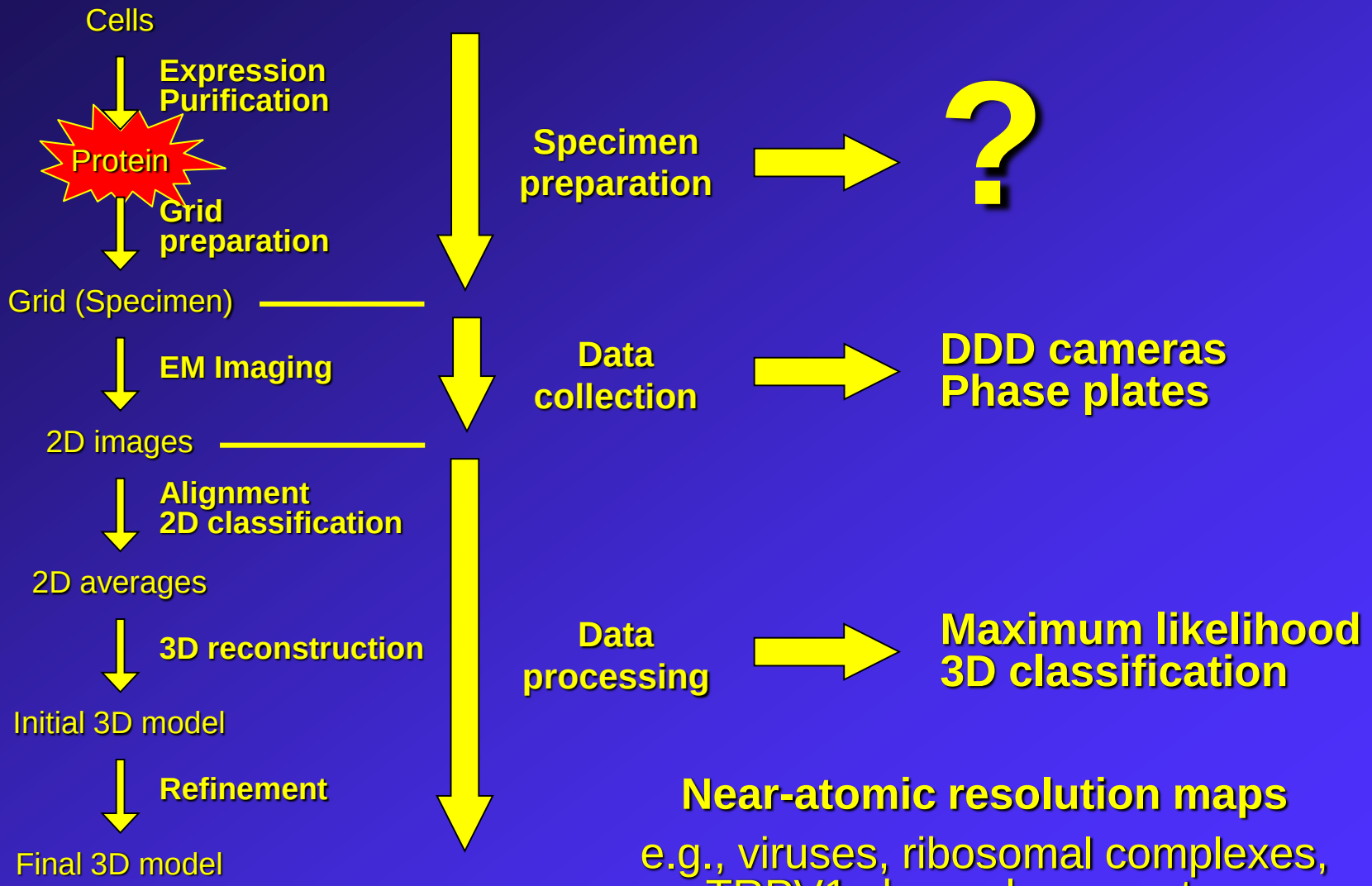
Workshop on Advanced Topics in EM Structure Determination: Where do we go from here?

Challenges remaining for specimen preparation

The Scripps Research Institute
La Jolla, November 2014

Single-particle EM workflow



Near-atomic resolution maps

e.g., viruses, ribosomal complexes, TRPV1 channel, γ-secretase

Why do we need specimen preparation?

Biological specimens consist of up to 80% of water

→ **COLLAPSE OF STRUCTURE** because of dehydration in EM vacuum

stain embedding	**BUT**: resolution limitation
vitrification	**BUT**: low contrast

Biological specimens consist of light atoms, such as C, N, O, H

→ **LOW CONTRAST** because electron scattering ~ atomic number Z

stain embedding	**BUT**: resolution limitation
high defocus	**BUT**: CTF correction required

→ **BEAM DAMAGE** because σ_{el}/σ_{in} = Z/19 (~ 2 inelastic per elastic scattering event)

short exposures	**BUT**: noisy images (low signal-to-noise ratio, SNR)
low temperature	**BUT**: only reduces the effects of beam damage

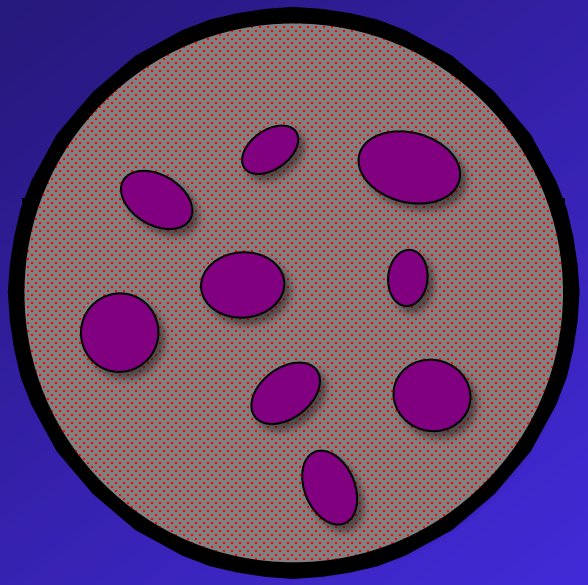
***Bottom line*:** negatively stained specimens: limited resolution but better SNR
vitrified specimens: "unlimited" resolution but poor SNR

Preparing good grids

Specimen preparation

EM grid
(copper, gold, molybdenum
new materials → Lori Passmore)

Carbon film
(continuous or holey;
new substrates
→ Lori Passmore)



Glow discharging
(renders carbon film hydrophilic,
negative or positive charge)

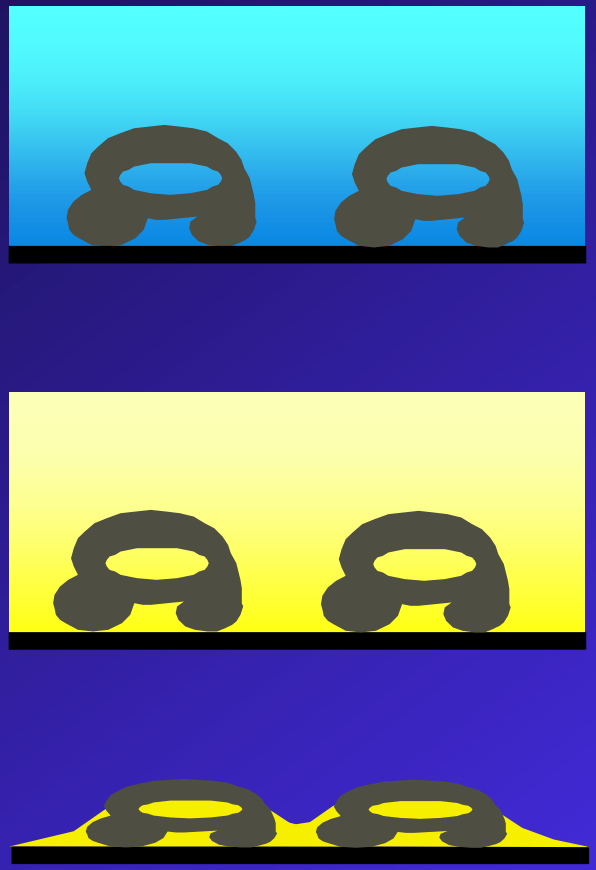
Apply specimen
(usually with pipette;
new approach: Spotiton
→ Clint Potter

Negative staining

Ice embedding (vitrification)

Negative staining

Ohi *et al.* (2004) *Biol. Proced. Online* 6: 23-34



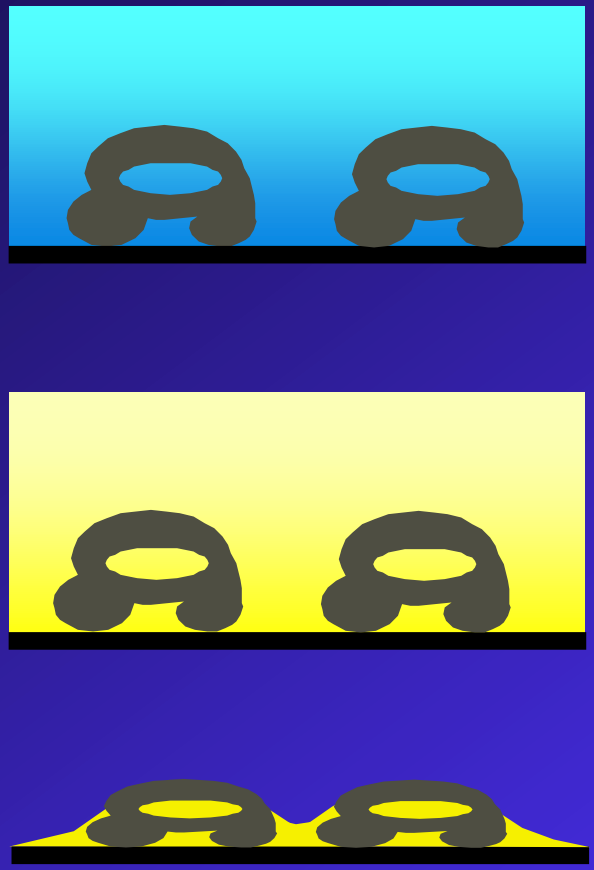
Stains:

uranyl formate (finer grain, but unstable)
uranyl acetate (coarser grain, but stable)
– higher contrast and radiation-resistant
– but acidic (pH ~4)
note, however, it also functions as fixative !

sodium silicotungstate
sodium phosphotungstate
ammonium molybdate
aurothioglucose
– neutral pH
– but lower contrast and less radiation-resistant

Negative staining

Ohi *et al.* (2004) *Biol. Proced. Online* 6: 23-34

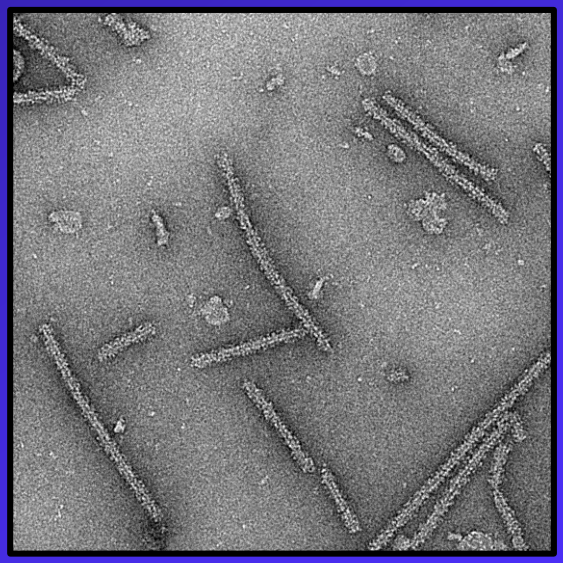


Stain embedding:

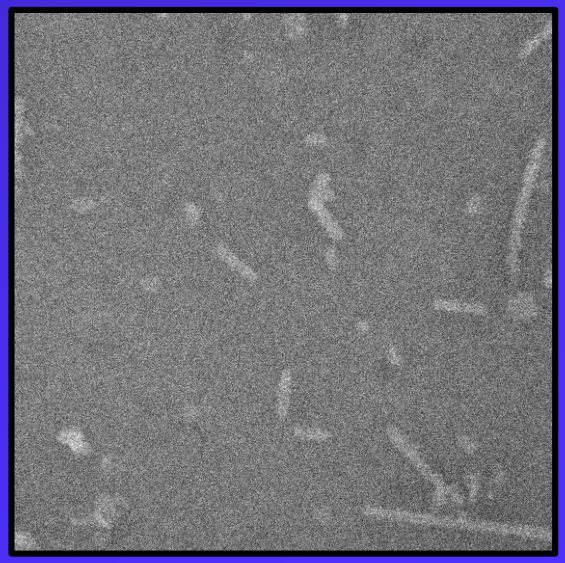
thin staining better for 2D analysis
– better contrast (important for small proteins)

thick staining better for 3D analysis
– better representation (important for 3D features)

thin staining

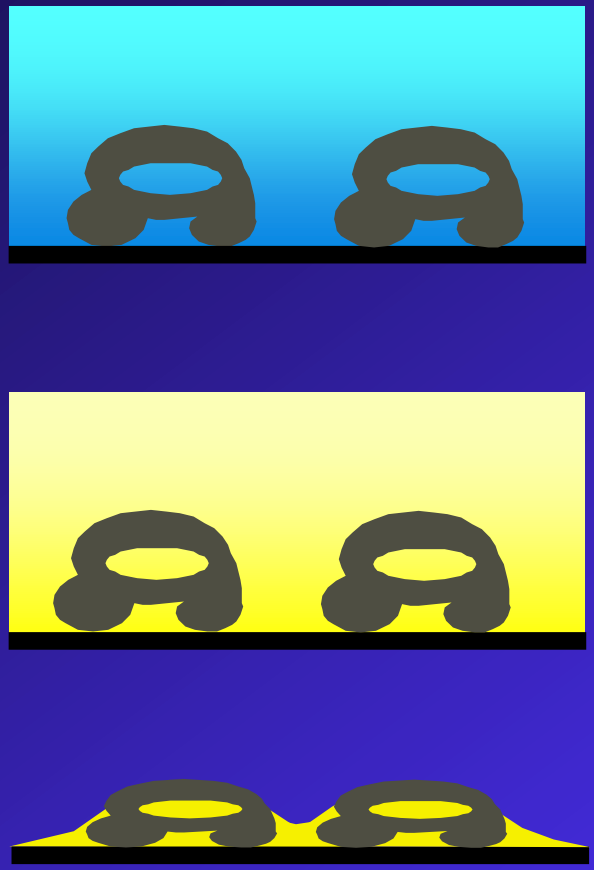


thick staining



Negative staining

Ohi *et al.* (2004) *Biol. Proced. Online* 6: 23-34



Pros:

