CNRS
INSERM
Université de Strasbourg

Parc d'Innovation - 1, rue Laurent Fries - ILLKIRCH - C.U. DE STRASBOURG - FRANCE

**1) Short summary of general concepts**

**2) Single particle image processing and 3D reconstruction**

→ **getting prepared for the practicals**

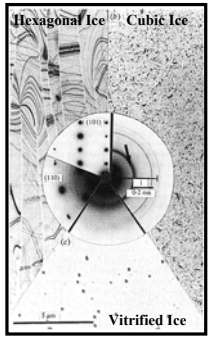
**Bruno Klaholz 2015, ReNaFoBis, Oléron School**
http://www.igbmc.fr
http://igbmc.fr/Klaholz

# Electron microscopy: application examples - Summary

| | |
|---|---|
| *Negative staining* | **2D observation** + 3D reconstruction |
| *Spreading* | **2D observation only** |
| *Shadowing* | **2D observation only** |
| *Cryo-EM* | (2D observation +) **3D reconstruction** |
| *2D crystals* | (2D observation +) **3D reconstruction** |
| *Tomography of cellular structures* | **(2D observation +) 3D reconstruction** |
| *Freeze-fracture* | **2D observation only** |

# The importance of cryo-approaches

**Vitreous ice:**
**forms by flash-cooling, is metastable and converts to crystalline ice modifications:**
**- cubic ice, forms when vitreous ice is warmed up above -135°C → keep samples below ~-135°C**
**- hexagonal ice, forms when water is (relatively slowly) cooled down at atmospheric pressure**
**(is typical source of contamination in cryo-EM)**

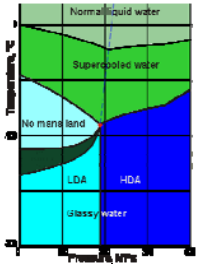



**Dubochet *et al.*, 1988**

**cooling rate required to obtain vitreous ice: ~$10^4$ K / s**

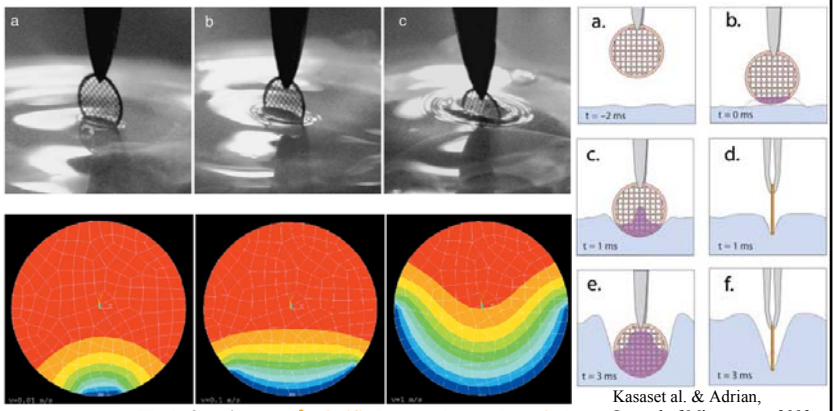
**Boiling and melting points of liquid ethane: -88.7 °C / -183.3 °C, temperature of liquid nitrogen: -196 °C**

**phase diagram of water**




http://www1.lsbu.ac.uk/water/amorph.html

# The importance of cryo-approaches




Kasaset al. & Adrian, Journal of Microscopy, 2003

**Benefits?**

**Flash-freezing:**
→ **vitrified water (amorphous ice)**
→ **specimen conservation (frozen-hydrated)**
→ **very weak ice sublimation in the vacuum of the microscope**
→ **fixation of particle orientations**

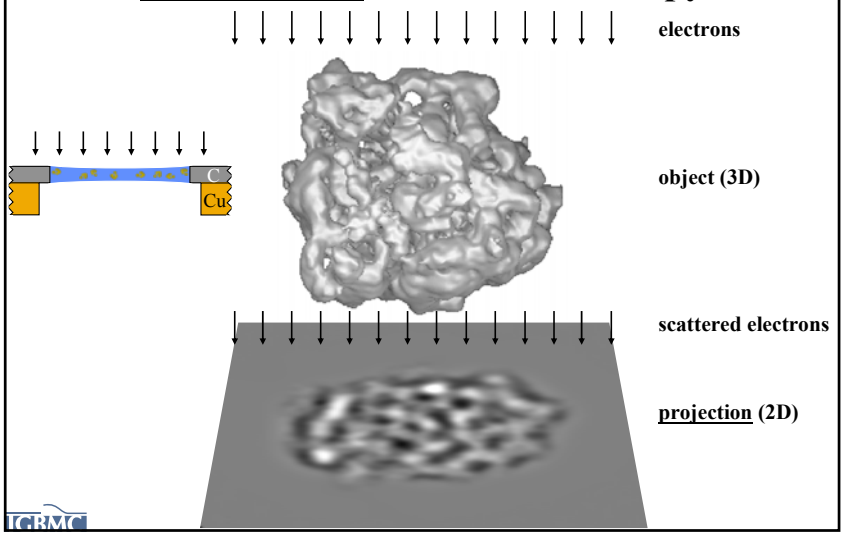
# Sample preparation

**A new research project – steps to cross for single particle cryo-EM:**

- functional studies
- purification and biophysical characterization of complexes (very important)
- optimization of the sample for imaging:
  - fast freezing (vitreous ice)
  - buffer composition
  - support type (holey carbon vs continuous carbon film)
  - concentration (~0.5 mg/ml)
  - ice thickness
  - absence of contamination


IGBMC

# Transmission electron microscopy



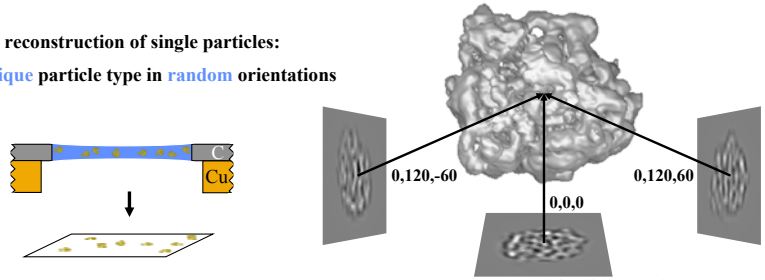

IGBMC

## Structure determination by 3D reconstruction

**Assumptions?**

**3D reconstruction of single particles:**

**unique particle type in random orientations**

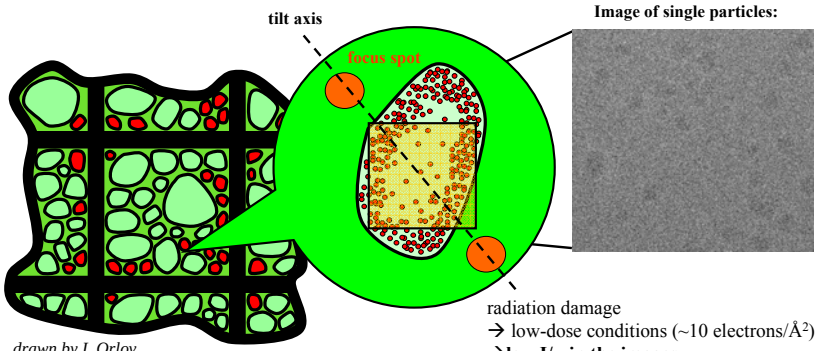



**Reconstruction provides a 3D density map**, i.e. a **structure** with all internal features

IGBMC

# Sample irradiation

**Concept of "low-dose": reduce sample irradiation**




*drawn by I. Orlov*

radiation damage
→ low-dose conditions (~10 electrons/Å$^2$)
→ **low I/σ in the images**
→ **use averaging techniques to recover the signal**

IGBMC

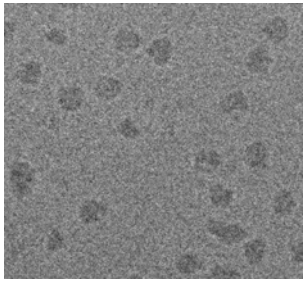
# Single particle image processing and 3D reconstruction.



