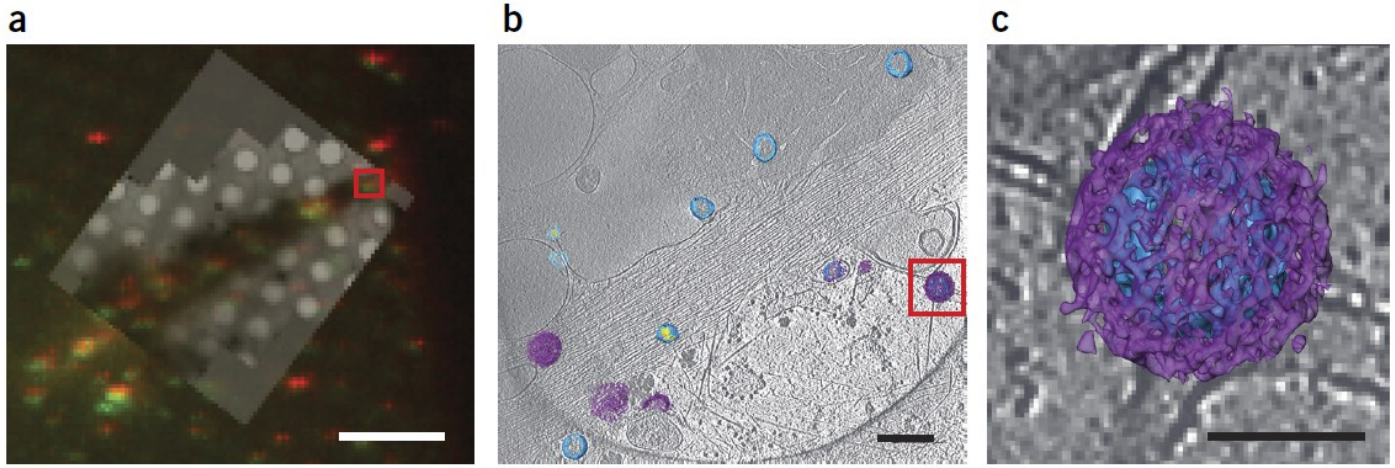
Koshouey et al., 2017

Contraste très faible en cryo-TEM : il était difficile de visualiser des objets inférieurs à 5 nm de diamètre (~150 kDa). Nouvelles technologies (caméra à détecteur direct d'électrons, phase plate) permettent de visualiser des objets de plus en plus petits (< 100 kDa)