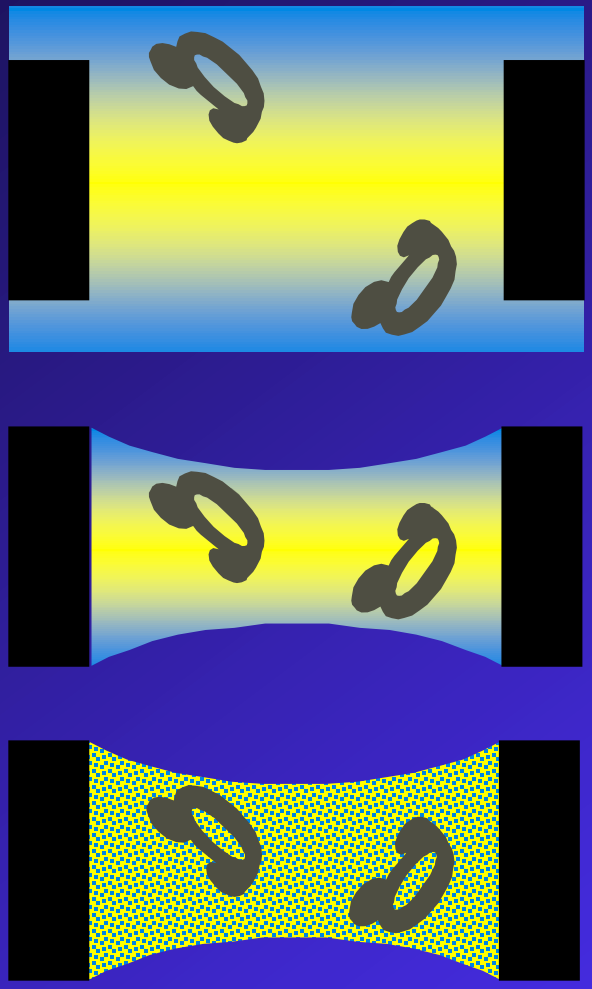
- easy and quick
- good contrast
- induces preferred orientations

Cons:

- limits resolution to ~20 Å
- introduces artifacts
  - incomplete stain embedding
  - adsorption deformation
  - specimen flattening upon drying
- induces preferred orientations

Cryo-negative staining approach 1

Adrian *et al.* (1998) *Micron* 29: 145-160



Pros:

- good contrast
- induces random orientations
- less preparation artifacts

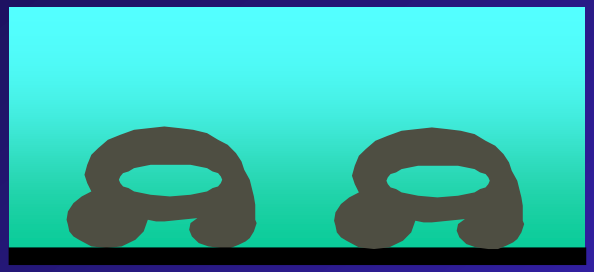
Freezing:
prevents specimen flattening

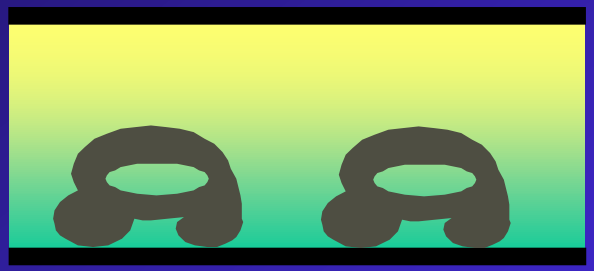
Cons:

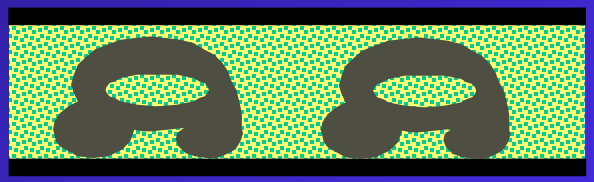
- limits resolution to ~20 Å
- induces random orientations
- high ionic strength can cause complexes to dissociate

Cryo-negative staining approach 2

De Carlo & Stark (2010) *Methods Enzymol.* 481: 127-145







Pros:

- good contrast
- induces preferred orientations
- less preparation artifacts

Addition of glycerol:
– minimizes adsorption artifacts
– minimizes specimen flattening
– serves as cryo-protectant

Carbon sandwich:
reduces incomplete stain embedding

Freezing:
prevents further specimen flattening

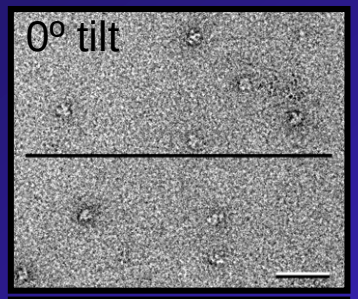
Cons:

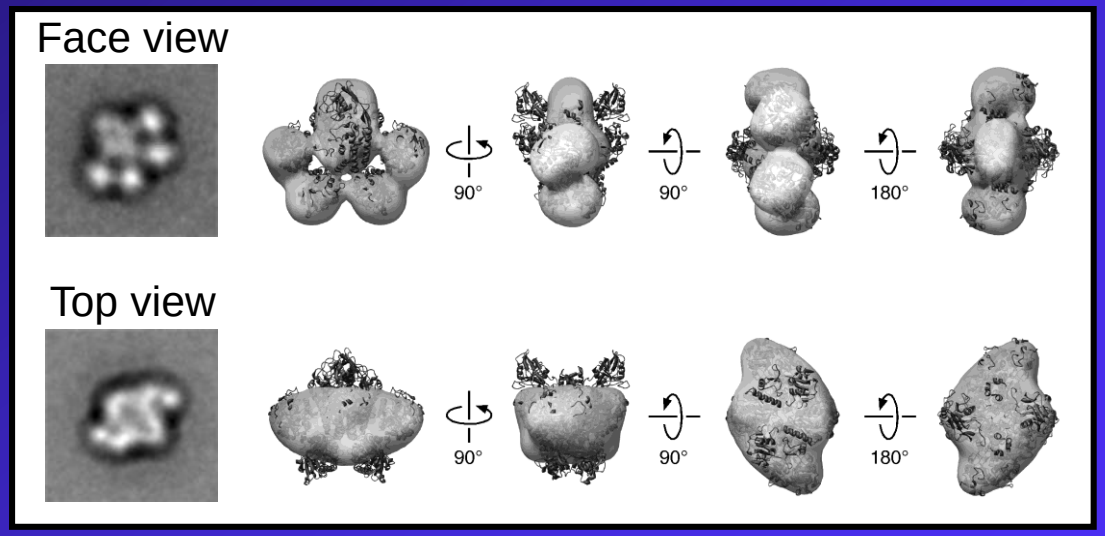
- limits resolution to ~20 Å
- induces preferred orientations
- really, really painful !!!

Conventional *versus* cryo-negative staining

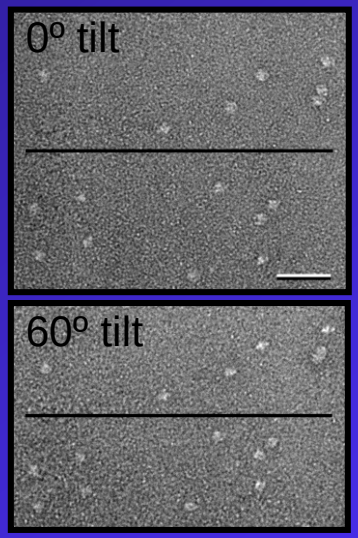
Random conical tilt reconstruction of the Tf-TfR complex

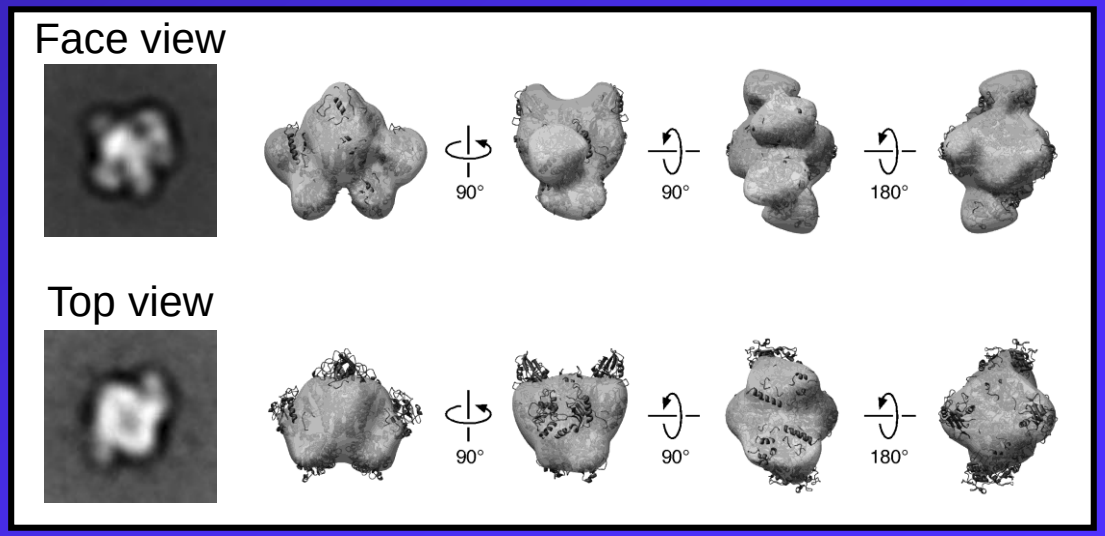
Conventional negative staining





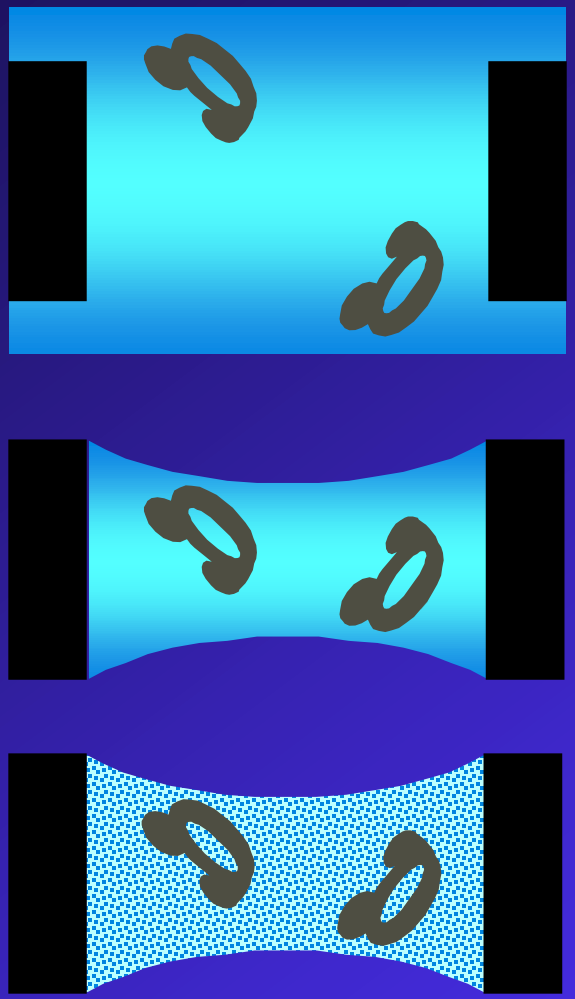
Cryo-negative staining





Vitrification

Adrian *et al.* (1984) *Nature* 308: 32-36



Pros:

- near-native conditions
- no limitation on resolution
- induces random orientations

Cons:

- low contrast
- not very reproducible
- more tedious
- induces random orientations

What is a good cryo-EM grid?

good amorphous ice
– not crystalline ice
– no “leopard skin” pattern
– no contamination

appropriate ice thickness
– typically as thin as possible

clearly visible particles
– particle size and shape
– buffer composition
– defocus, movie mode, phase plates

good particle distribution
– in holes
– dense but particles not touching
– randomly distributed orientations

Variables in grid preparation

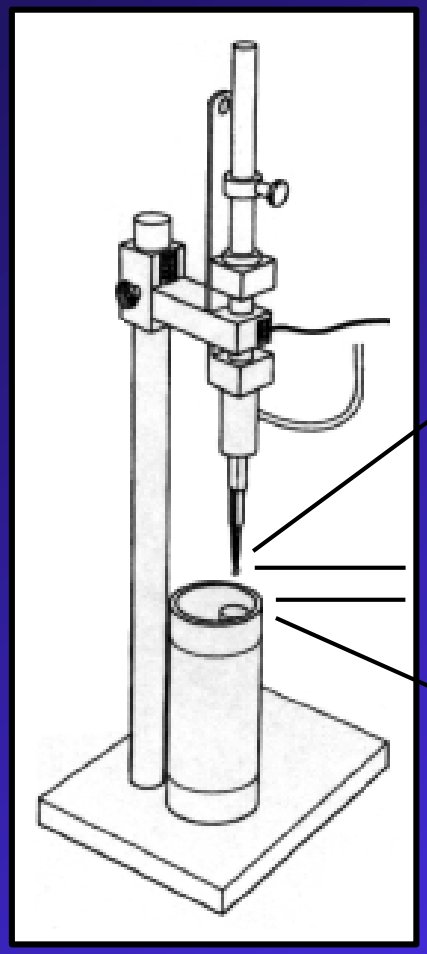
grid
– type of grid and substrate
– batch and age of grid
– glow discharging

sample
– concentration
– buffer composition
– detergent or “others”

freezing (blotting & drying)
– blotting time (physical water removal)
  – single- or double-side blotting
– waiting time (evaporation)
  – temperature & humidity
– multiple sample applications

Good amorphous ice

Not crystalline ice



Accel-
eration

to prevent crystallization, temperature has to decrease faster than 10^5-10^6 K/s
Dubochet *et al.* (1988)
Q. Rev. Biophys. 21: 129-228