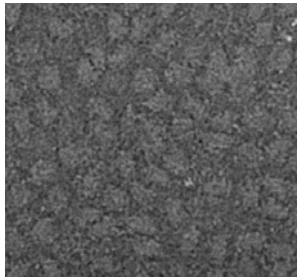
**I. Pre-processing**
- Digitization of micrographs (negatives); not needed if CCD images
- particle selection, « boxing »
- correction of the contrast transfer function
- band-pass filtering and normalisation of particle images

**II. Structure determination**
- particle centering / alignments
- MSA (multivariate statistical analysis) + classification
- angle assignment
    - angular reconstitution
    - projection matching
- 3D reconstruction
- structure refinement
- resolution assessment
- map interpretation; fitting of known structures, atomic model building…

**I. Pre-processing**
- Digitization of micrographs (negatives); not needed if CCD images



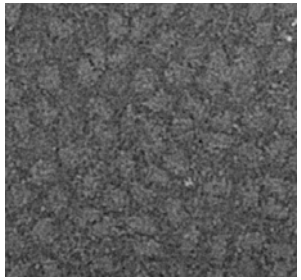
CCD image (positive contrast)



micrograph (negative contrast)

**I. Pre-processing**
- Digitization of micrographs (negatives)

high-resolution scanner (5000 dpi)



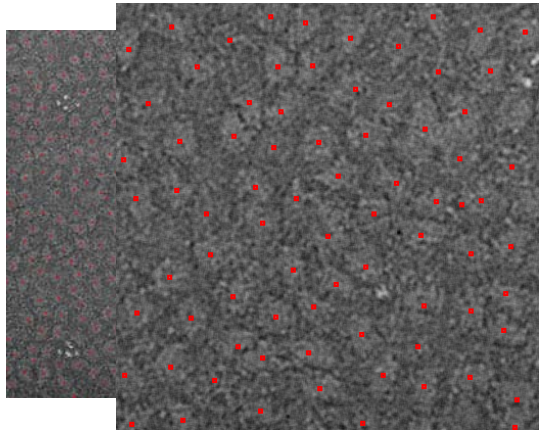
**Sampling** = Pixel size / Magnification

**5 µm / 50 000 = 1 Å / pixel at specimen level**

**I. Pre-processing**
**- particle selection, « boxing »**

**What is important when selecting particles?**

**Proper centering!**



ICBMC

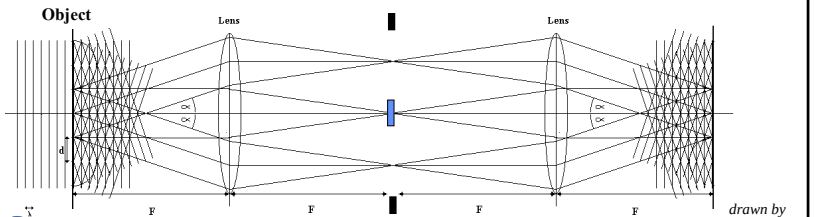
**I. Pre-processing**
**- correction of the contrast transfer function**

**Background:** **Phase Contrast Microscope**

- Small phase difference difficult to observe
- Limited to study of thin specimens (<1000Å)
- **Phase contrast is converted to amplitude contrast by defocusing the specimen**
- Same technique used in light microscopy to study unstained specimens
- Staining provides strong amplitude contrast but may affect the macromolecular structure

**Scherzer focus**

**under-focus (e.g. -2 µm)**

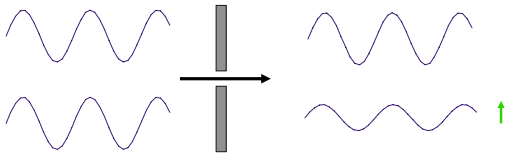



*drawn by M. van Heel*

ICBMC

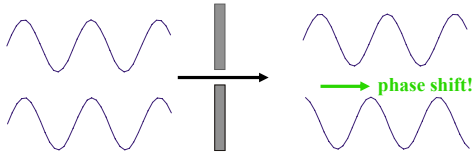
**Specimen contrast**

**Amplitude contrast (inelastic scattering, absorption)**



**amplitude change! (absorption by sample)**

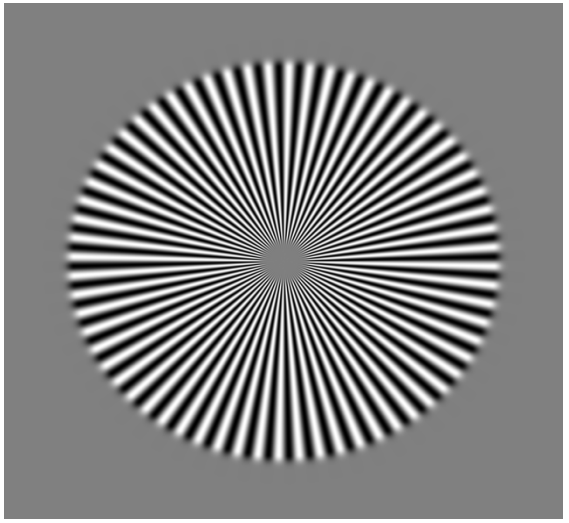
**Phase contrast (elastic scattering)**



**phase shift! (delay through sample)**

ICBMC

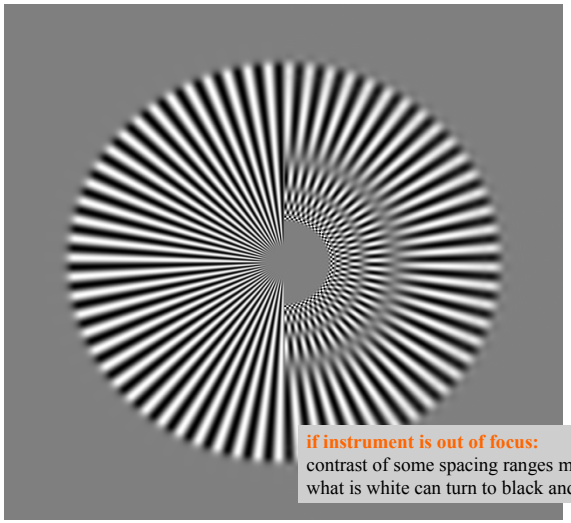
**Siemens stars: a whole range of spacings / frequencies / resolution in a single image**



*drawn by M. van Heel*

ICBMC

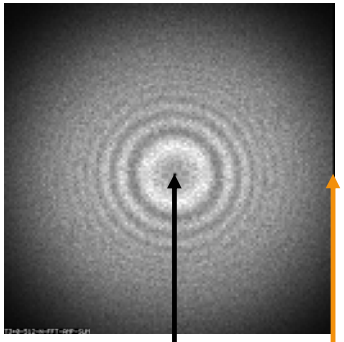
Calculated effect of an electron microscopical PhCTF on the image of a Siemens star



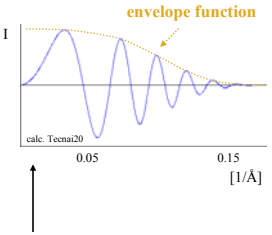
**if instrument is out of focus:**
contrast of some spacing ranges may be reversed, what is white can turn to black and vice versa.

ICBMC

**Powerspectrum = Fourier Transformation of the image**



**Profile of the intensity distribution**




**very low resolution difficult to measure**

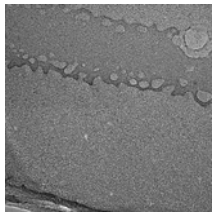
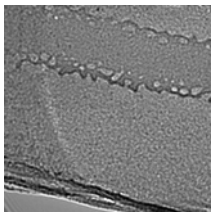
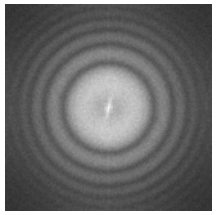
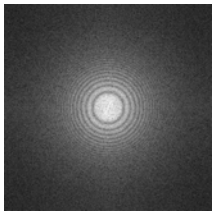
**low resolution**
**low frequency**

**high resolution**
**high frequency**

→ **averaging techniques**

ICBMC

**Defocusing: effect on contrast, power-spectrum and max. resolution with FEG**

| | defocus -1 µm | defocus -10 µm |
|---|---|---|
| **image** |  |  |
| **power-spectrum** |  |  |

***Defocus?***

***Focal point is after the sample (under-focus)***

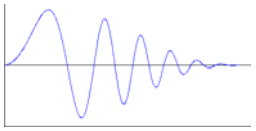
ICBMC

***Images: continuous carbon area from holey carbon grid, cryo, Tecnai20, CCD 2Kx2K, 1024 pixel mode***

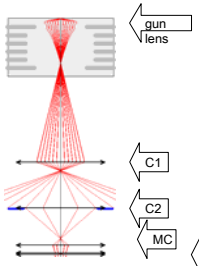
**I. Pre-processing**

- **correction of the contrast transfer function**

**Weak-phase biological specimen → weak-contrast image**



**CTF function:** $T(k) = - \sin (\pi/2\, C_s\, \lambda^3 k^4 + \pi\, \Delta f\, \lambda\, k^2)$




CTF depends on parameters of imaging system:

- ***λ***, wavelength defined by accelerating voltage (e.g. high tension 200 kV --> λ = 0.025 Å)
- ***Cs***, spherical aberration coefficient, determines the quality of objective lens
- other, constant parameters from electron microscope

} **constant**

- ***Δf***, the **defocus value**, (offset from focus, "Scherzer" focus, underfocus: **focal point is below the sample**)
- ***k***, spatial frequency

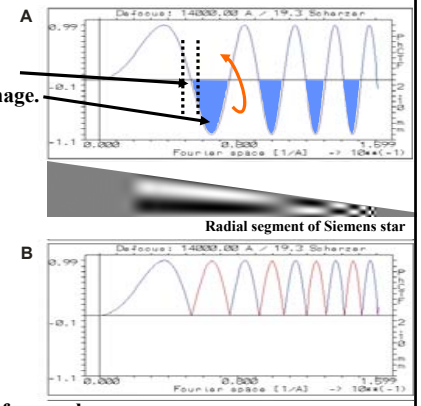
} **variable**

ICBMC

**I. Pre-processing**

- correction of the contrast transfer function (CTF)

No information transfer at the zero crossing;
CTF leads to contrast inversion within the image.