Microscopies Corrélatives



Plitzko *et al.*, 2009



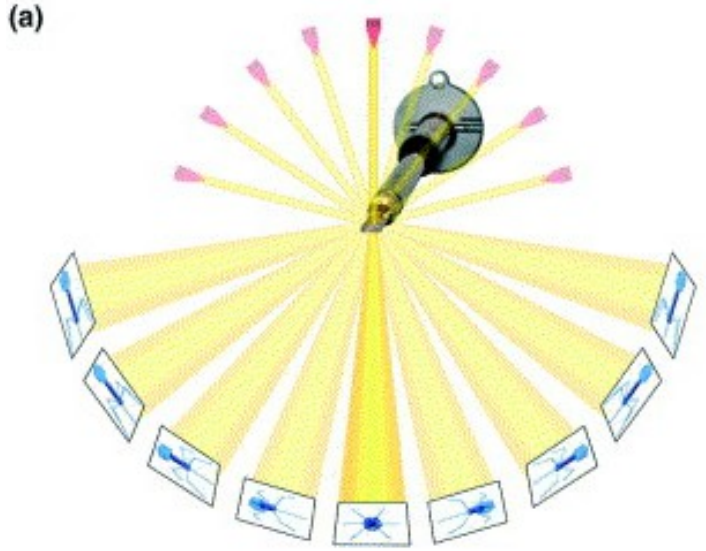
Hampton *et al.*, 2016

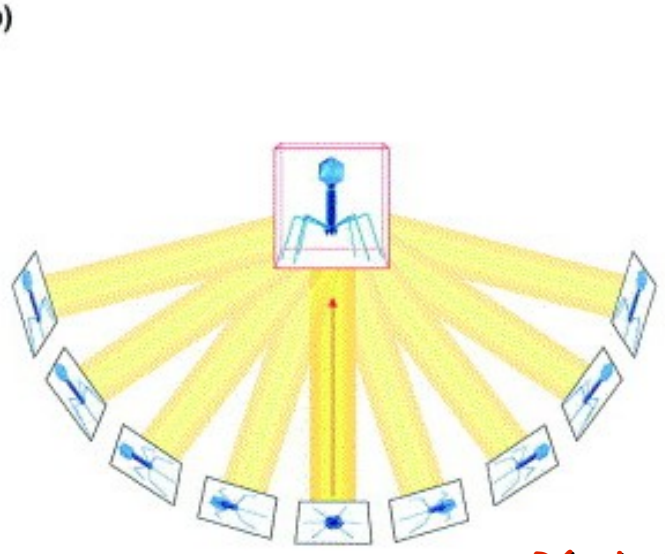
Figure 10 | Cryo-CLEM imaging of retroviral endocytosis and fusion. (**a**–**c**) Double-labeled HIV-1 particles pseudotyped with avian sarcoma and leukosis virus (ASLV) Env glycoprotein. (**a**) Cryo-CLEM of ASLV Env pseudotyped HIV-1 particles bound to CV-1/TVA950 cells. Central region is the overlay of the cryo-EM montage onto the cryo-fLM image. Red square indicates the tomography data in **b** and **c**. (**b**) Tomographic slice, with segmentation, of ASLV Env pseudotyped HIV-1 particles undergoing endocytosis. The viral membrane (light blue) and mature core (yellow) are rendered. Clathrin cages (purple) surround several viral particles. (**c**) Enlargement of one clathrin cage (purple) surrounding a viral particle (light blue). Scale bars, (**a**) 50 µm, (**b**) 250 nm, (**c**) 100 nm.