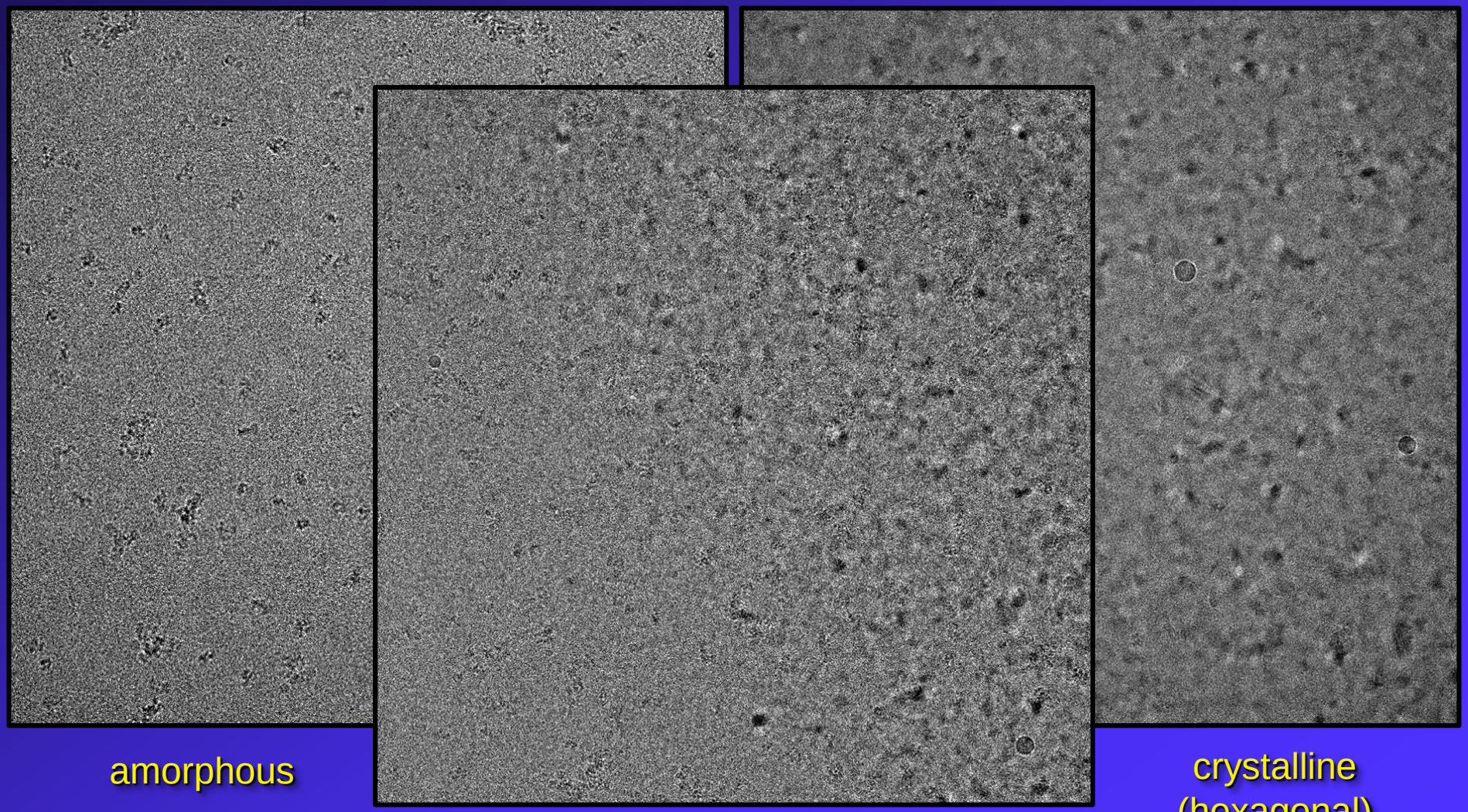
thermal conductivity of sample
– water is poor thermal conductor
→ sample has to be thin

] Distance

wetting characteristics of coolant
– LN_2 bad (Leidenfrost effect)
→ ethane cooled by LN_2

Good amorphous ice

Not crystalline ice



amorphous

crystalline
(hexagonal)

Good amorphous ice

Not crystalline ice

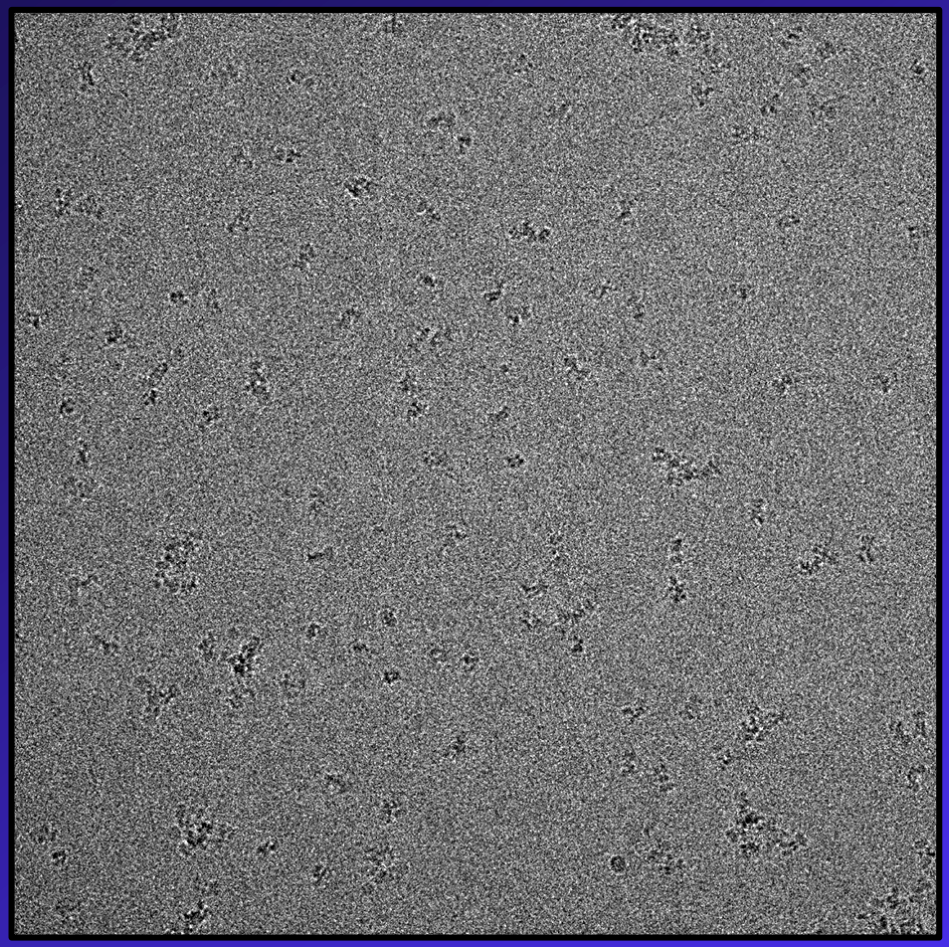
Homemade Plungers

FEI Vitrobot

Gatan Cryoplunge

Good amorphous ice

No “leopard skin” ice



“normal” ice

“leopard skin” ice

Good amorphous ice

No “leopard skin” ice

aka: alligator ice
snake skin ice
turtle ice

3D EM list

The leopard skin is comprised of nano-ice/salt crystals, they might come from slow ice contamination in high vacuum environment with small leaks which were observed by number of labs, and another possibility is the solution containing certain type of salts/agents which are precipitated or crystallized during freezing, the third one could be the cooling liquid was too 'warm'.
You may try to freeze your sample at different lab or freeze a simple buffer like 20 mM NaCl using your Vitrobot.

Shall we discuss ?

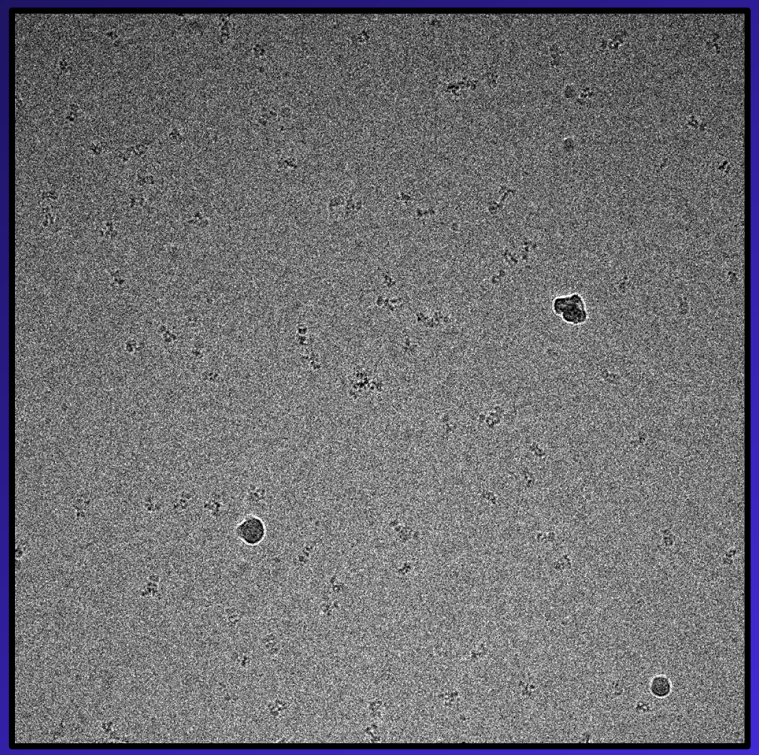
We would see it from time to time, occasionally when an ethane tank was nearing empty.
As such we attributed to impurities that impeded freezing.
It could also be just something in the buffer or sample that alters the freezing.

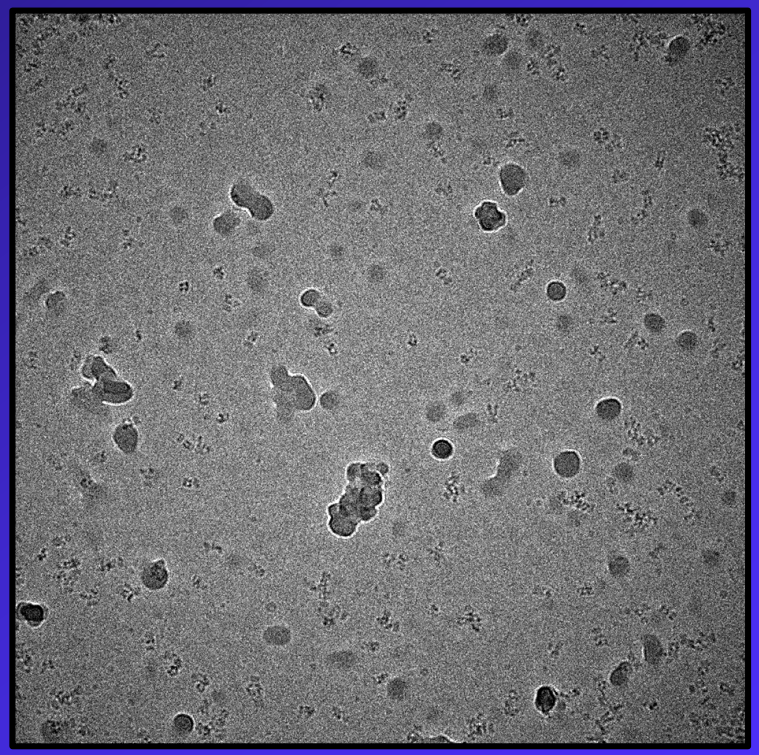
My experience is also that it comes and goes and does not have much to do with the vitrification itself or the grids. More likely to me is also mild exposure to "warm" air or surfaces during transfer. Having the goniometer opened and closed during inspection also sometimes affected the behaviour. My latest trick is pre-pumping the airlock on our T20 more than once before transfer of the holder.

I believe the reason is minor exposure to warm & moist air, typically during the transfer of the cryo holder.

Good amorphous ice

No contamination





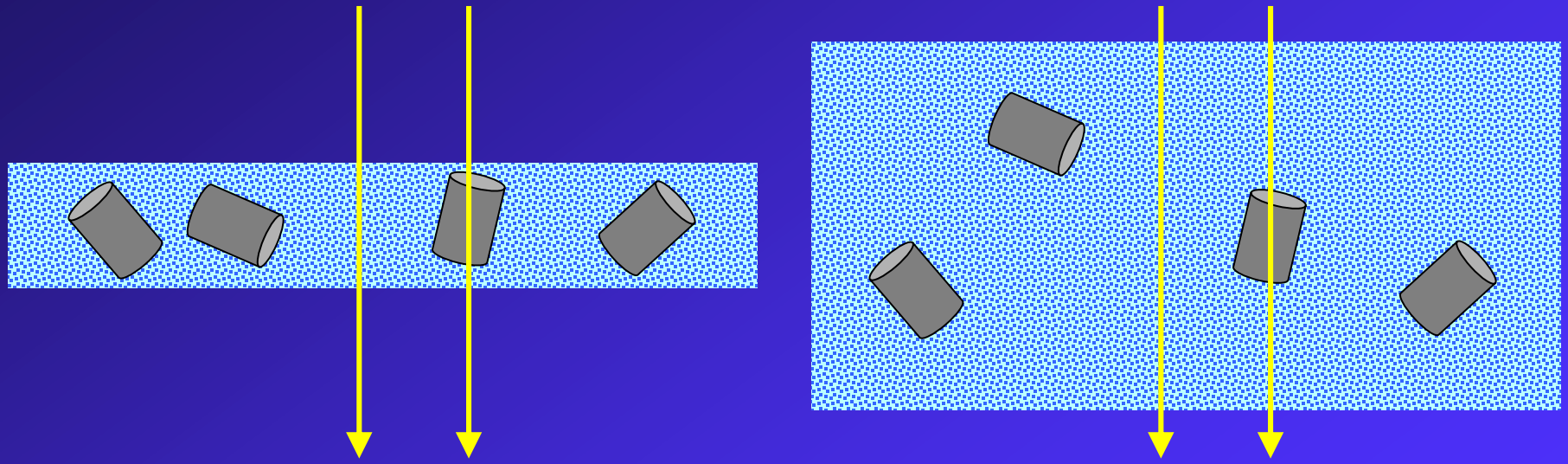
– can happen during grid transfers due to air humidity
  → fast transfers
  → low-humidity environment

– can be due to water in LN_2 used for grid preparation or storage
  → use narrow neck dewars and keep dewars dry

Appropriate ice thickness

Thin ice is usually better

density of protein: ~1.36 g/cm^3
density of pure water: 1.00 g/cm^3



thin ice provides better contrast, which is especially important for small proteins

Appropriate ice thickness

Factors that affect ice thickness

– thickness of the carbon film
→ can be increased by evaporating extra carbon on grid

– hydrophobicity of carbon film
→ can be changed by glow discharging

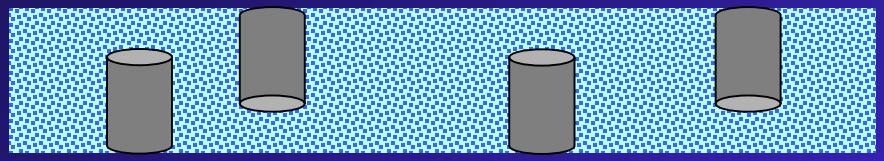
– blotting
→ time of blotting
→ single- or double-sided blotting

– time between blotting and freezing (evaporation)
→ can be controlled by temperature and humidity
→ however: only water evaporates → changes buffer

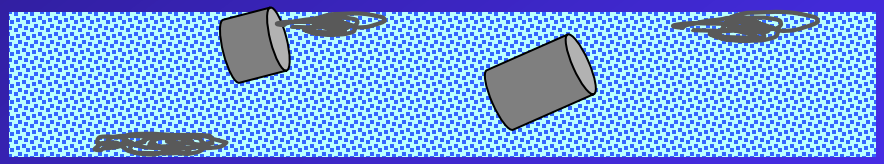
Optimal combination has to be determined empirically for every new sample (and grid batch)