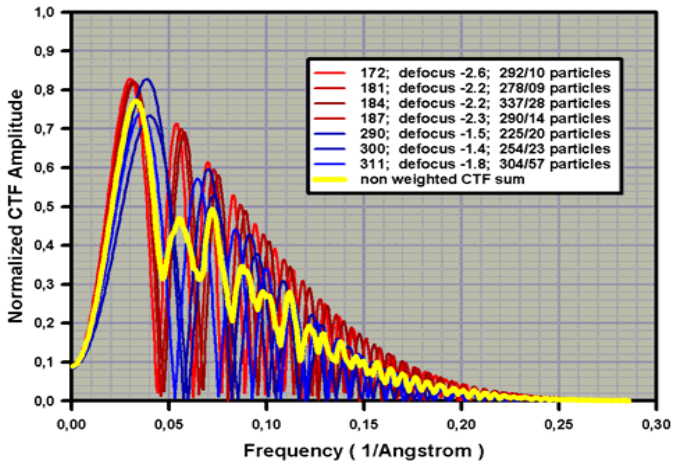



CTF-correction? By "phase flipping":

Therefore: data collection over range of defocus values,
e.g. under-focus -1.0 – 3.0 µm



**Combination of powerspectra from different defocus images**

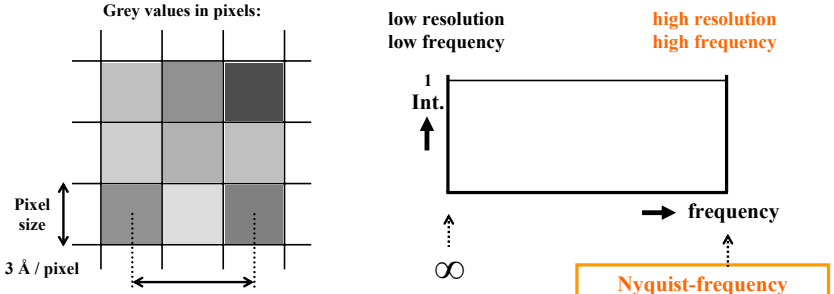



*(spectra also take the envelope function into account)*



**I. Pre-processing**

- band-pass filtering and normalisation of particle images




Maximum resolution = 2 x pixel size = Nyquist-frequency



**I. Pre-processing**

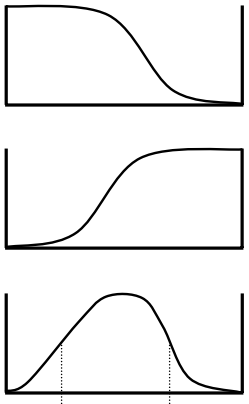
- band-pass filtering and normalisation of particle images

Combination of high-pass and low-pass filters:

low-pass

high-pass

band-pass



Removes:

- low frequency contribution (scanner, etc.)
- high frequency noise

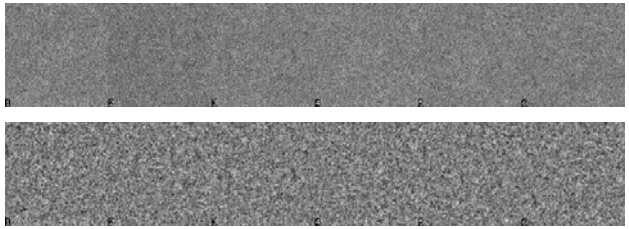
Particle size e.g. 200 Å

Effective high reso. e.g. 8 Å

**I. Pre-processing**

**- band-pass filtering and normalisation of particle images**

**Effect of bandpass filter:**



**Removes:**

**- low frequency contribution (scanner, etc.)**

**- high frequency noise**

ICBMC

**Single particle image processing and 3D reconstruction.**

**I. Pre-processing**

**- Digitization of micrographs (negatives); not needed if CCD images**

**- particle selection, « boxing »**

**- correction of the contrast transfer function**

**- band-pass filtering and normalisation of particle images**

**II. Structure determination**

**- particle centering / alignments**

**- MSA (multivariate statistical analysis) + classification**

**- angle assignment**

**- angular reconstitution**

**- projection matching**

**- 3D reconstruction**

**- structure refinement**

**- resolution assessment**

**- map interpretation ; fitting of know structures, atomic model building…**

ICBMC

**II. Structure determination**

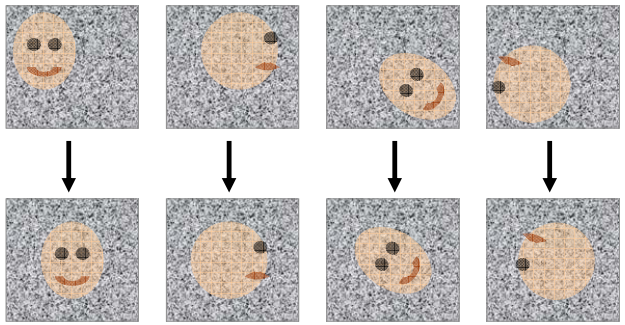
**- particle centering / alignments**

***Overview of the concept:***
***align – classify – reconstruct in 3D***

**"reference-free" alignment (if structure unknown)**
**(or multiple reference alignment, if similar structure already known)**

Translational Alignment

- Requires reference image(s) to align to



*drawn by*
*A. Patwardhan*

ICBMC

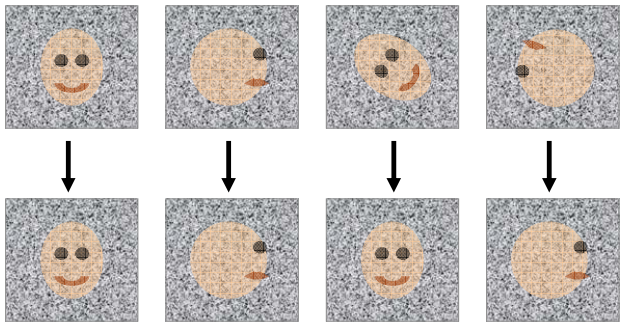
**II. Structure determination**

**- particle centering / alignments**

***Overview of the concept:***
***align – classify – reconstruct in 3D***

Rotational Alignment

- Requires reference image(s) to align to



*drawn by*
*A. Patwardhan*

ICBMC

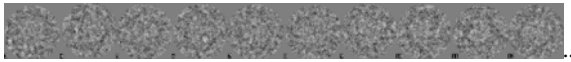
## II. Structure determination
## - particle centering / alignments

**"reference-free" alignment (if structure unknown)**
**(or multiple reference alignment, if similar structure already known)**


...

aligned to total sum of particles

provides centered particle images:

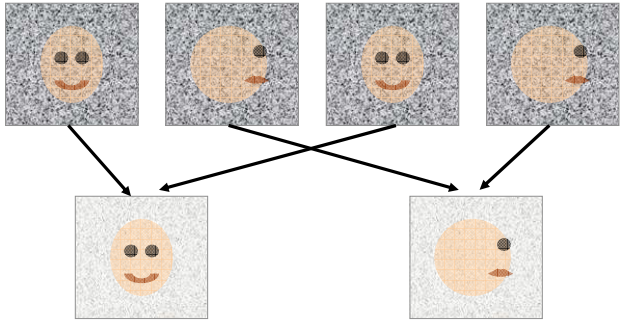

...

ICBMC

*Overview of the concept:*
*align – classify – reconstruct in 3D*

## II. Structure determination
## - MSA (multivariate statistical analysis) + classification

- Combine like views to improve signal to noise: **how?** **Requires statistical analysis of the pixel intensities**

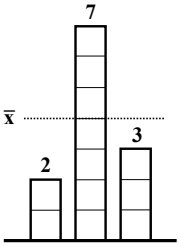


*drawn by*
*A. Patwardhan*

ICBMC

## Some "basic statistics"

**example for 3 values (measurements)**




"how variable are the data"

$n = 3$

$\Sigma x_i = 2+7+3 = 12$

$\bar{x} = 1/n \Sigma x_i = 1/3 (2+7+3) = 4$

$\sigma^2 = 1/n \Sigma (x_i - \bar{x})^2 = 1/3 [(2-4)^2 + (7-4)^2 + (3-4)^2 ] = 14/3$ variance

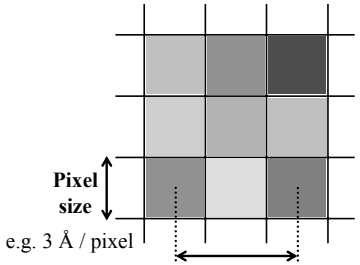
$\sigma = \sqrt{1/n \Sigma (x_i - \bar{x})^2} = \sqrt{14/3} = 2.16$ standard deviation

"normalization"

ICBMC

**Images are composed of pixels:**

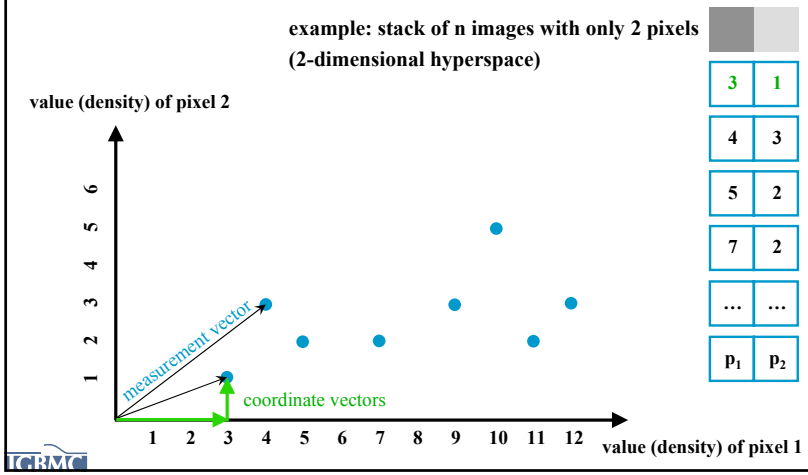
**Grey values in the pixels, each is a measurement:**




Maximum resolution = 2 x pixel size (**Nyquist frequency!**)