Tomographie électronique

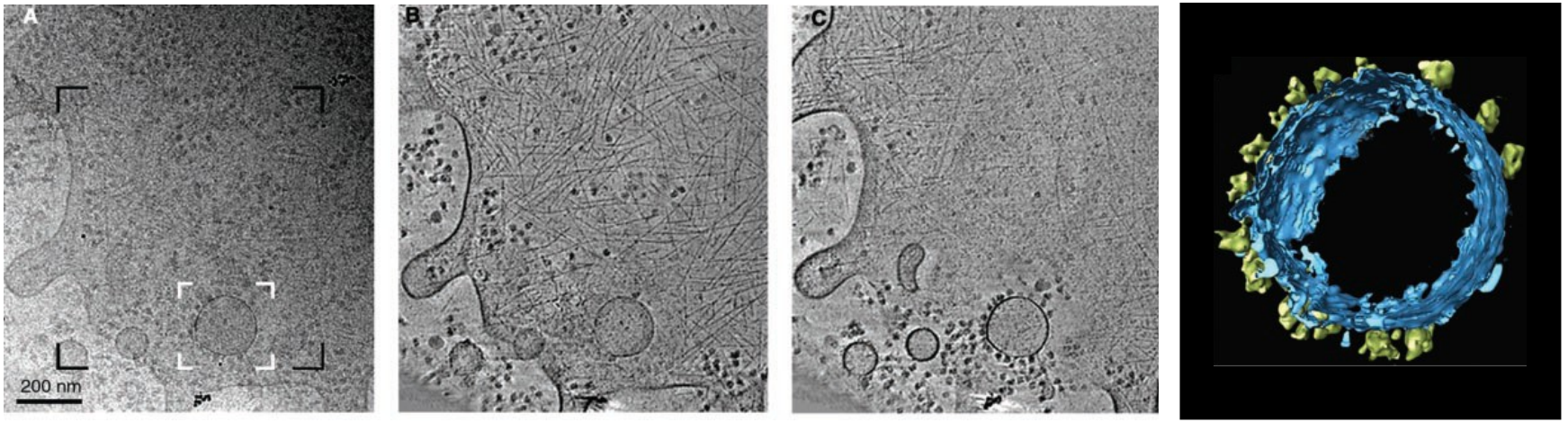
Série de projections inclinées

Reconstruction d'un volume 3D par rétro-projection





Résolution 2 – 5 nm



Medalia *et al.*, 2002

Subtomogram averaging

Parts of tomograms are extracted, aligned and averaged using SPA approaches

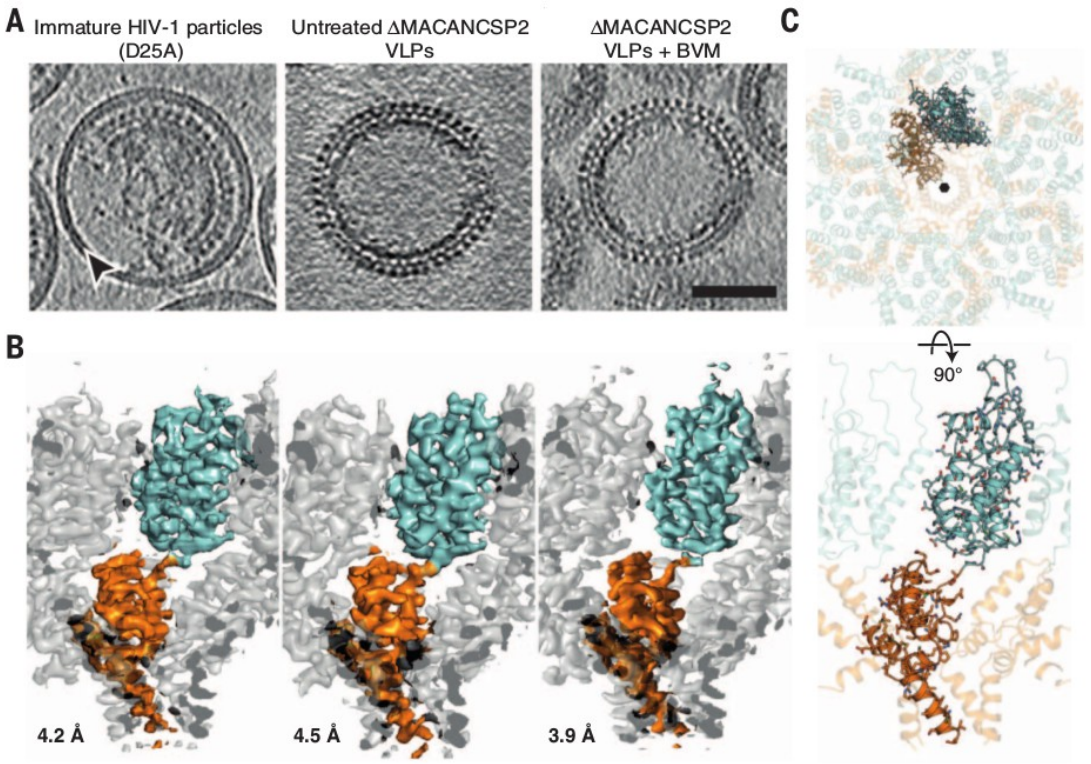
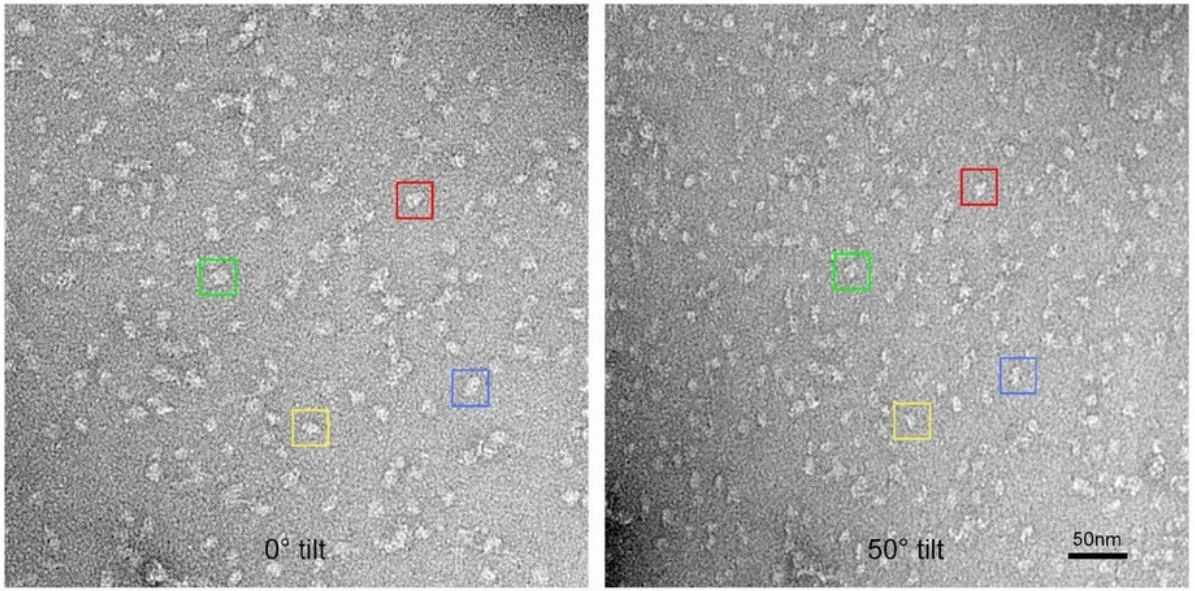