Appropriate ice thickness

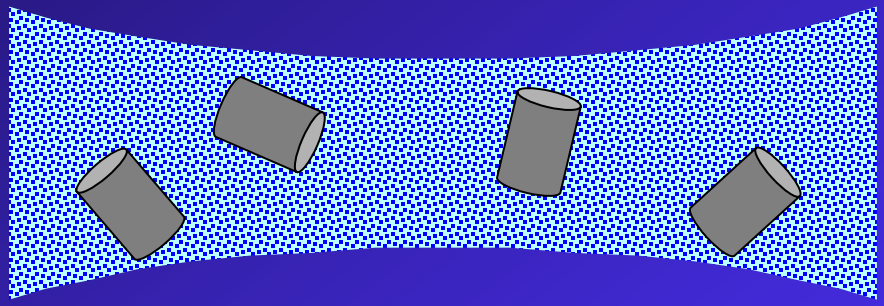
Why thicker ice can sometimes be better

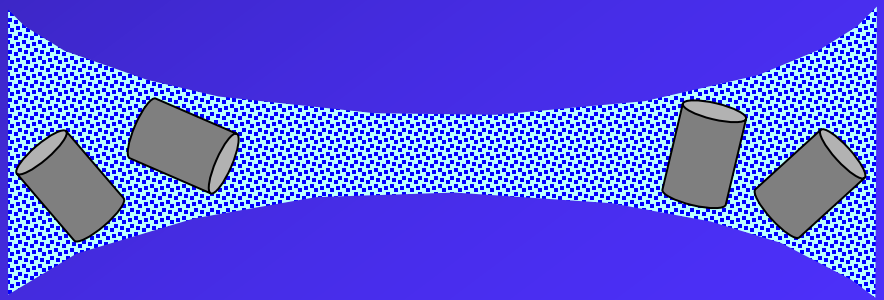


air/water interface can induce proteins or complexes to adopt preferred orientations



air/water interface can induce proteins to denature or complexes to dissociate





– problem greater for samples with detergents (lower surface tension)
– problem can be alleviated by using carbon with a smaller hole size

Clearly visible particles

Factors that affect visibility of particles (contrast)

– particle size and shape
  → need to scatter sufficient electrons to be visible
    (even if particles can be seen, it does not mean that they can be aligned)
  → globular particles are easier to see than extended particles of same MW

– ice thickness
  → should be as thin as possible

– buffer composition
  → density of protein: 1.36 g/cm^3
    density of pure water: 1.00 g/cm^3
    density of glycerol: 1.26 g/cm^3

beware of high concentrations of:
– glycerol
– sugars
– salt
– detergent

Clearly visible particles

How to improve the visibility of particles (contrast)

– negative staining
  → all the known problems (limited resolution, deformations, …)

– high defocus
  → limits achievable resolution

– record long movies
  → use full movie for processing
  → use less (or weigh) frames for final reconstruction (RELION version 1.3)

– phase plates
  → presentations by Rado Danev and Wah Chiu

Good particle distribution

Why are the bloody particles not in the bloody holes ?
And how can I get them there ???

– sample is too dilute
→ use higher protein concentration
→ adsorb to a thin carbon (or graphene) film or a lipid monolayer

– protein/complex prefers to stick to carbon film
→ change grid batch or vary glow discharge conditions
→ adsorb to a thin carbon (or graphene) film or a lipid monolayer
→ apply specimen twice

– protein/complex is too big for the thickness of the carbon film
→ evaporate carbon onto holey carbon grid

– protein/complex denatures or dissociates on air/water interface
→ use chemical fixation
→ adsorb to a thin carbon (or graphene) film or a lipid monolayer

Notes: – carbon film and lipid monolayer will reduce image contrast
– adsorption to any substrate can induce preferred orientations

Good particle distribution

Factors that affect particle distribution

– sample concentration
→ not linear

– ice thickness
→ varies across the grid and even within a hole
→ more problematic for samples containing detergent
(possible advantage of using amphipols and Nanodiscs)

– buffer composition
→ aggregation through ionic interactions
→ change salt concentration
→ aggregation through hydrophobic patches
→ add detergent

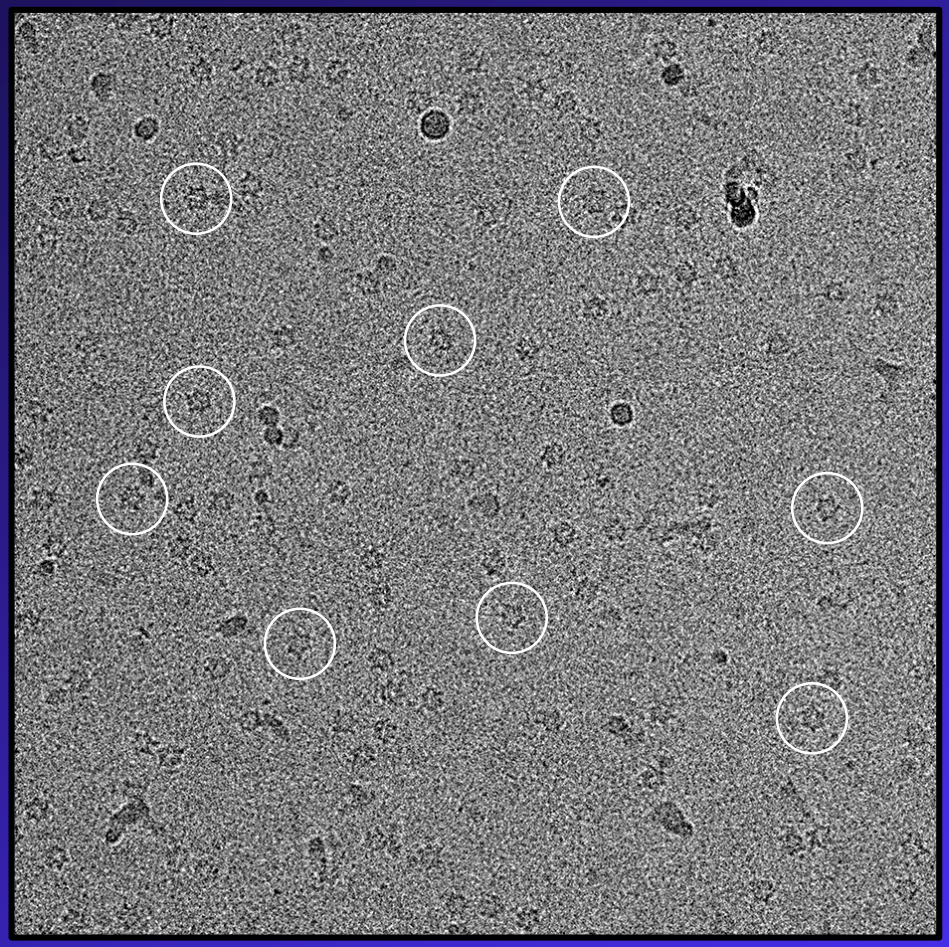
Good particle distribution

What to do when the particles adopt preferred orientation

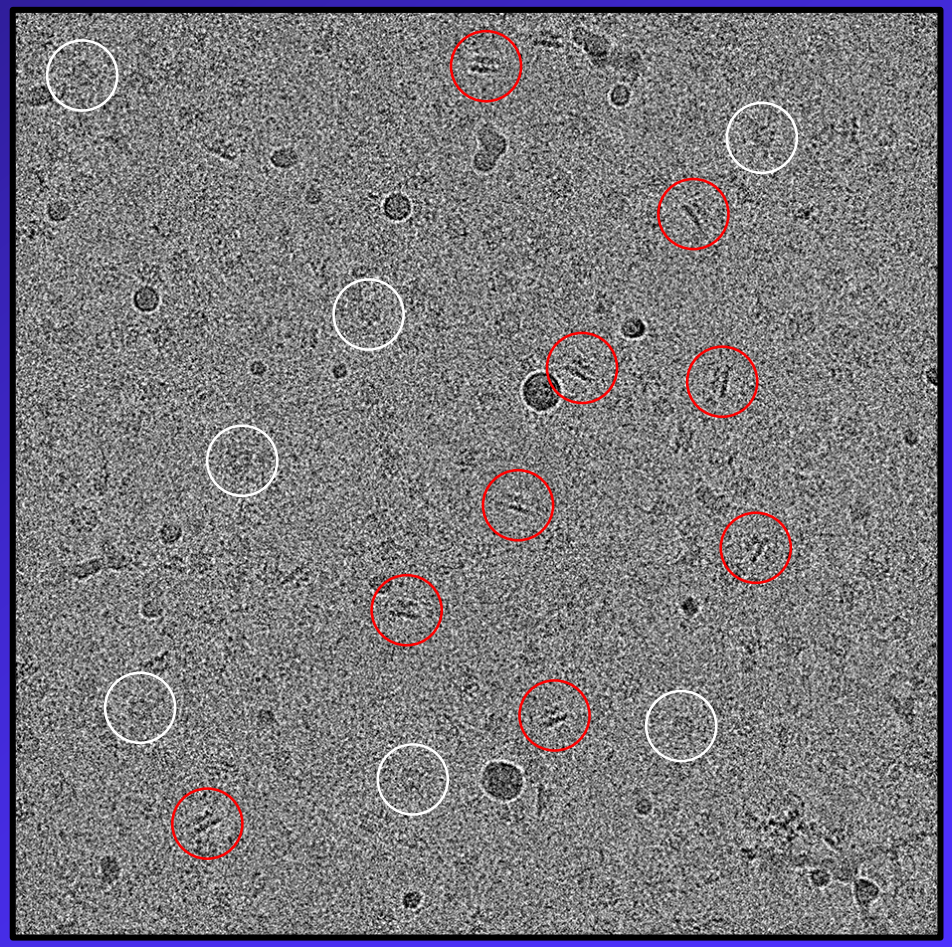
– try thicker ice

Good particle distribution

What to do when the particles adopt preferred orientation



thin ice



thick ice

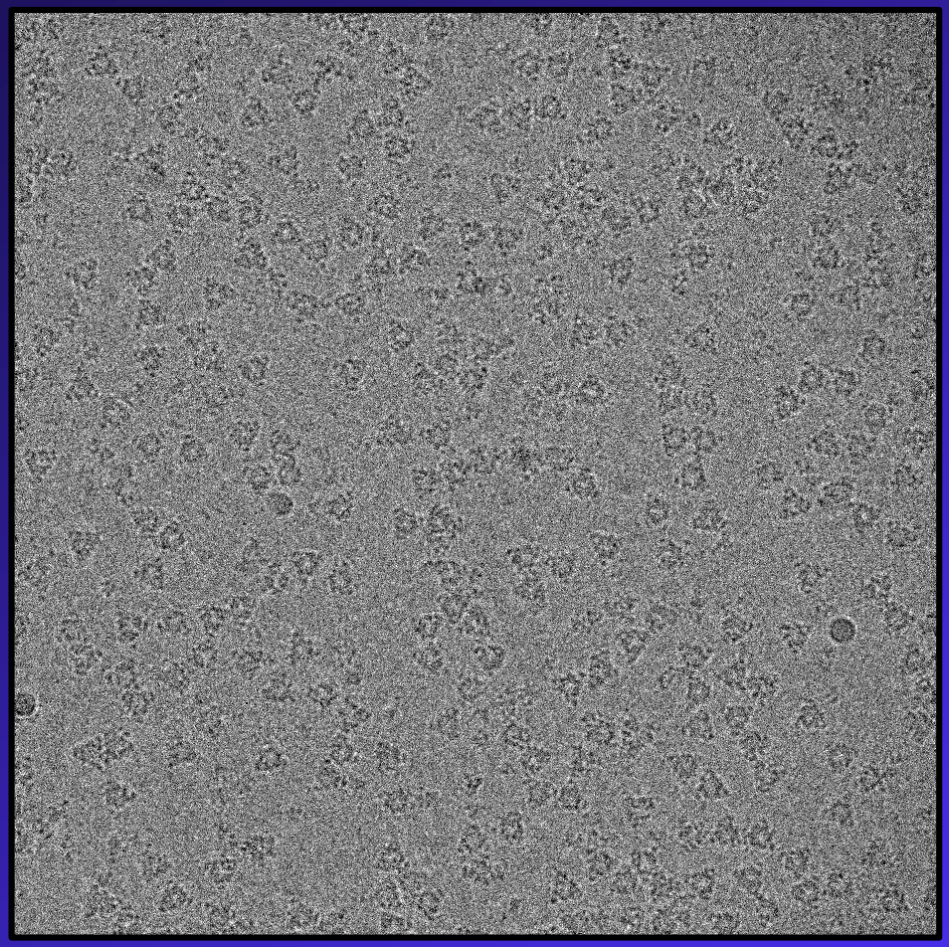
Good particle distribution

What to do when the particles adopt preferred orientation

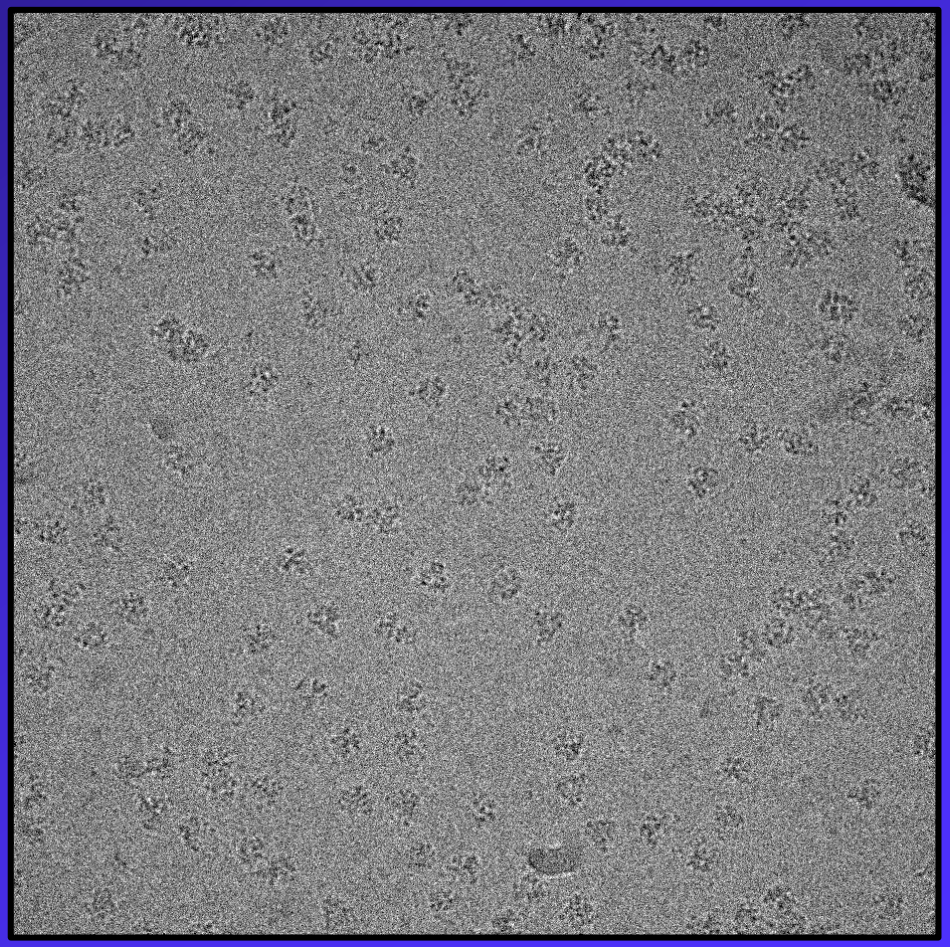
– try thicker ice
– try to add some detergent