ICBMC

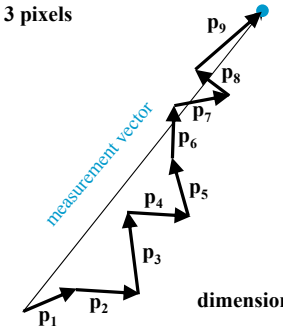
**example: stack of n images with only 2 pixels (2-dimensional hyperspace)**




| | |
|---|---|
| 3 | 1 |
| 4 | 3 |
| 5 | 2 |
| 7 | 2 |
| … | … |
| $p_1$ | $p_2$ |

**extension to more dimensions,**
**example: image with 3 x 3 pixels**

| | | |
|---|---|---|
| $p_1$ | $p_2$ | $p_3$ |
| $p_4$ | $p_5$ | $p_6$ |
| $p_7$ | $p_8$ | $p_9$ |




**dimensionality in image space: $p^2$**

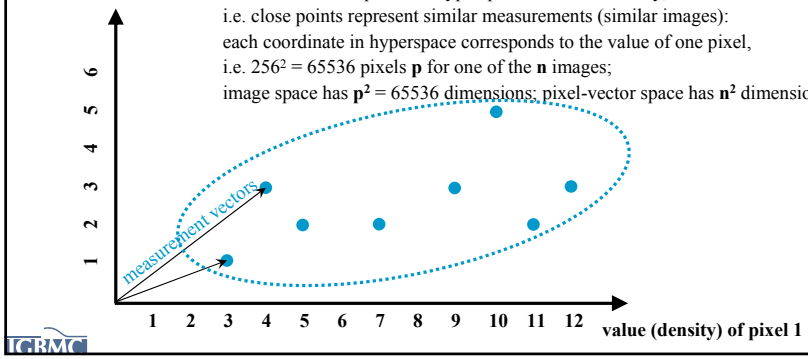
**length of each coordinate vector $\triangleq$ intensity of pixel $p_i$**

**Large number of comparable measurements: stack of n images with the same size (e.g. 256x256)**

**Describe image pixels as a vector of numbers (individual pixel values / densities).**
**This vector is positioned in hyperspace:**

→ each measurement corresponds to a point in hyperspace,
all measurements thus create a ***data cloud***.

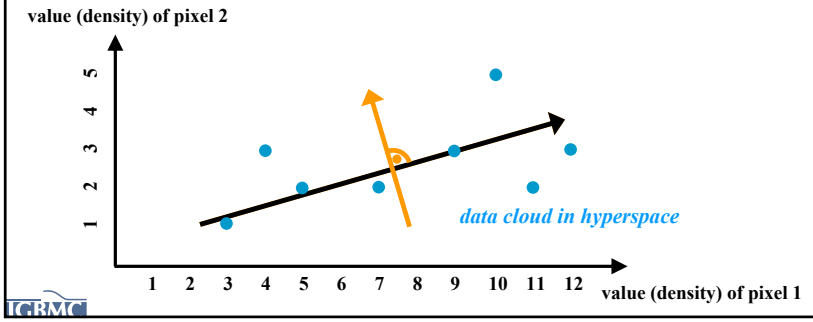
Distance between points in hyperspace defines similarity,
i.e. close points represent similar measurements (similar images):
each coordinate in hyperspace corresponds to the value of one pixel,
i.e. $256^2 = 65536$ pixels **p** for one of the **n** images;
image space has $\mathbf{p^2} = 65536$ dimensions; pixel-vector space has $\mathbf{n^2}$ dimensions.



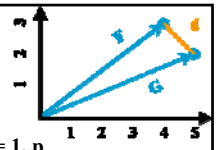

**Aim of MSA: adapt the coordinate system of the hyperspace to the shape of the data cloud.**

First axis of the rotated coordinate system will correspond to the largest elongation of the data cloud,
**i.e. highest variance**; the next (orthogonal) axis corresponds to the next-strongest variance.

**Data reduction**: use first few components that describe the strongest variations / main components;
reduction of dimensionality: e.g. 20 or 50 instead of $p^2$ (here 65536): gain factor **>1000**;
reduction of noise: by omitting higher components which describe only variance in noise;
main directions of variations: given by eigenvectors



**Similarity between two measurement vectors F and G:**
**inner product of the vectors (= correlation = covariance):**




$C_{FG} = 1/p \, \Sigma \, F_a \cdot G_a$ covariance $a = 1, p$

$C_{FF} = 1/p \, \Sigma \, F_a \cdot F_a = 1/p \, \Sigma \, F_a^2$ variance $a = 1, p$

**Euclidian square distance:** *e.g. unit cube:*

$d = \sqrt{1^2+1^2+1^2} = \sqrt{3}$

$D^2_{FG} = \Sigma \, (F_a - G_a)^2$

$= \Sigma \, F_a^2 + \Sigma \, G_a^2 - 2 \, \Sigma \, F_a \cdot G_a$ $a = 1, p$

**(variances in F and G) minus (2 · correlation between F and G):**

**short distance means high correlation**

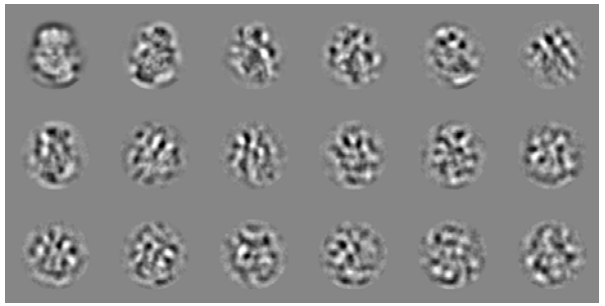
**PCA metric (principal component analysis)**

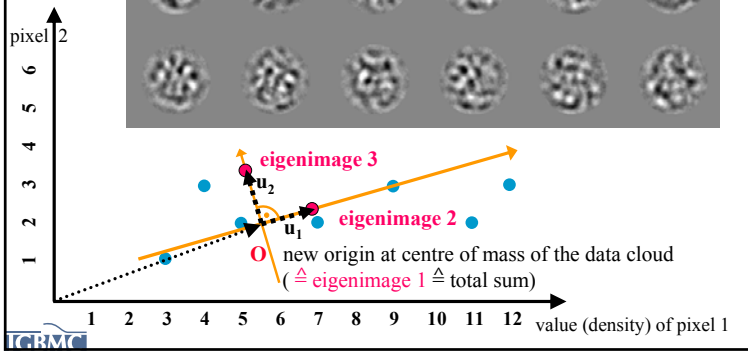
**[for comparison: discrete correlation function: $C_{FG} = \Sigma \, F_a \cdot G_{(a-x)}$, with shift x;**

**cross-correlation coefficient CCC = $\Sigma \, F_a \cdot G_a$ / sqr ( $\Sigma \, F_a^2 \cdot \Sigma G_a^2$) ]**

ICBMC

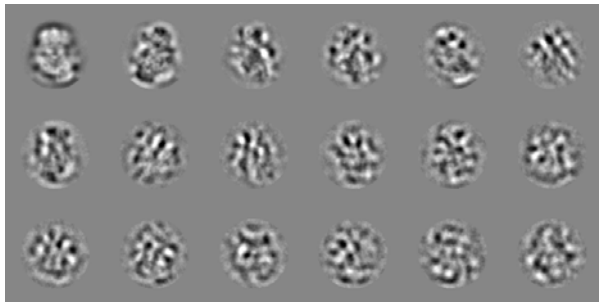
**example: eigenimages**






ICBMC

**example: eigenimages**



**example: first 18 eigenimages of a data set**

The entire data set can be reconstituted from a linear combination of the eigenimages, or simply approximated by a small subset of eigenimages (data reduction!)

**"$a \cdot u_1 + b \cdot u_2 + \ldots$"**

ICBMC

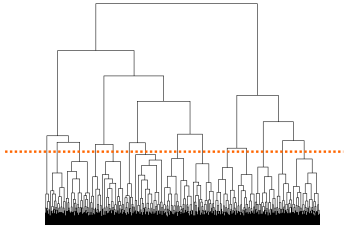
**II. Structure determination**

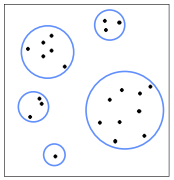
**- MSA (multivariate statistical analysis) + classification**

multivariate statistical analysis (MSA), related to principal component analysis:
data set can be represented as a linear combination of images, each describing the highest differences within the data

→ data compression

→ images with statistically similar pixel intensity distribution can be grouped = classified into groups of images describing similar views of the 3D object





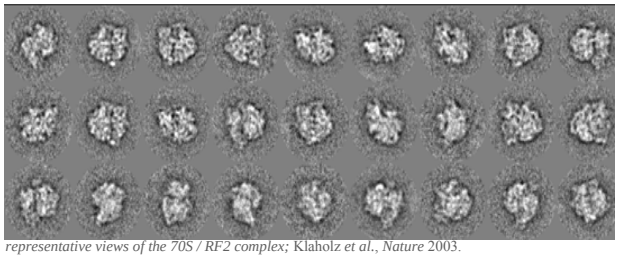
**hierarchical ascendant classification**

ICBMC

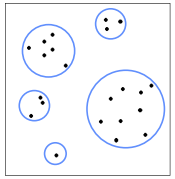
**II. Structure determination**

**- MSA (multivariate statistical analysis) + classification**

signal enhancement after
- classification by MSA
- hierarchical ascendant classification
- averaging of particles representing same views into class averages:



*representative views of the 70S / RF2 complex;* Klaholz et al., *Nature* 2003.