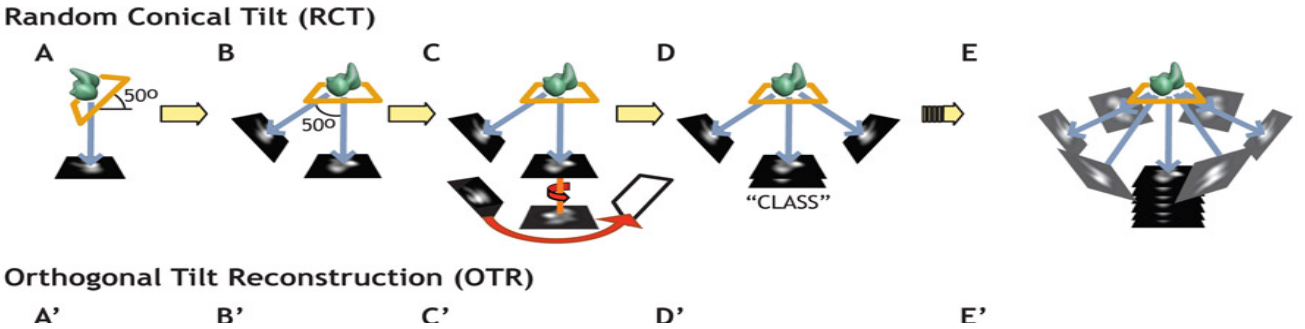
Fig. 1. Structure of the immature HIV-1 CA-SP1 lattice at 3.9 Å. (**A**) Computational slices through tomograms of immature HIV-1 (D25A mutant), untreated ΔMACANCSP2 VLPs, and BVM-treated ΔMACANCSP2 VLPs. The arrowhead marks the membrane bilayer in the left panel. Scale bar, 50 nm. (**B**) Electron densities of CA-SP1 from the samples shown in (A), viewed perpendicular to the lattice, generated by subtomogram averaging. One CA-SP1 monomer is highlighted in color, with the CA-NTD in cyan and the CA-CTD and SP1 in orange. The resolutions of the determined structures are noted. (**C**) The refined atomic model, viewed from outside of the virus (top) and rotated by 90°, shown in the same view as in (B) (bottom). The sixfold symmetry axis is indicated with a hexagon.