Good particle distribution

What to do when the particles adopt preferred orientation



without detergent



with detergent

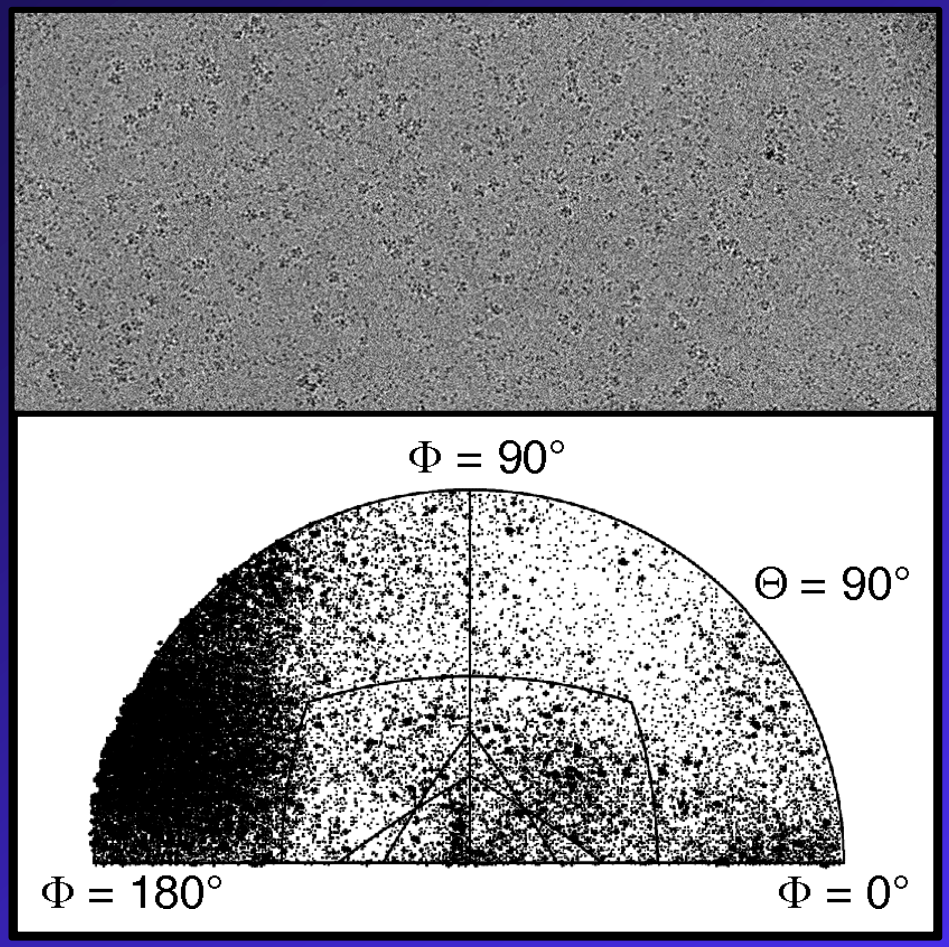
Good particle distribution

What to do when the particles adopt preferred orientation

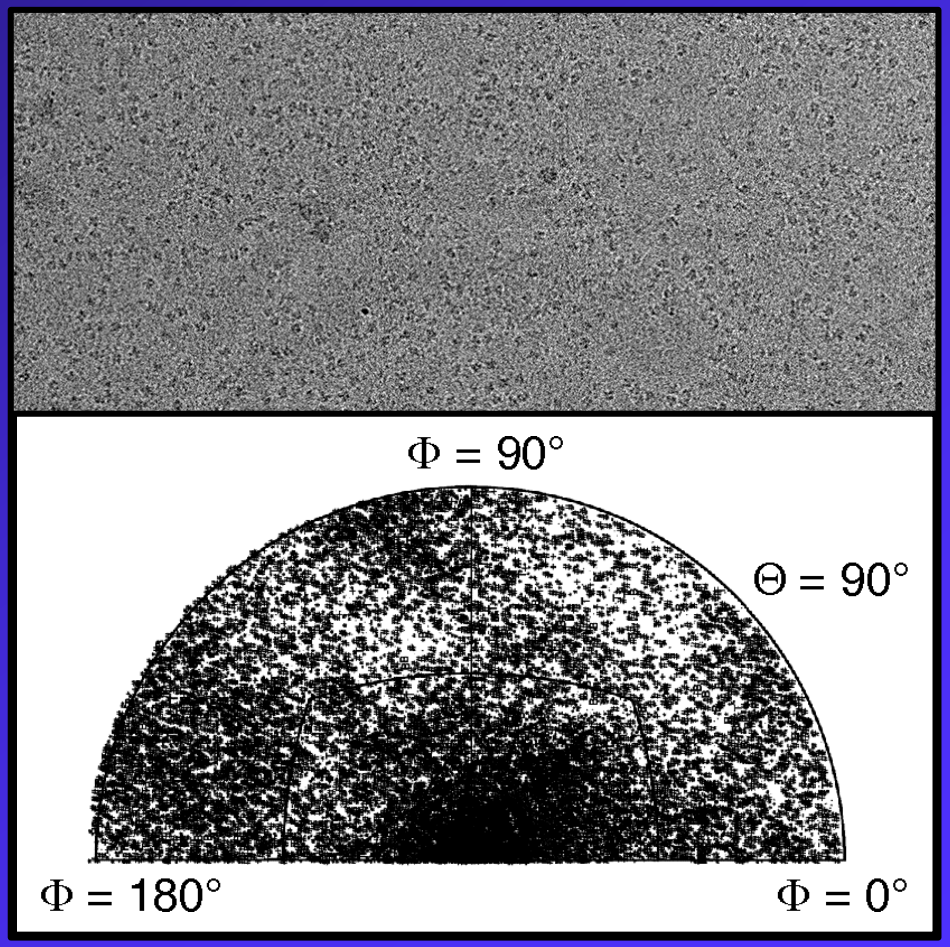
– try thicker ice
– try to add some detergent
– try to adsorb to a substrate

Good particle distribution

What to do when the particles adopt preferred orientation



without carbon



on thin carbon

Good particle distribution

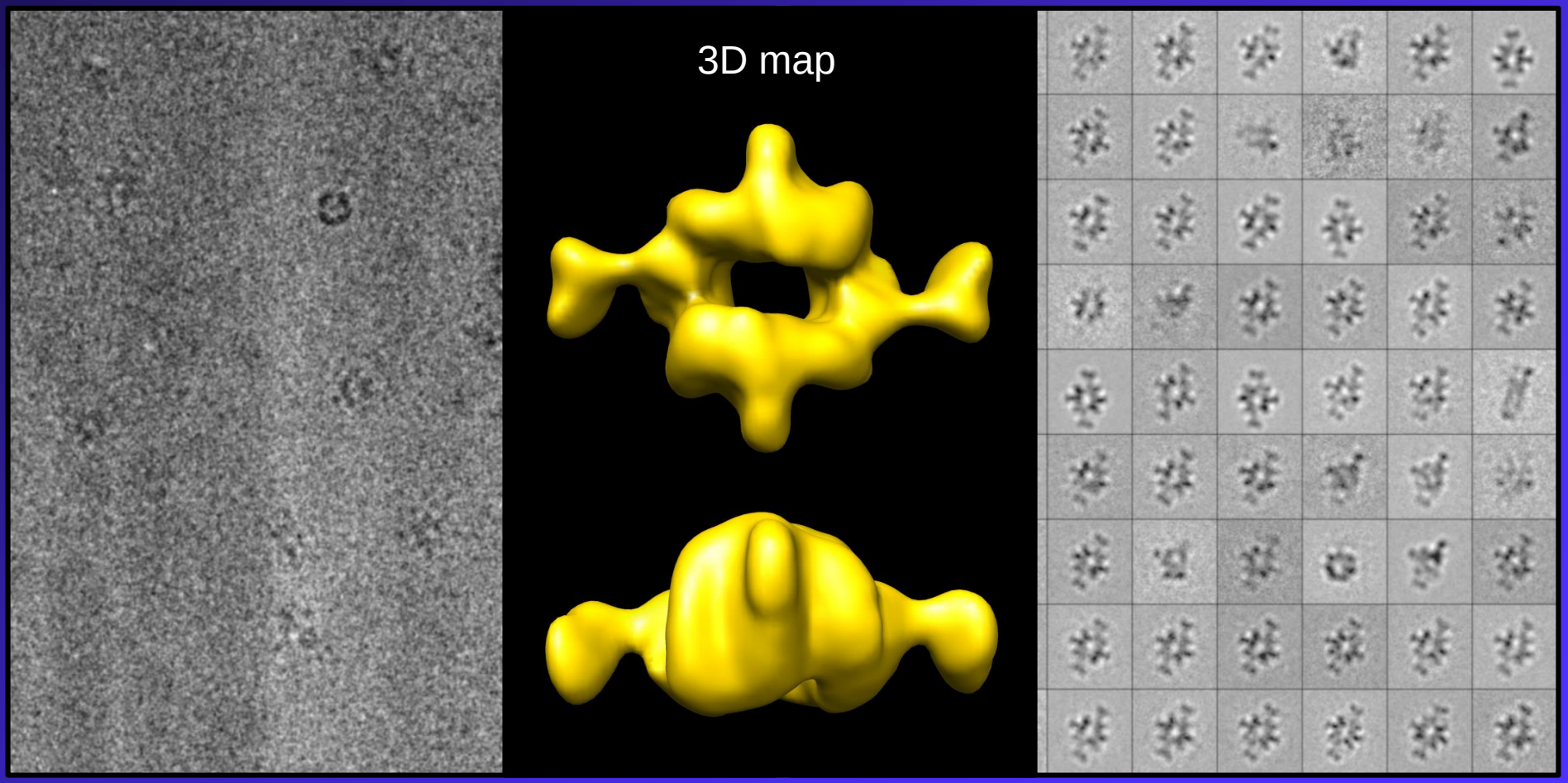
What to do when the particles adopt preferred orientation

– try thicker ice
– try to add some detergent
– try to adsorb to a substrate
– try to change the buffer composition
– try to change glow discharge conditions
– try to add tags to the protein

– give up and tilt the bloody grid
… Have fun !
… Kiss your high resolution good-bye !

Good particle distribution

What to do when the particles adopt preferred orientation



raw image (on film)

class averages

The perfect grid

High-contrast particles that are perfectly distributed and adopt randomly distributed orientations

The best images

Thon rings in all directions beyond the Nyquist frequency

The ideal processing

You made it all the way to the “Publish” button in RELION

Nice map ...

but where are the side chains ???

Preparing good protein (for EM)

What is good protein for EM?

size

the bigger the better
~250 kDa currently minimum for near-atomic resolution

symmetry

the higher the better
pseudo-symmetry can be problematic

shape

globular better than extended
“extra features” highly beneficial

homogeneity

the more homogeneous the better, but
heterogeneity now more manageable due to
– better image quality (DDD cameras)
– new software tools (e.g., 3D classification)

What is good protein for EM?

best case scenario:

large, globular and highly symmetric molecule with little heterogeneity
→ viruses, virus-like particles
→ atomic model almost guaranteed

typical samples → resolution ???

Shall we discuss ?

worst case scenario:

small, extended and asymmetric molecule with high degree of heterogeneity
→ our samples (tethering complexes, cell-surface receptors, etc.)
→ condemned to negative-stain EM studies

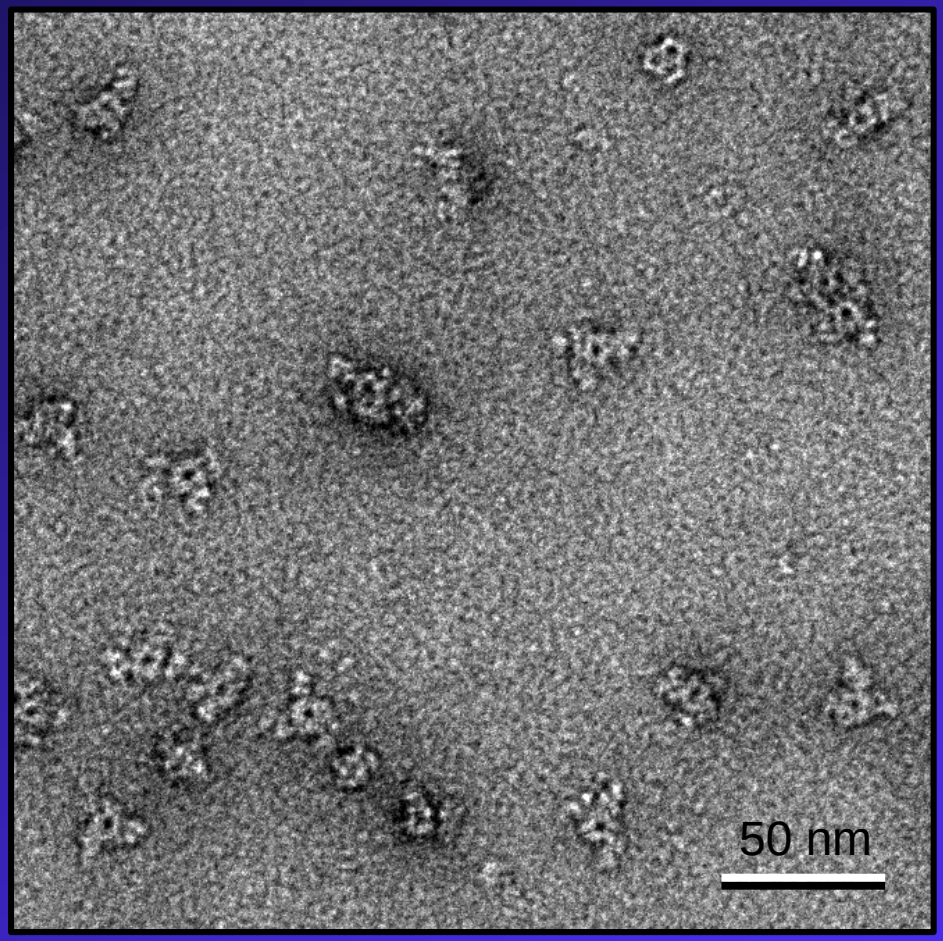
different characteristics can compensate for each other

– a large molecule does not need high symmetry → e.g., ribosome
– a large molecule can tolerate some heterogeneity → e.g., ribosome
– a highly symmetric molecule can be smaller → e.g., some filaments

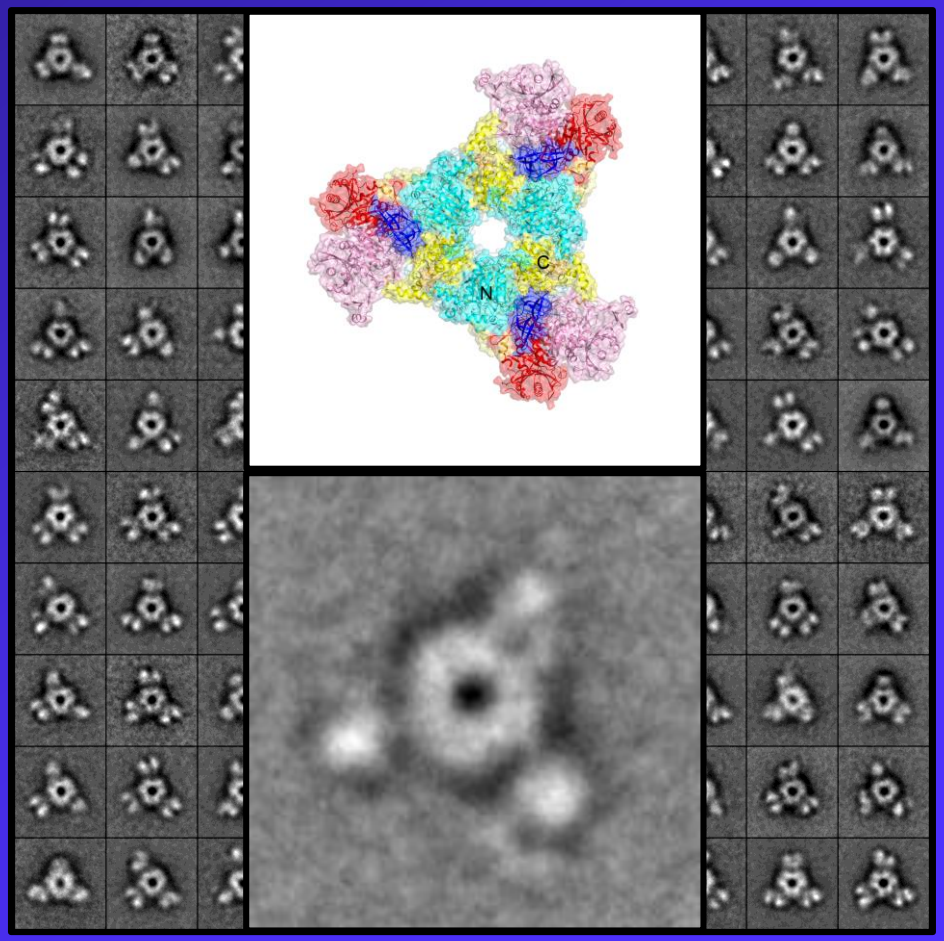
Protein size

Big and oligomeric is not always enough for cryo-EM !