Average the images of each class

***Typical class averages of ribosome particle images***



**Correct terms are important:**

**A classification is based on a statistical analysis:**

**- multivariate statistical analysis (MSA) provides information on variance (variability) which serves to merge similar images into class averages (classes); is *independent* of a reference**

**- classes *are NOT*: the sum of images that correlate best with a reference (through a multi-reference alignment)**



**Some references**


- L. Borland & M. van Heel; Classification of image data in conjugate representation spaces, *J. Optic. Soc. Am.* A, 7 (1990) 601-610.
- M. van Heel *et al.;* Single-particle cryo electron microscopy: towards atomic resolution; Quart. Rev. Biophys. 33 (2000) 307-369.
- M. van Heel, Multivariate Statistical Classification of Noisy Images (Randomly Oriented Biological Macromolecules) *Ultramicroscopy* 13 (1984) 165-183.
- M. van Heel, Classification of very large electron microscopical image data sets, *Optik* 82 (1989) 114-126.
- E.R. Malinowski, *Factor Analysis in Chemistry*, 3rd ed. (2002)
- Benzécri J.-P. , L'Analyse des Données Vol 2, L'analyse des correspondances (1973-1980) Dunod Paris.
- Frank J: Three-Dimensional Electron Microscopy of Macromolecular Assemblies, Oxford University Press (2006).
- van Heel M, Frank J: Use of multivariate statistics in analyzing the images of biological macromolecules, *Ultramicroscopy* 6 (1981) 187-194.
- van Heel M: Multivariate Statistical Classification of Noisy Images (Randomly Oriented Biological Macromolecules), *Ultramicroscopy* 13 (1984a, )165-183.
- Ward JH: Hierarchical grouping to optimize an objective function. *J. Amer. Statist. Assoc*. 58 (1982) 236-244.




## What can we apply MSA to?

**- 2D classification (reference-free alignment: only centered data, not rotationally aligned)**

**- alignment by classification (alignment against class averages or a typical eigenimage)**

**- analysis of symmetry (through symmetry in the eigenimages)**

**- local MSA (focus on an area with high structural variability)**

**- re-classification of class averages belonging to an object view**

**- size-classification (e.g.** White et al., *J. Mol. Biol.* **336** (2004) 453-460).

**- 3D classification of structures (separation of mixed particle populations): particles:3D-SC, sub-tomograms**

**- classification of powerspectra (sorting of defocus classes)**

**Important to do before MSA:**
- **normalisation**
- **filtering**
- **centered data (aligned if for structure refinement)**
- **define MSA area: MSA mask**

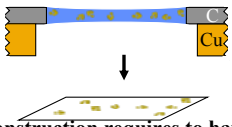
*Determining structures of multiple conformational states in a single sample*

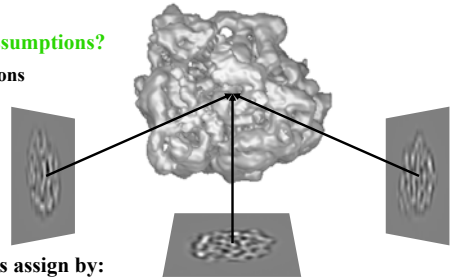

ICBMC

**II. Structure determination**
**- angle assignment**
**- angular reconstitution (in early stage of structure determination)**
**- projection matching (if structure already well refined): find best correlation between input image and reference images from 3D re-projections)**

**3D reconstruction of single particles: assumptions?**

**unique particle type in random orientations**



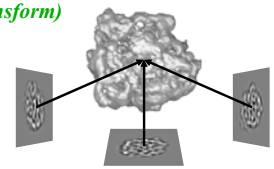



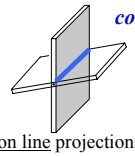
**Reconstruction requires to have angles assign by:**
- angular reconstitution (in early stage of structure determination), or
- projection matching (if structure already refined; reference-dependent; bias), or
- maximum likelihood

ICBMC

*Real space* *Fourier space*

| | | | |
|---|---|---|---|
| *map* | 3d density distribution | 3D FFT | 3D Fourier transform |
| back-project ↑ project ↓ | | | extract central section ↓ insert ↑ |
| *class average* | 2D projection | 2D FFT | 2D central section |
| back-project ↑ project to line ↓ | | | extract central line ↓ insert ↑ |
| *sinogram line (Radon transform)* | 1D projection | 1D FFT | 1D central line |

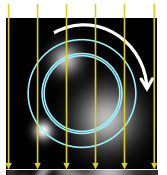


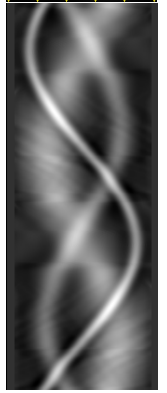



common line projections theorem
Theorem of the central section.

ICBMC

**II. Structure determination**
**- angle assignment**
**- angular reconstitution**



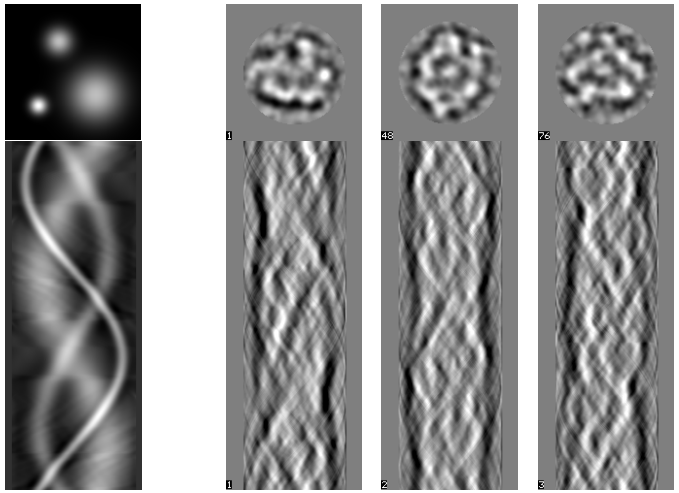



*amplitude-square-root filtered*

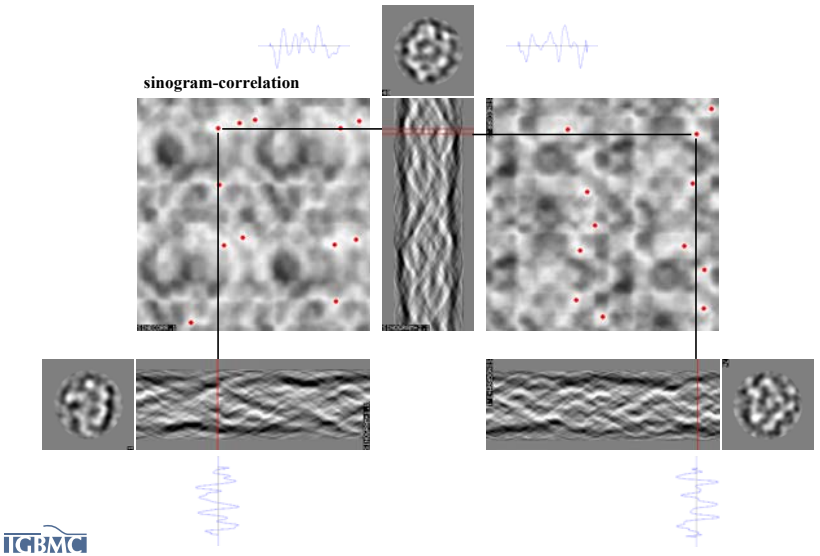
**sinogram = line-projection of the 2D image (also called Radon transform)**

ICBMC

Select 3 clearly different views (here: class average numbers 1,48,76):



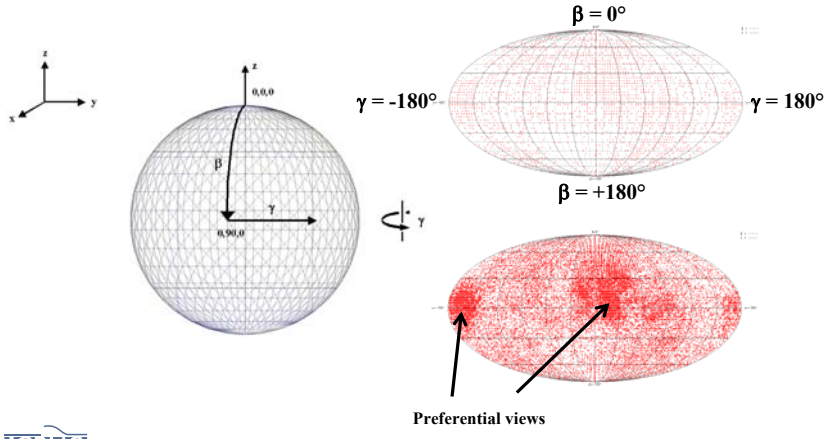
ICBMC




ICBMC

**II. Structure determination**
**- angle assignment**
**- angular reconstitution**

**Particle angles plotted on sphere:**




ICBMC

**II. Structure determination**
**- angle assignment**
**- angular reconstitution**

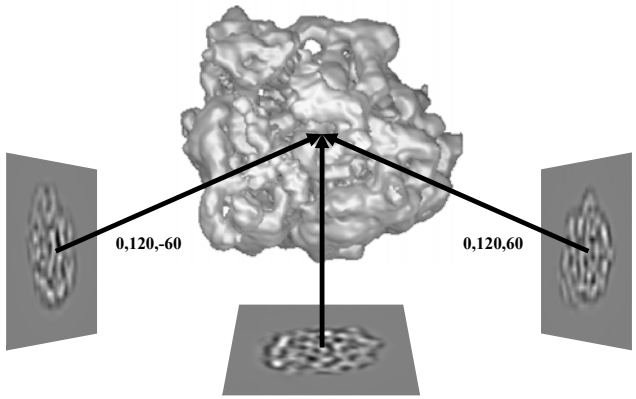
**In case of *ab initio* structure determination by reference-free alignment and angular reconstitution:**