Schur et al., *Science* 2016 DOI: 10.1126/science.aaf9620

Tomography-like approaches : random conical tilt, orthogonal tilt reconstruction

2 images of the same particles field, one tilted compaired to the other



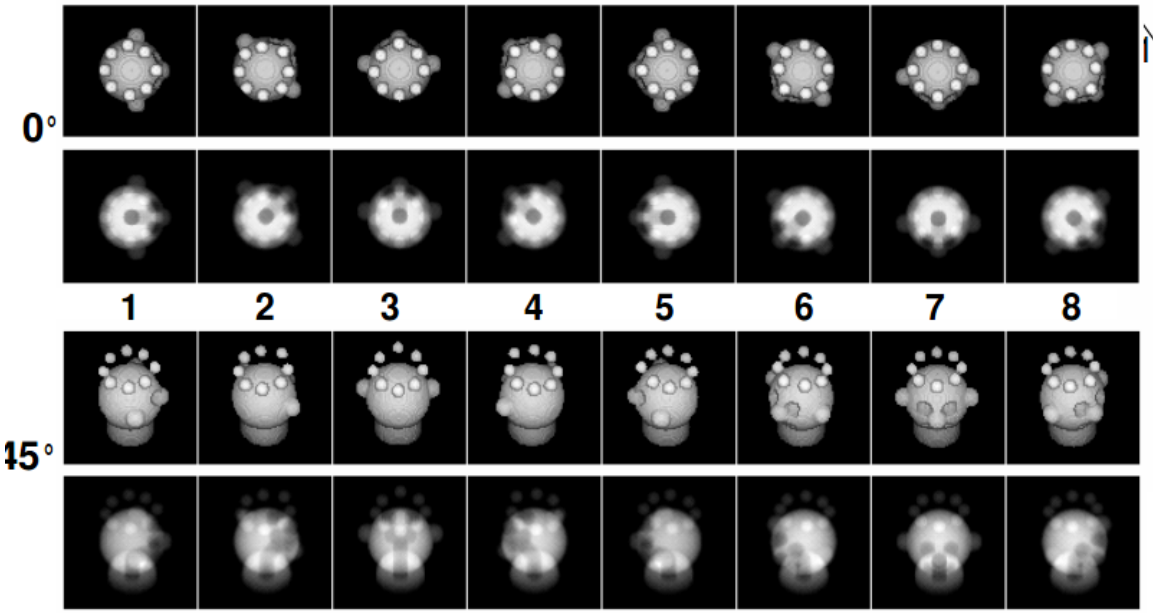


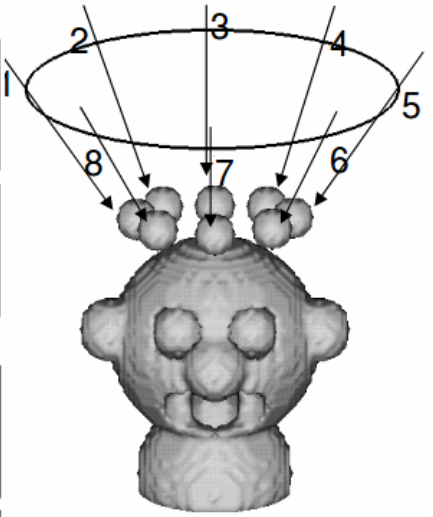
From A. Leschziner, https://software.rc.fas.harvard.edu/leschziner_public/index.php/Main_Page