Long-chain acyl-CoA carboxylase
120 kDa, hexamer → 720 kDa

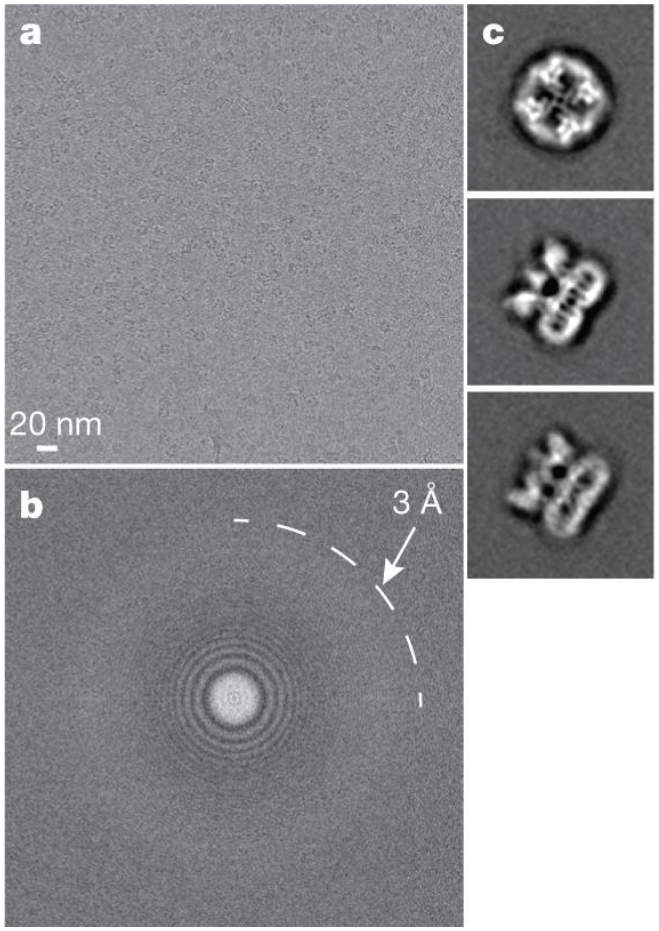


Tran *et al.* (2014)
Nature, in press



Protein size

Small is not what it used to be for cryo-EM !

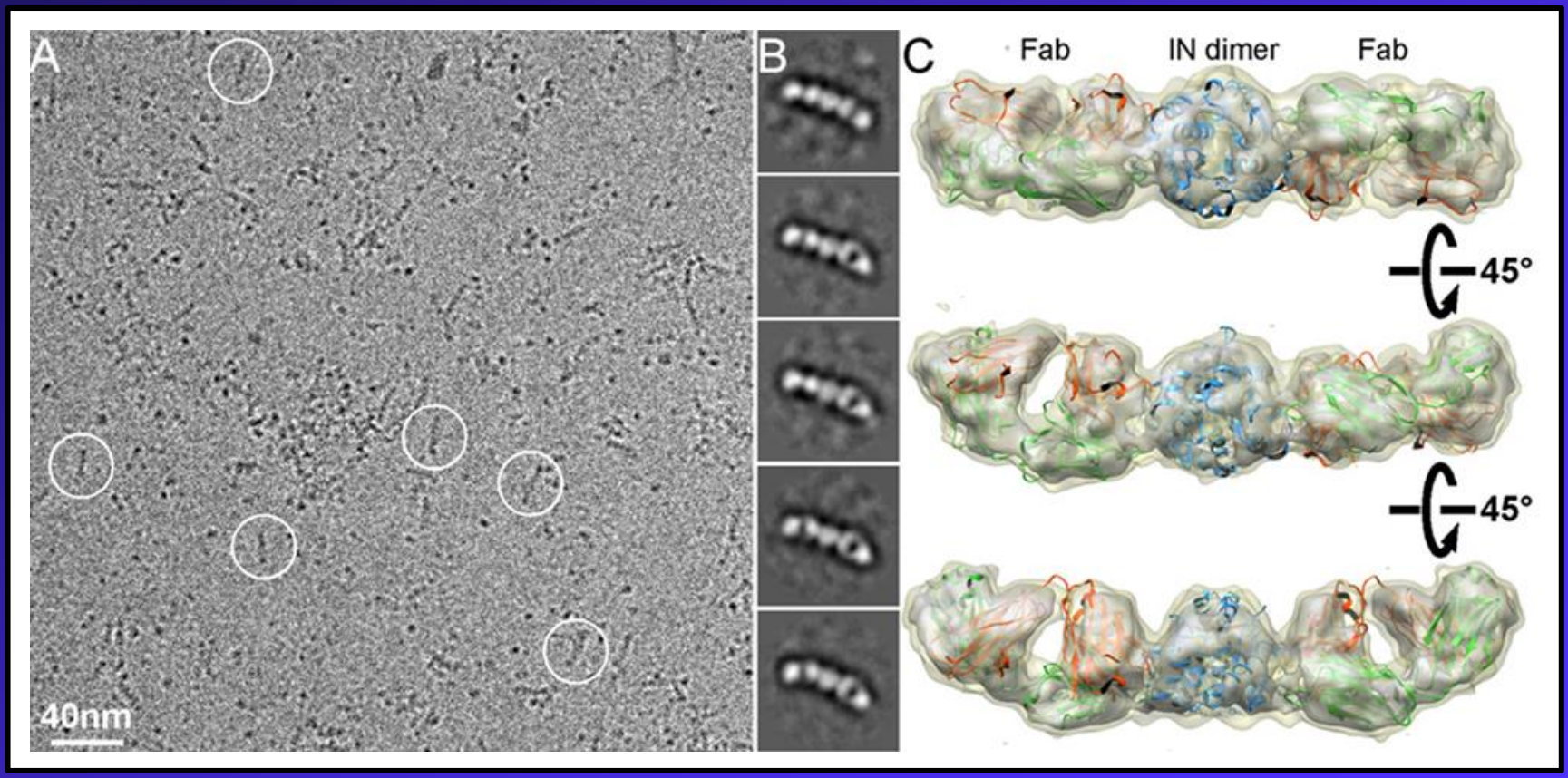


TRPV1 channel at 3.4 Å

Liao *et al.* (2013) *Nature* 504: 107-112

Protein size

What if the protein is well-behaved but too small ?
Fab labeling – Wu et al. (2012) *Structure* 20: 582-592



HIV-1 integrase dimer (~65 kDa) in complex with two Fabs (total of ~165 kDa) at a resolution of 13.3 Å (FSC = 0.5) , 10.2 Å (FSC = 0.143)

Protein size

What if the protein is well-behaved but too small ?
Fab labeling – Wu et al. (2012) *Structure* 20: 582-592

Fab labeling is a powerful approach for small proteins

– Fab increases particle size
– Fab adds an additional marker for alignment
– Fab density provides an inherent quality control of the 3D reconstruction

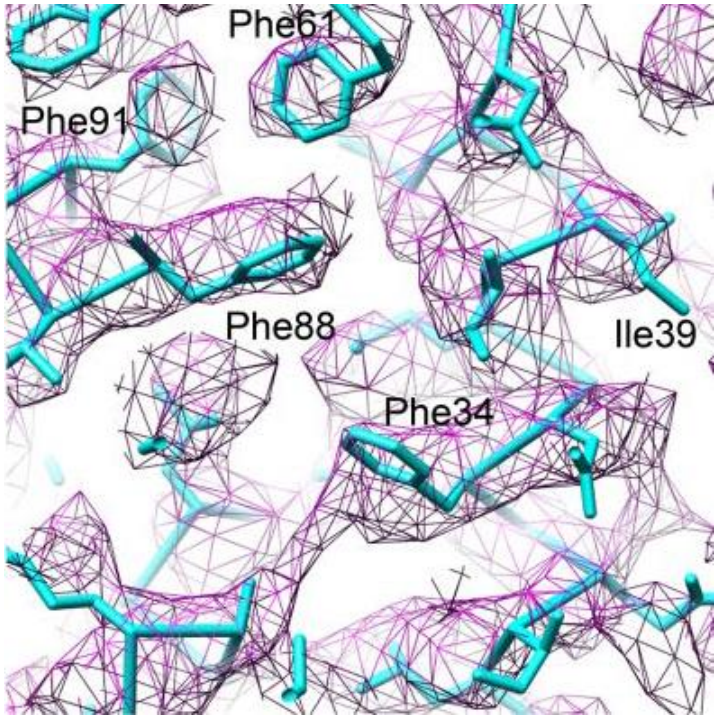
It can be tricky to find a suitable Fab

– (usually) should not affect protein function
– should have tight binding (low off rate)
– should not introduce structural variability

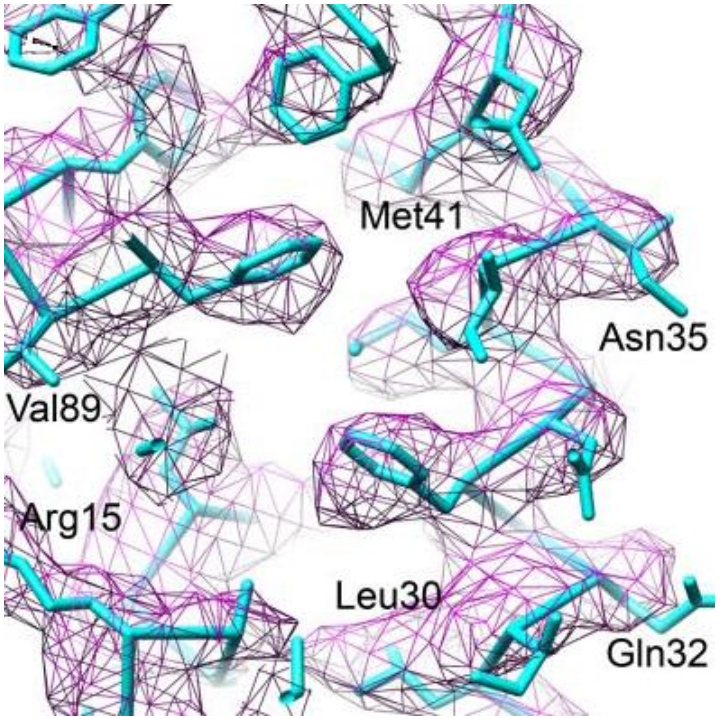
Protein symmetry

True symmetry is always helpful

Rotavirus double-layer particle



icosahedral averaging only



icosahedral averaging and
13-fold non-icosahedral averaging

Zhang *et al.* (2008) *Proc. Natl. Acad. Sci. USA* 105: 1867-1872

Protein symmetry

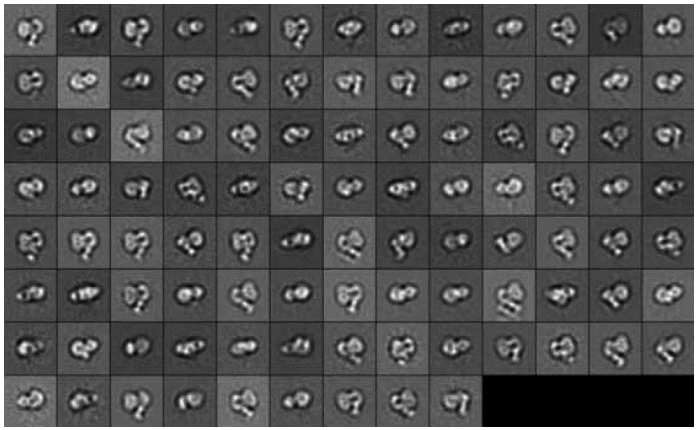
Pseudo-symmetry can be problematic

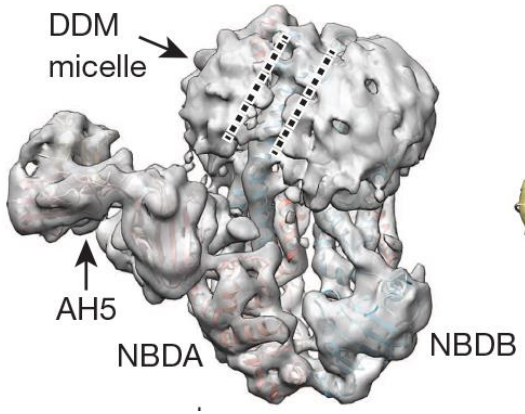
The problem pseudo-symmetry causes depends on the combination of:
– degree of deviation from true symmetry
– size of molecule
– resolution of map

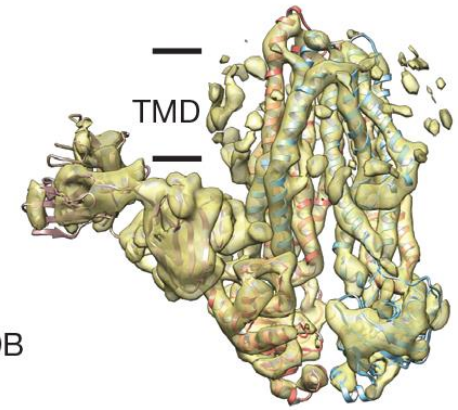
Determining whether symmetry is "true" or "pseudo" can potentially be sorted out by computational means
→ AAA+ ATPases
→ presentation by Frank DiMaio

Fabs again !

heterodimeric ABC exporter – TmrAB (~135 kDa) at 8.2 Å resolution







Kim *et al.* (2014) *Nature.* Epub ahead of print.