**Does not allow to determine handedness, requires either:**

- **random conical tilt (Radermacher *et al.*, J. Microsc. 1987)**
- **tomography**
- **phase residual error using a tilt pair (Rosenthal & Henderson, JMB 2003)**
- **fitting of crystal structures**

ICBMC

## II. Structure determination
## - 3D reconstruction

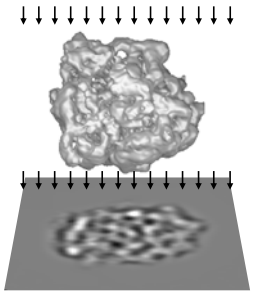



back-projection → 3D reconstruction

ICBMC

## II. Structure determination
## - 3D reconstruction

*Transmission electron microscopy*

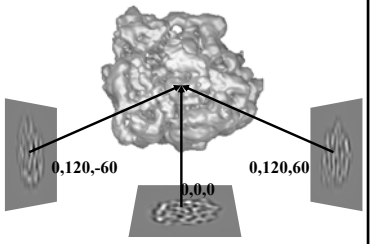


electrons

object (3D)

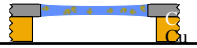
scattered electrons

projection (2D)




back-projection → 3D reconstruction

**All *internal* features in a 2D view!**



ICBMC

## Some basic concepts of cryo-EM & 3D reconstruction

**Correct terms are important (be precise and rigorous in science :-)**

**By cryo-EM, we obtain:**

- a "3D reconstruction" (initial or refined)
- a "cryo-EM map" or "density map"
- a "structure"

**technically:**

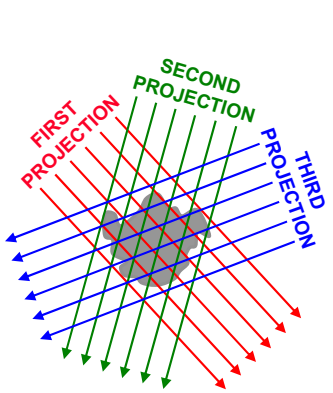
- back-projection
- angular reconstitution
- random conical tilt
- tilt series / tomogram

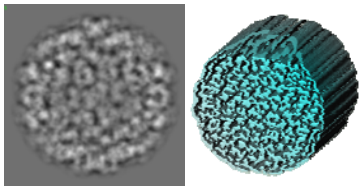
**NOT:**

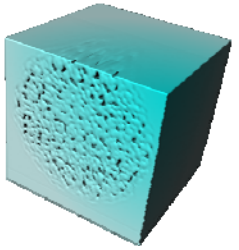
- an "envelope" (would be SAXS or neg. stain. EM)
- a "volume", units would be Å$^3$ (e.g. volume of a pocket, volume x density = mol. mass)
- a "surface", units would be Å$^2$ (e.g. interaction surface between 2 proteins)
- a "model", would be a molecular model *fitted to* the map (crystallography/cryo-EM)
  or a model *compatible with* SAXS data or NMR restraints;
  other "models": "homology model", "hypothetical model", "working model"

ICBMC

## II. Structure determination
## - 3D reconstruction





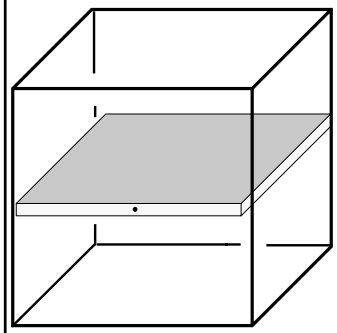
*drawn by I. Orlov*

ICBMC

sun rays: projections

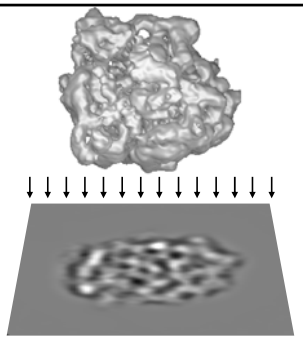
ICBMC

**Theorem of the central section**



reverse Fourier transform

Fourier transform



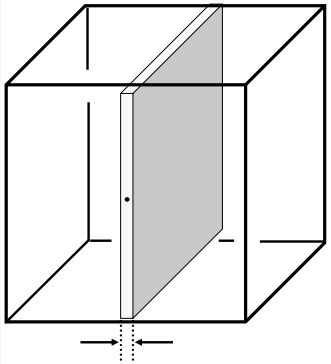
*central section in Fourier space*

**Theorem of the Central Section:**
**The FT of a 2D projection corresponds to the central section of the 3D Fourier transform of the density (i.e. plane perpendicular to the direction of the projection).**

**D. DeRosier & A. Klug, *Nature* 1968; R. G. Hart, *Science* 1968.**

ICBMC

**3D reconstruction by feeding in central sections in Fourier space**

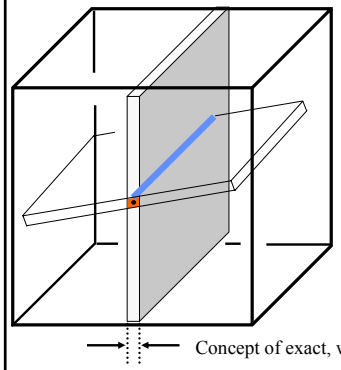


**1/D, D particle size**

***(all representations in Fourier space)***

ICBMC

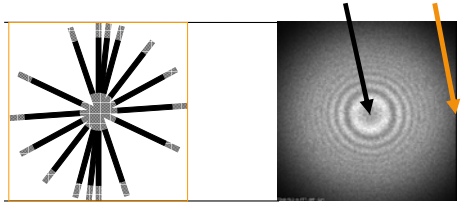
**3D reconstruction by feeding in central sections in Fourier space**



**1/D**

**Reminder: power-spectrum:**

**low resolution** **high resolution**
**low frequency** **high frequency**

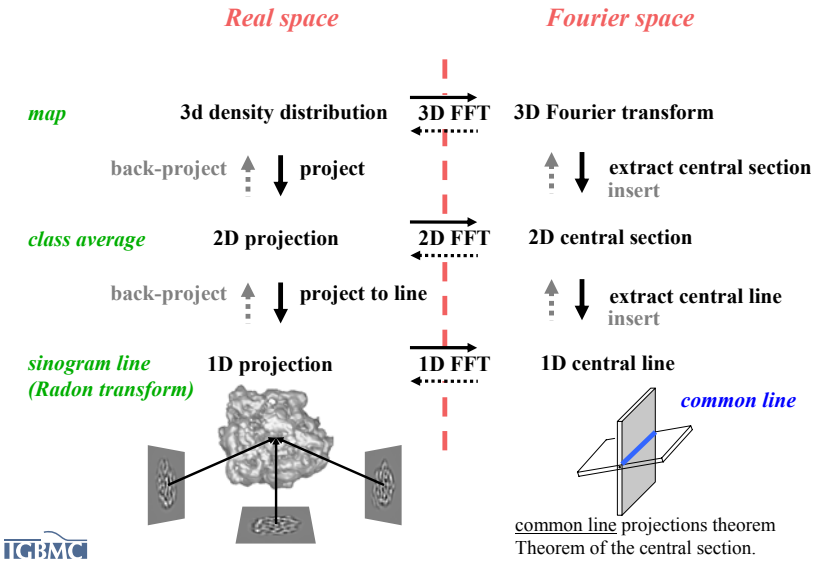


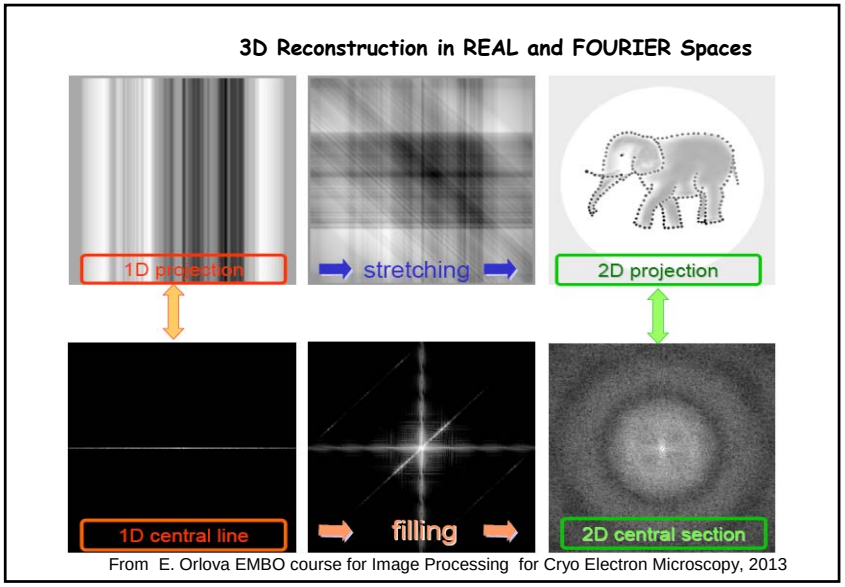
(representation in 2D for simplification)