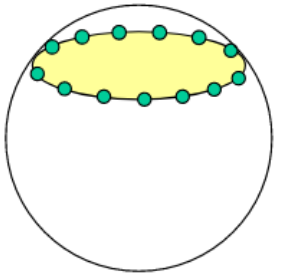
Tomography-like approaches

Random conical tilt

With only two exposures a conical tilt series can be generated



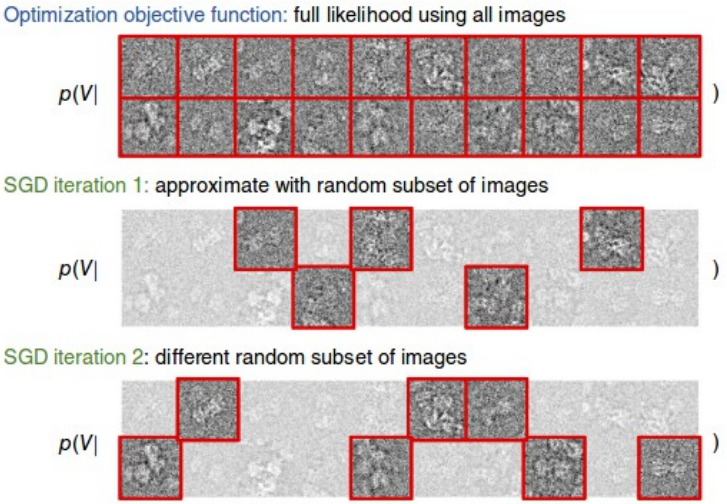


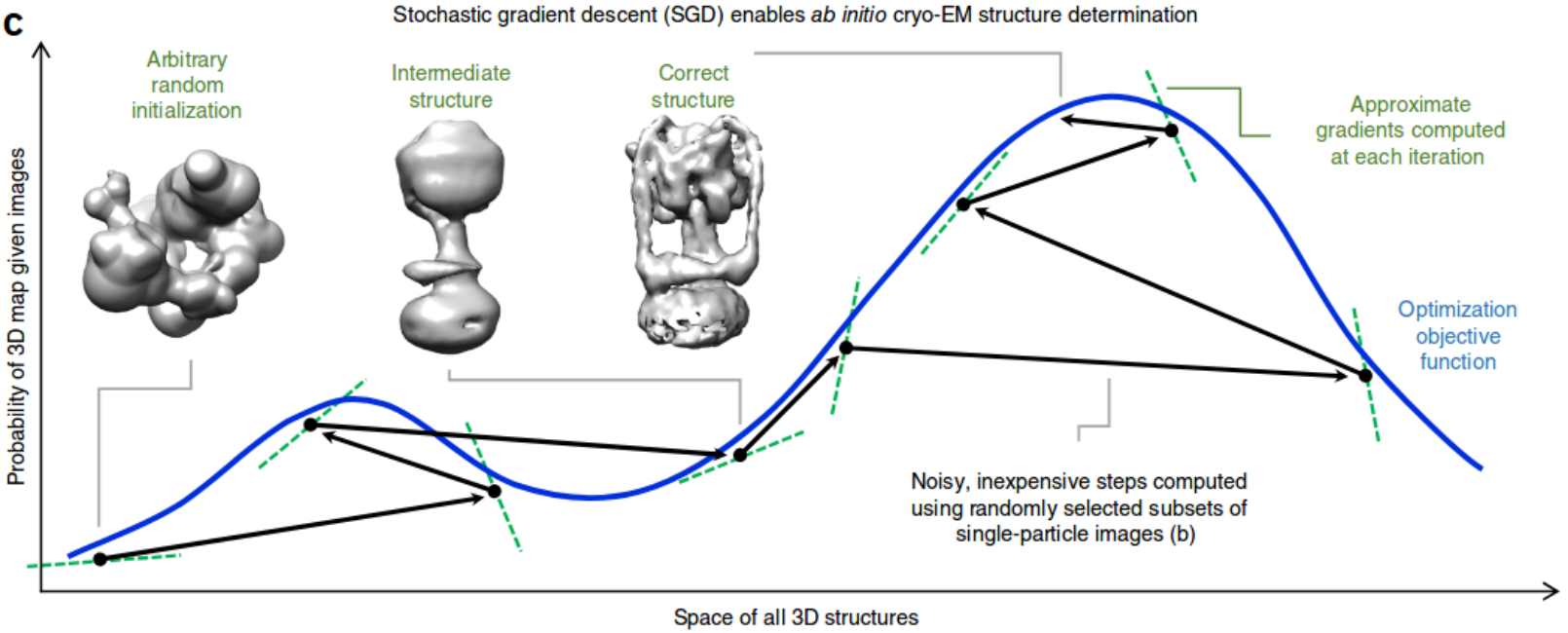


Angular distribution represented on a spherical angular map

Radermacher, M., Wagenknecht, T., Verschoor, A. & Frank, J. Three-dimensional reconstruction from a single-exposure, random conical tilt series applied to the 50S ribosomal subunit of *Escherichia coli*. *J Microsc* **146**, 113-36 (1987).

Probabilistic approaches
Stochastic gradient descent (SGD) : Relion, cryoSPARC





Punjani et al., Nature Methods2017