Protein shape

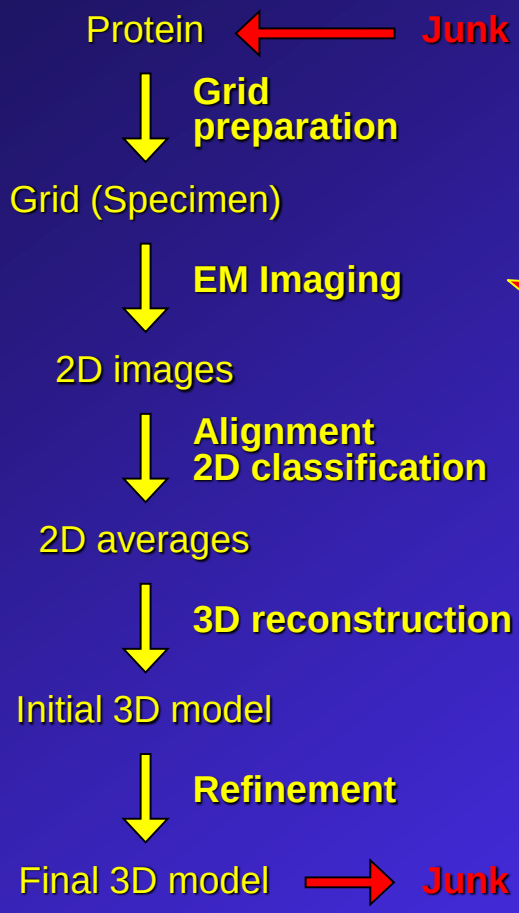
For proteins of same MW:
– globular protein easier to see
– extended protein easier to align

“extra features” highly beneficial for alignment (see Fabs)

Protein homogeneity

The importance of biochemistry !

The universal principle:
Junk In → Junk Out



Invest time to optimize protein purification !

Every sample is heterogeneous – it is just a question to what extent !

If the sample is good, the rest is now “easy”

Check the quality of the protein by:

SDS-PAGE	→ somewhat informative
Gel filtration	→ somewhat more informative
Negative-stain EM	→ a picture is worth more than a thousand gels and profiles

Protein homogeneity

The advantages (and pitfalls) of negative-stain EM images

SDS-PAGE

information on sample composition
– contaminations
– which subunits of complexes
– degradation of proteins

Negative-stain EM

additional information on
– whether contaminations are troublesome
– whether all complexes have all subunits or a mixture of different subcomplexes

Gel filtration

information on sample homogeneity
– sharp, symmetric peak
→ compositional homogeneity
– broad peak and/or shoulders
→ aggregation and/or instability

Negative-stain EM

additional information on
– shape of particles
→ conformational homogeneity
– size of particles
→ extent of aggregation and/or compositional instability

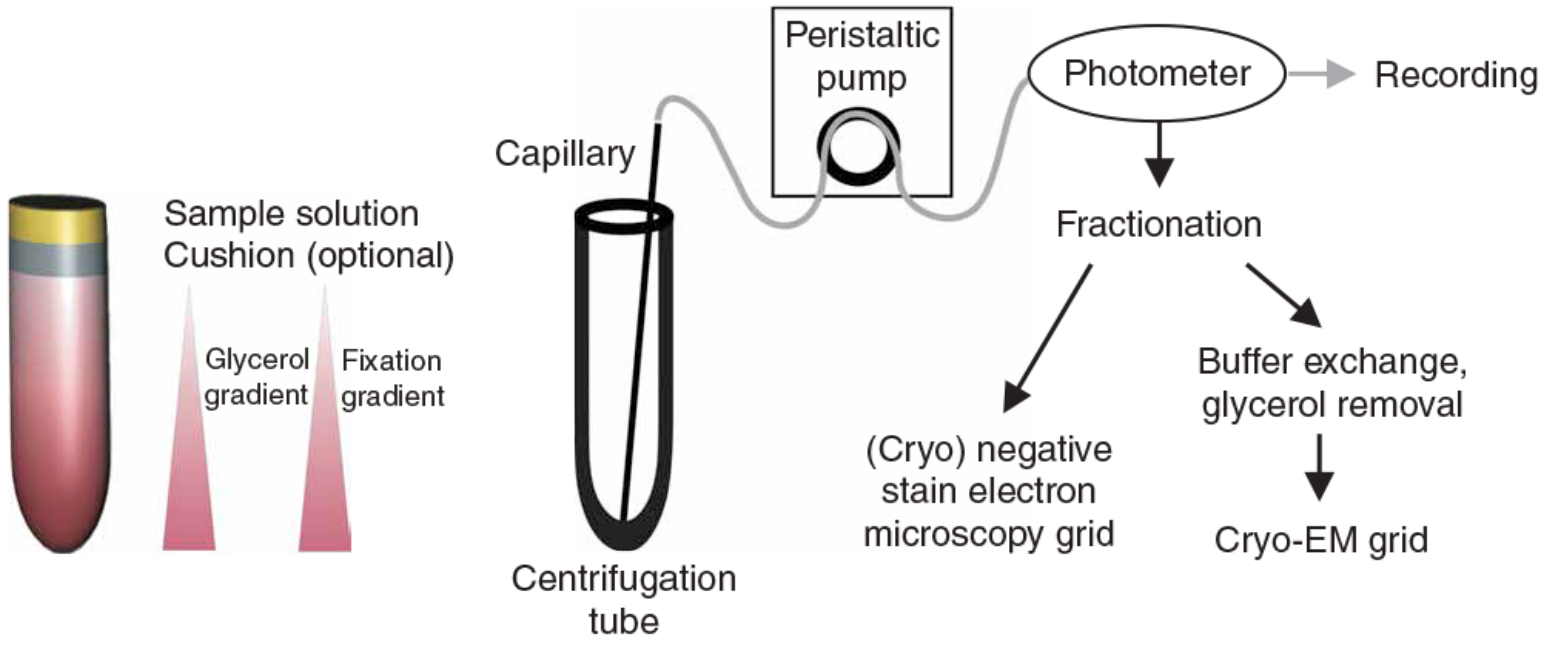
– adsorption to carbon film can sometimes induce artificial heterogeneity
– different shapes does not necessarily mean heterogeneity (different orientations)
– heterogeneity does not necessarily mean protein does not form 3D crystals (ordering effect of crystal lattice)

Protein homogeneity

Optimize compositional homogeneity – chemical fixation

Cross-linking with low concentrations of glutaraldehyde
→ optimize cross-linking conditions (assess by SDS-PAGE and negative-stain EM)

GraFix – sophisticated cross-linking approach



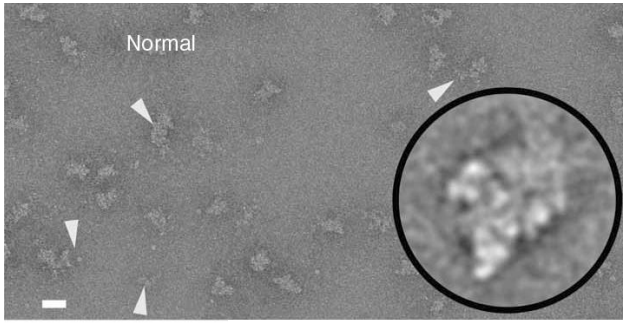
Kastner *et al.* (2008) *Nat. Methods* 5: 53-55

Protein homogeneity

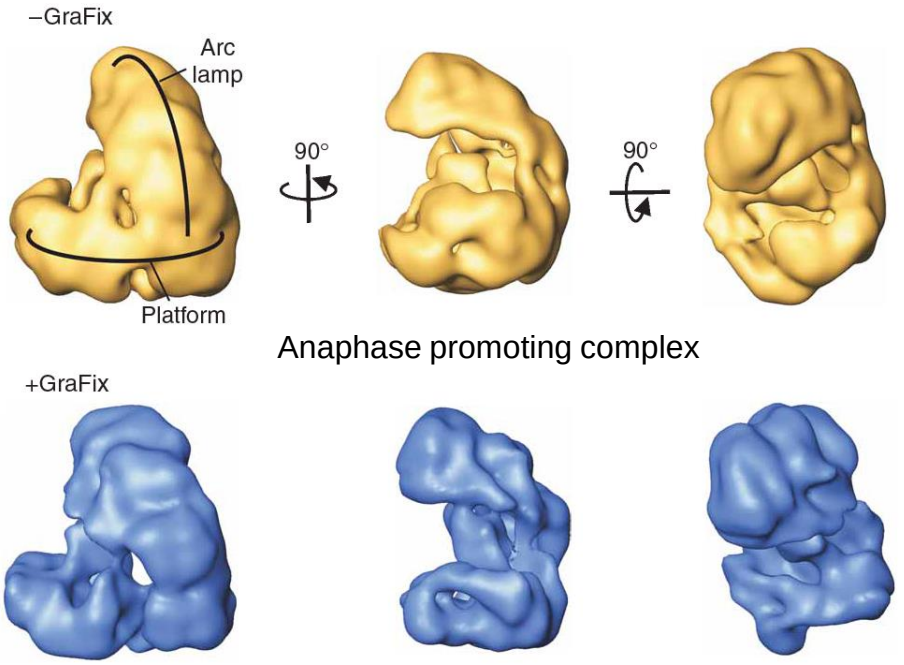
Optimize compositional homogeneity – chemical fixation

Cross-linking with low concentrations of glutaraldehyde
→ optimize cross-linking conditions (assess by SDS-PAGE and negative-stain EM)

GraFix – sophisticated cross-linking approach



B complex spliceosomes



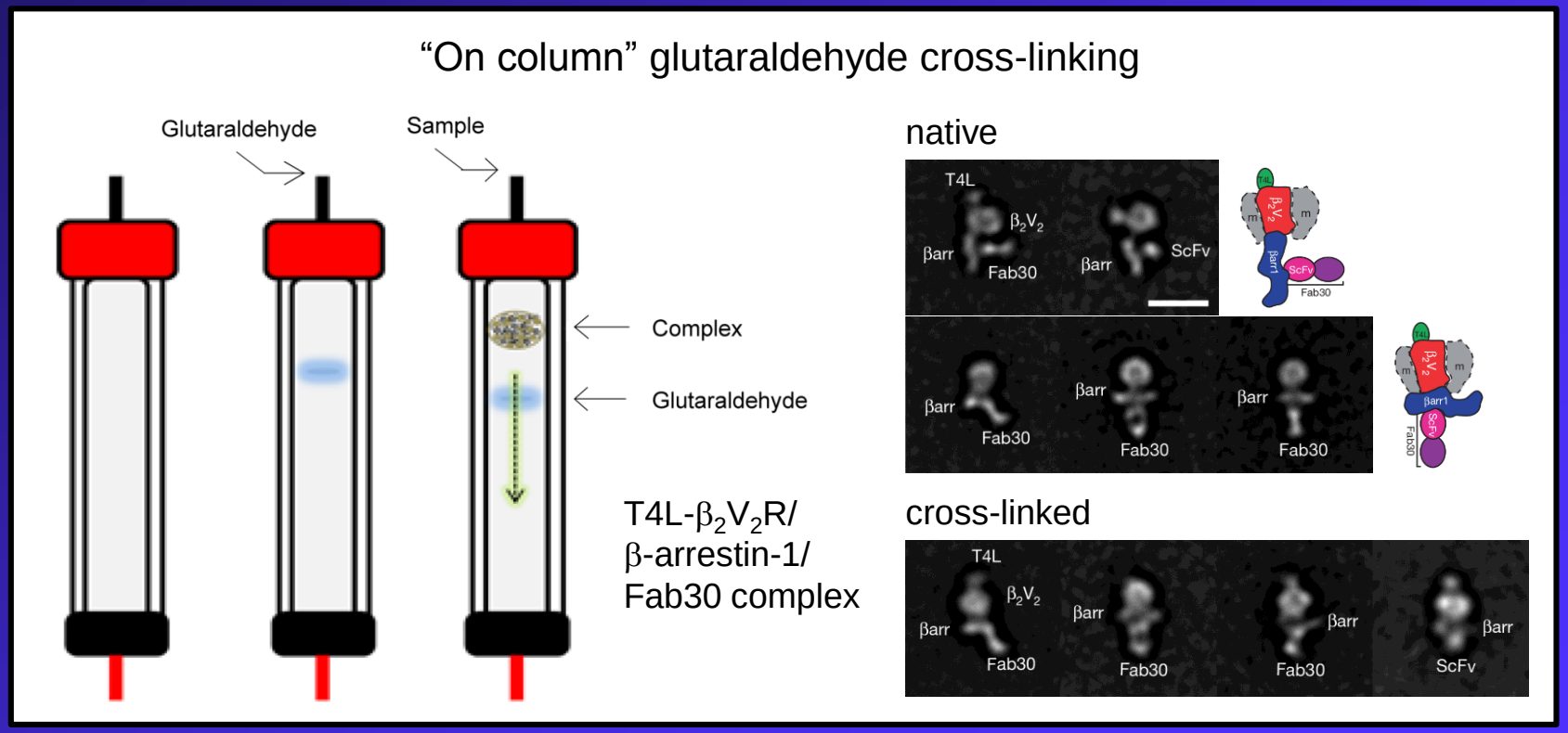
Anaphase promoting complex

Kastner *et al.* (2008) *Nat. Methods* 5: 53-55

Protein homogeneity

Optimize compositional homogeneity – chemical fixation

Cross-linking with low concentrations of glutaraldehyde
→ optimize cross-linking conditions (assess by SDS-PAGE and negative-stain EM)



Shukla *et al.* (2014) *Nature* 512: 218-222

Protein homogeneity

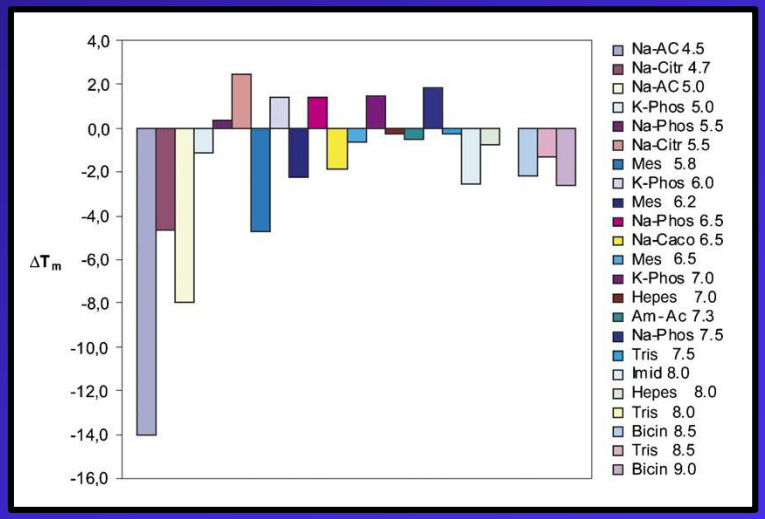
Optimize compositional homogeneity – sample buffer

– Buffer optimization based on functional assays is tedious (and not always possible)

– Biophysical properties (e.g., homogeneity, solubility, stability) predictive of successful protein crystallization

→ Fluorescence-based thermal stability assay as a high-throughput screen for buffer optimization and ligand-induced stabilization of proteins

– Thermofluor-based high-throughput stability optimization of proteins
Ericsson *et al.* (2006) *Anal. Biochem.* 357: 289-298



Changes in unfolding transition temperature (ΔT_m) for 17 proteins in 23 buffers

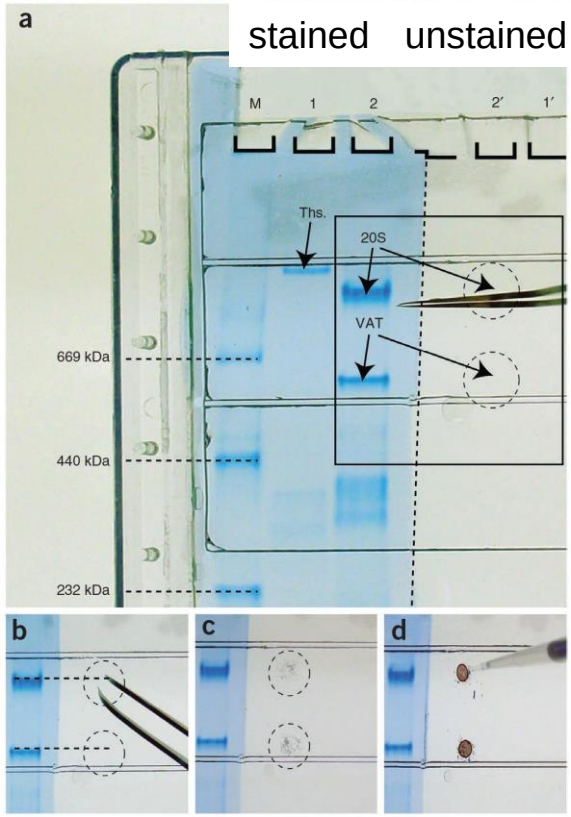
→ Studies by Holger Stark for protein complexes

Any news ???