Concept of exact, weighted filtered back-projection (WBP) reconstruction:
- back-projection
- filtered (partial removal of low frequencies to reduce overlap of central sections)
- weighting (down-weighting preferential orientations)

**(Harauz & van Heel, 1986; Radermacher *et al.*, 1986)**

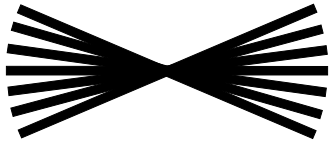
ICBMC



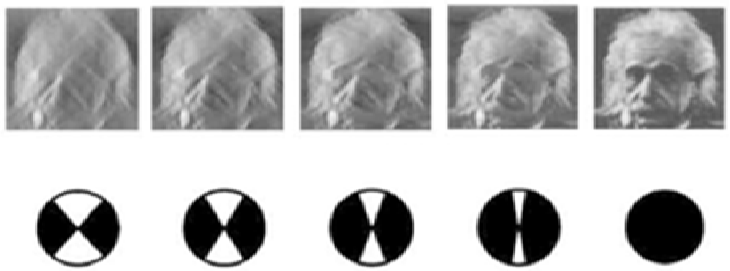
From E. Orlova EMBO course for Image Processing for Cryo Electron Microscopy, 2013

**Back projection of individual images**
**fills the reciprocal space by adding central sections:**



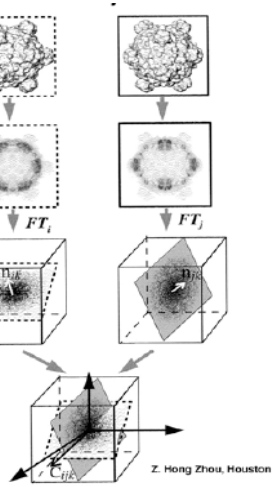
**central sections in Fourier space**

**Filling Fourier space**

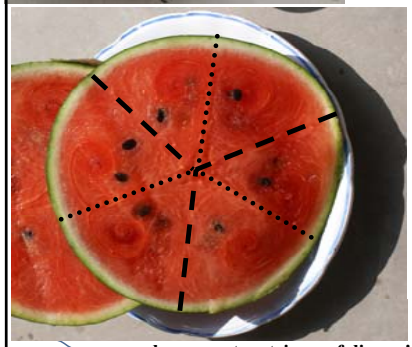


*(example from tomography)*

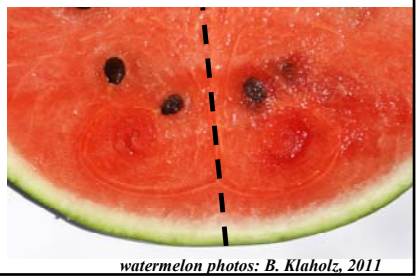
## 3D Reconstruction in REAL and FOURIER Spaces




*symmetries in nature…*



pseudo-symmetry: trimer of dimers!



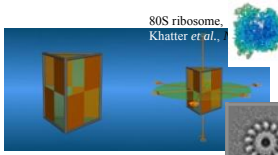
*watermelon photos: B. Klaholz, 2011*

IGBMC

## Point group symmetries of biological objects (e.g. used in single particle reconstructions)

$C_n$

80S ribosome, Khatter *et al.*, 2015

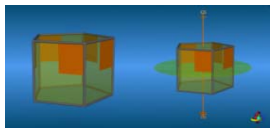


$C_1$

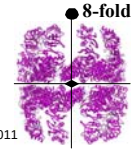
$C_{13}$

SPP1 bacteriophage Orlova *et al.*, *EMBO J* 2003.

$D_n$



$D_8$

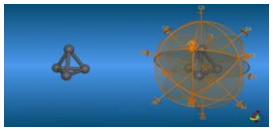


8-fold axis

2-fold axis

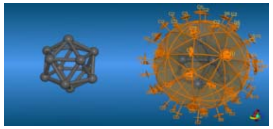
Chaperonine complex Zhang *et al.*, *Structure* 2011

T



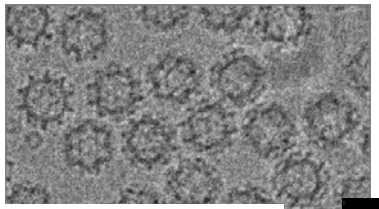
O

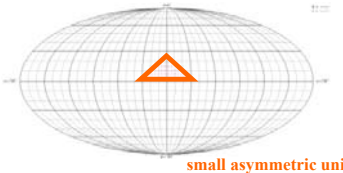
I



http://csi.chemie.tu-darmstadt.de

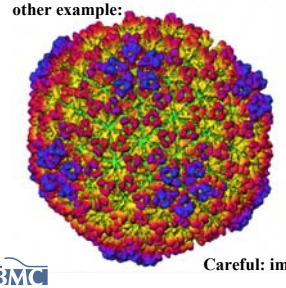
## Practical Example – Icosahedral Reconstruction

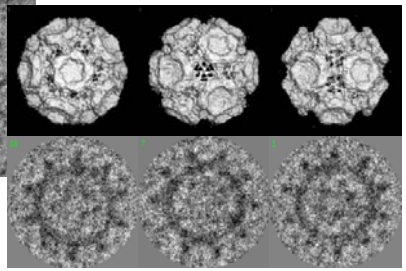




small asymmetric unit

other example:



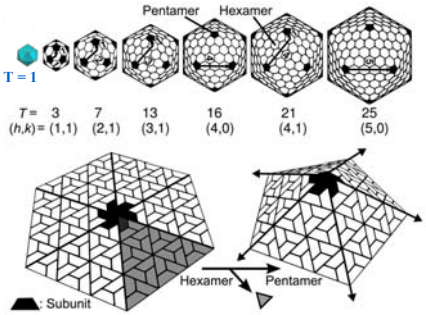


IGBMC

Careful: imposition of inexplicit symmetry can happen…

## Icosahedral Triangulation Number

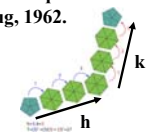
**The T-number is calculated by**
**(1) applying a grid to the surface of the virus with coordinates *h* and *k*,**
**(2) counting the number of steps between successive pentagons on the virus surface,**
**(3) applying the formula:** $T = h^2 + h \cdot k + k^2$ gives the number of structural units per face, = number of subunits in the asymmetric unit




always 12 pentamers,
but variable hexamer insertion

**Multi-symmetric polyhedra:**
**Goldberg, Tôhoku Math., 1937.**

**Concept of quasi-equivalence:**
**Caspar & Klug, 1962.**




http://viralzone.expasy.org/all_by_protein/1057.html

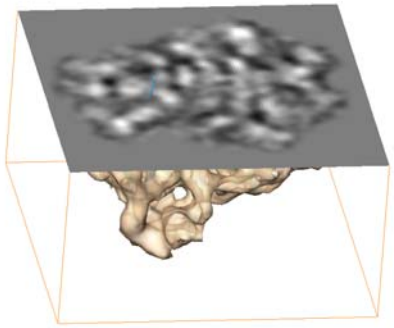
**The two integers h and k describe the number of hexamers (h+k-1)**
**one has to "walk over" to get from one pentamer to an adjacent pentamer.**



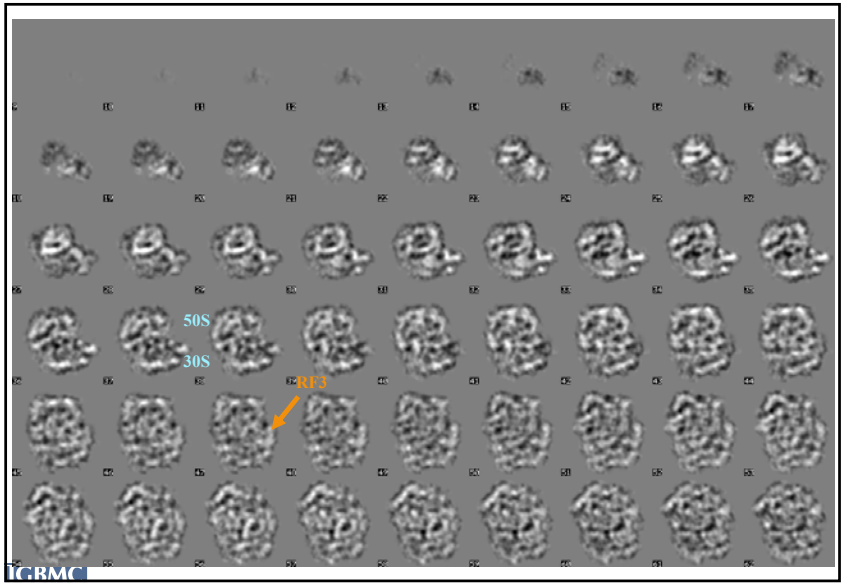
see also Johnson & Speir, *JMB* 1997. Mannige & Brooks, PLoS One, 2010

## II. Structure determination
## - 3D reconstruction

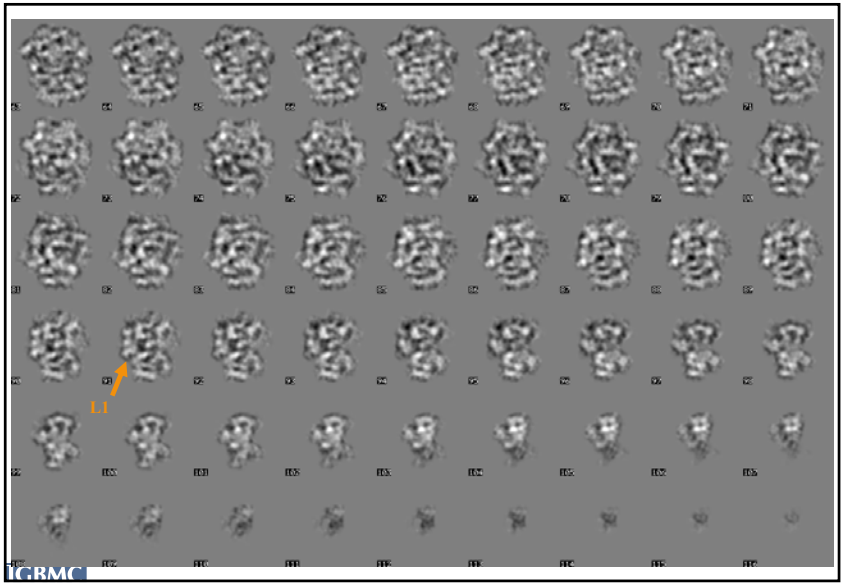
**<u>Representing 3D structures</u> as consecutive sections through the 3D structure:**





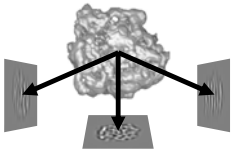


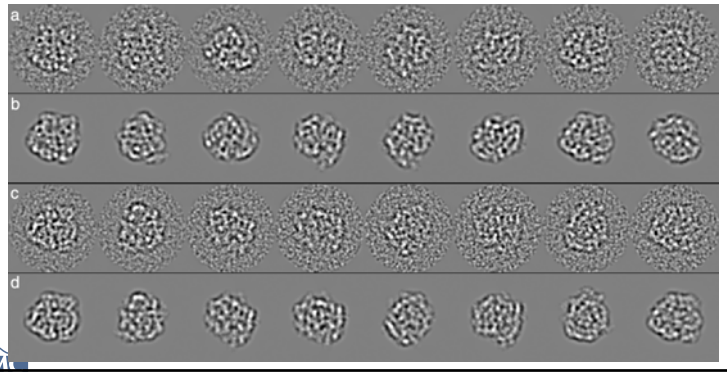


## II. Structure determination
## - 3D reconstruction

**cross-validation of angle assignment and image quality by comparison with re-projections according to the same angles**






IGBMC

## II. Structure determination
## - structure refinement

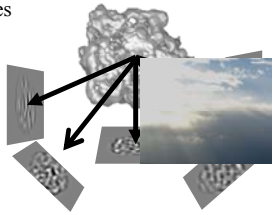
iteration

- centering/alignment
- variance analysis + classification
- angle assignment
- angular reconstitution → 3d-reconstruction
- reprojections = new references

*"phase" the particles by aligning/classifying them*

**- improve quality of angle assignment**

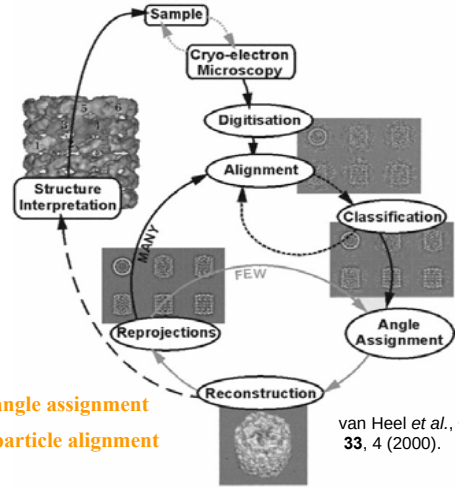
**- improve quality of particle alignment**



**equally distributed forward-projections**
(re-projections)

IGBMC

## II. Structure determination
## - overview




**- improve quality of angle assignment**

**- improve quality of particle alignment**

van Heel *et al.*, Quart. Rev. Biophys **33**, 4 (2000).

IGBMC

## II. Structure determination
## - resolution assessment

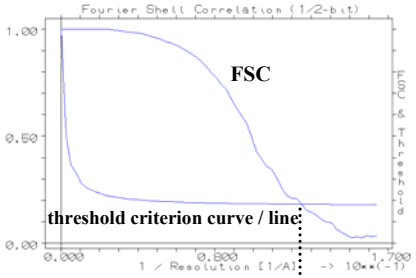
**Fourier Shell Correlation (FSC)**

**Particle data set**

**split into 2 halves**

**3D** **3D**

**calculate cross-correlation by shells in Fourier space**




**1 / 1.25 x 1 / 10 Å = 8 Å**

**Keep in mind: resolution is what you can resolve in the 3D map!**

IGBMC

## Some basic concepts of cryo electron microscopy

**Basic aspects:**

- "resolution" corresponds to "spatial frequency" in image processing (1/ Å)
- Nyquist frequency is = 2 x pixel size, e.g. 1 Å / pixel → Nyquist = 2 Å
- interpolations during 2D image alignment and 3D reconstruction limit the possible resolution to about 2/3 of the Nyquist frequency, i.e. here ~ 3 Å *(exception: super-reso)*

*pixels in 3D: "voxel"*

**Consider:**

- any correlation calculation (e.g. alignment) is biased by the reference used
- resolution estimation, criteria used:
  - 0.5, arbitrary, historically from the virus field, tends to underestimate resolution
  - 0.143 (Henderson) and ½ bit (van Heel)
  - 3 σ, not used anymore (over-estimation)
  - features in the map: can we see dsRNA helices (~10-12 Å resolution), α-helices (~8 Å), ß-sheets (~5 Å) or side chains (4-2.5 Å depending on size)?

IGBMC

**II. Structure determination**
**- map interpretation**

IGBMC

**II. Structure determination**
**- map interpretation ; fitting of crystal or NMR structures**

**Fitting procedures:**

- manual fitting (e.g. O, A. Jones, Acta Cryst. (1991))
- real space fitting
- reciprocal space fitting

**1) global search**

**2) refinement**

**e.g.** torsion-angle molecular dynamics

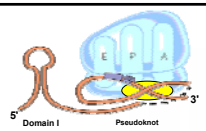
**- fit complete structures, domains, factors;**

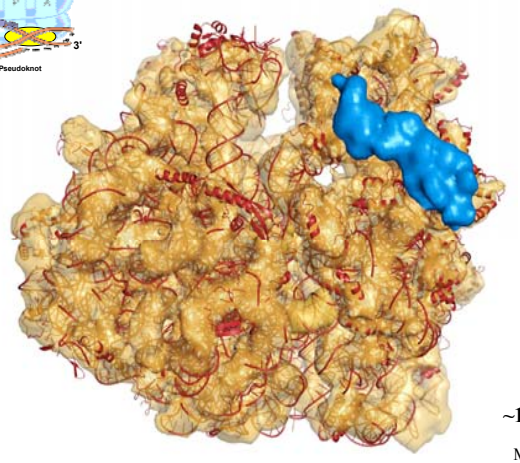
**Usually backbone is enough.**

**rigid body or flexible fitting**

**- use full maps or difference maps**

**Be careful with local minima and over-fitting!**



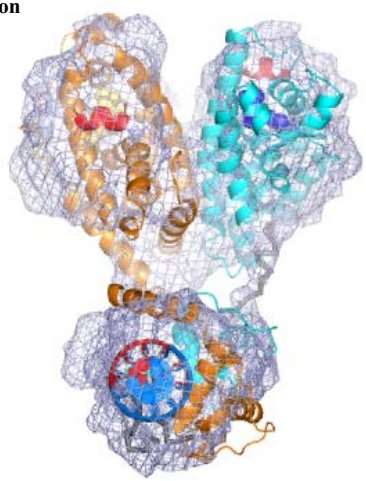
**~12 Å resolution**

Marzi *et al.*, *Cell* 2007.

**50S** **30S**

Fitting of 70S crystal structure from ***E. coli*** (Schuwirth *et al.*, 2005; *Science*, **310**, 827-834)

**II. Structure determination**
**- map interpretation**



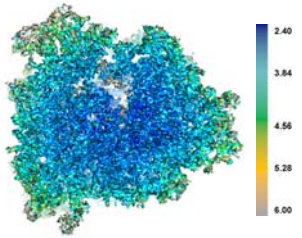
Orlov *et al.*, *EMBO J.* 2012.

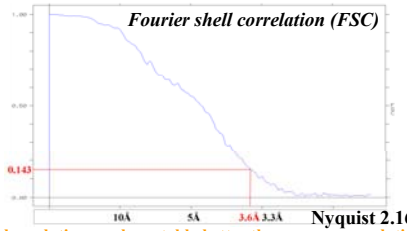
IGBMC

IGBMC

*Atomic structure of the human ribosome*

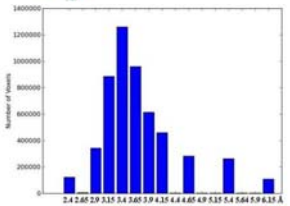
**Structure determination at < 3 Å resolution by single particle cryo-EM**

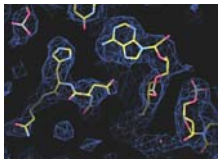



**Nyquist 2.16Å**

→ local resolution can be notably better than average resolution





**cryo-EM map**

**Khatter *et al.*, *Nature, 2015.***

**Some softwares:**
**(for single particle cryo-EM image processing)**

**- Imagic**
**- Spider**
**- FreAlign**
**- EMAN**
**- XMIPP**
**- Relion**
**- Scipion**
**- …**

***Plus other specific software for helical reconstructions, viruses etc.***