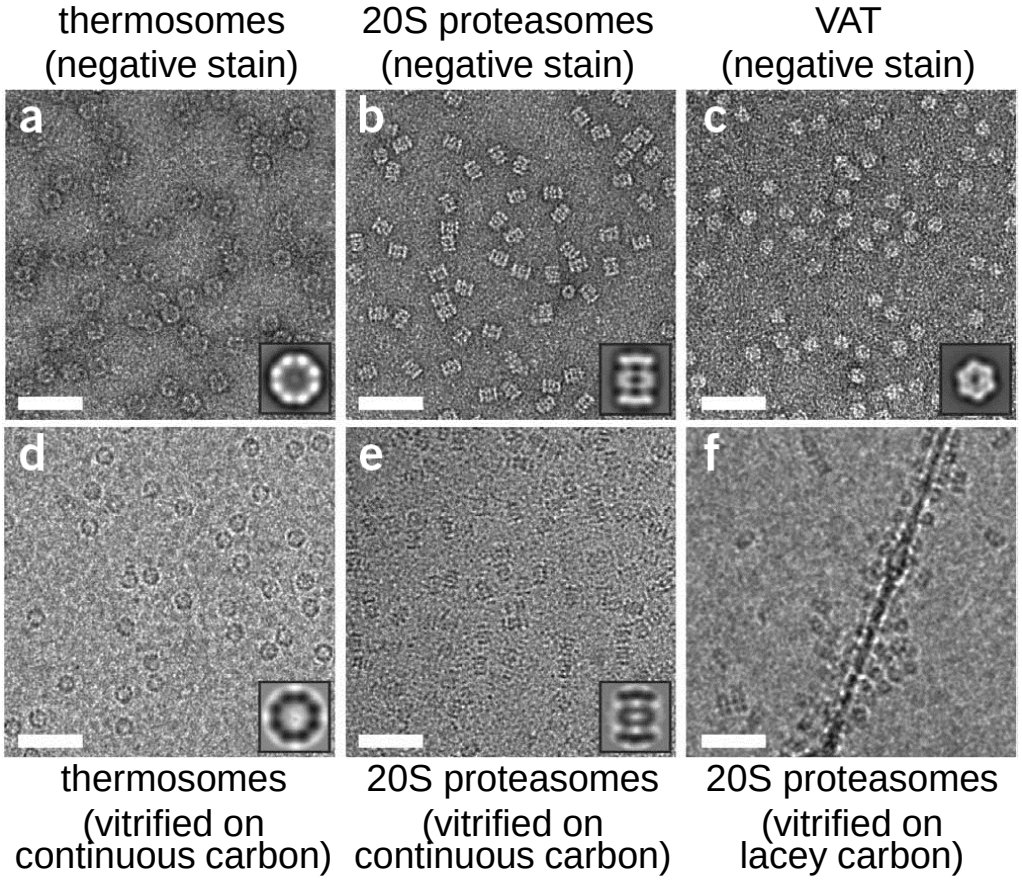
Protein homogeneity

Optimize compositional homogeneity – blotting from a native gel

Native gel



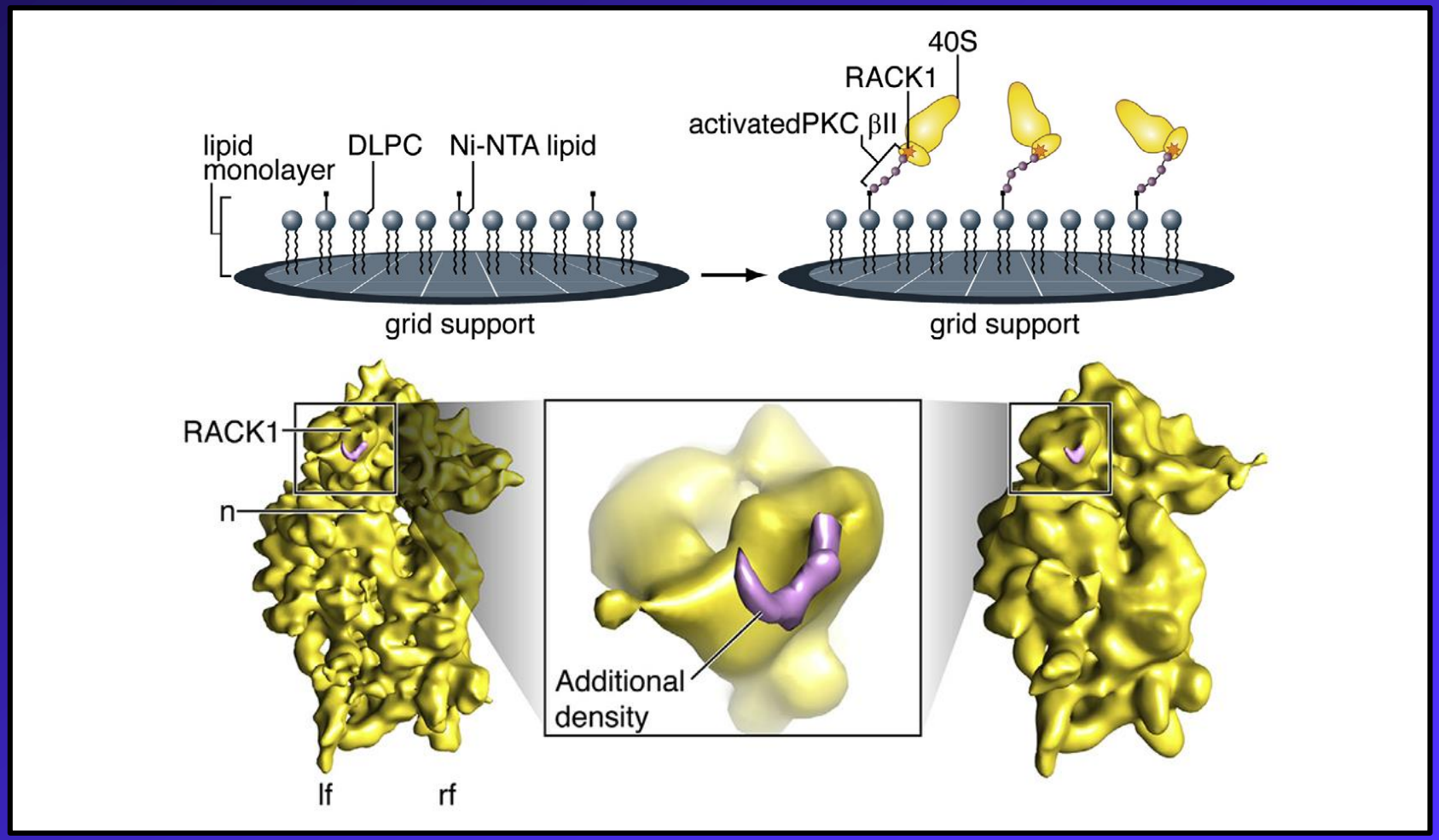
EM images



Knispel *et al.* (2012) *Nat. Methods* 9: 182-184

Protein homogeneity

Optimize compositional homogeneity – Affinity Grid



Sharma *et al.* (2013) *J. Struct. Biol.* 181: 190-194

Protein homogeneity

Optimize conformational homogeneity

– negative-stain EM may be the best (maybe only) way to "easily" assess conformational heterogeneity
  – high contrast → possible to identify heterogeneity
  – preferred orientations (calculate averages if necessary)

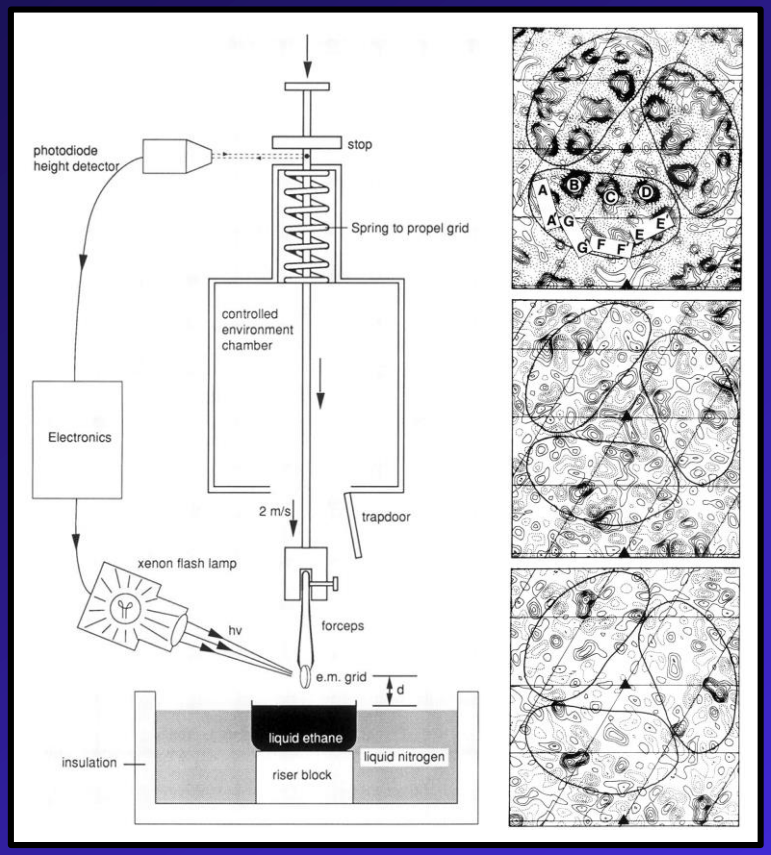
– often difficult to manage
  – add substrates, co-factors, ligands, regulators etc.
  – modify buffer (pH, ions, etc.) or try cross-linking
  – for membrane proteins: try amphipols instead of detergents

– heterogeneity has become much more manageable to deal with due to improved image quality (DDD cameras) and new software tools

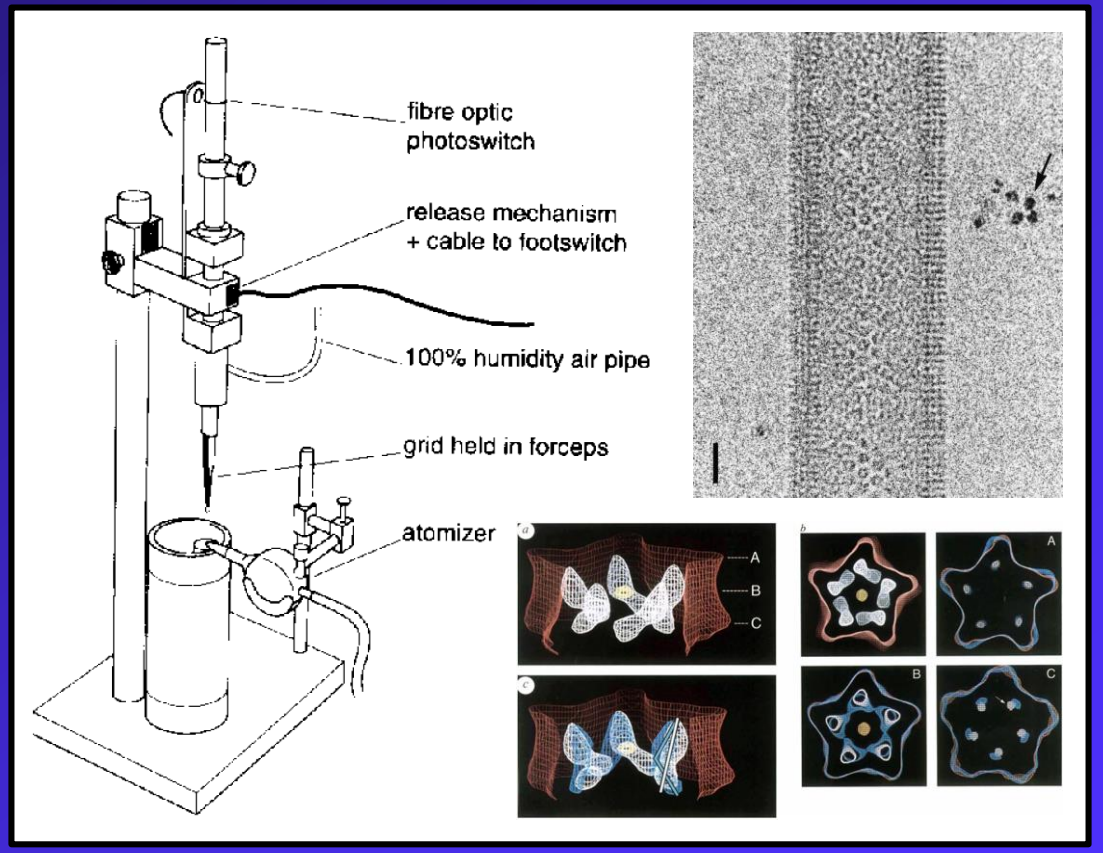
Conformational variability usually limits the resolution of the 3D maps and usually prevents building of atomic models

DIFFERENT CONFORMATIONS ARE ACTUALLY INTERESTING !

Time-resolved EM



Subramaniam *et al.* (1993)
EMBO J. 12: 1-8



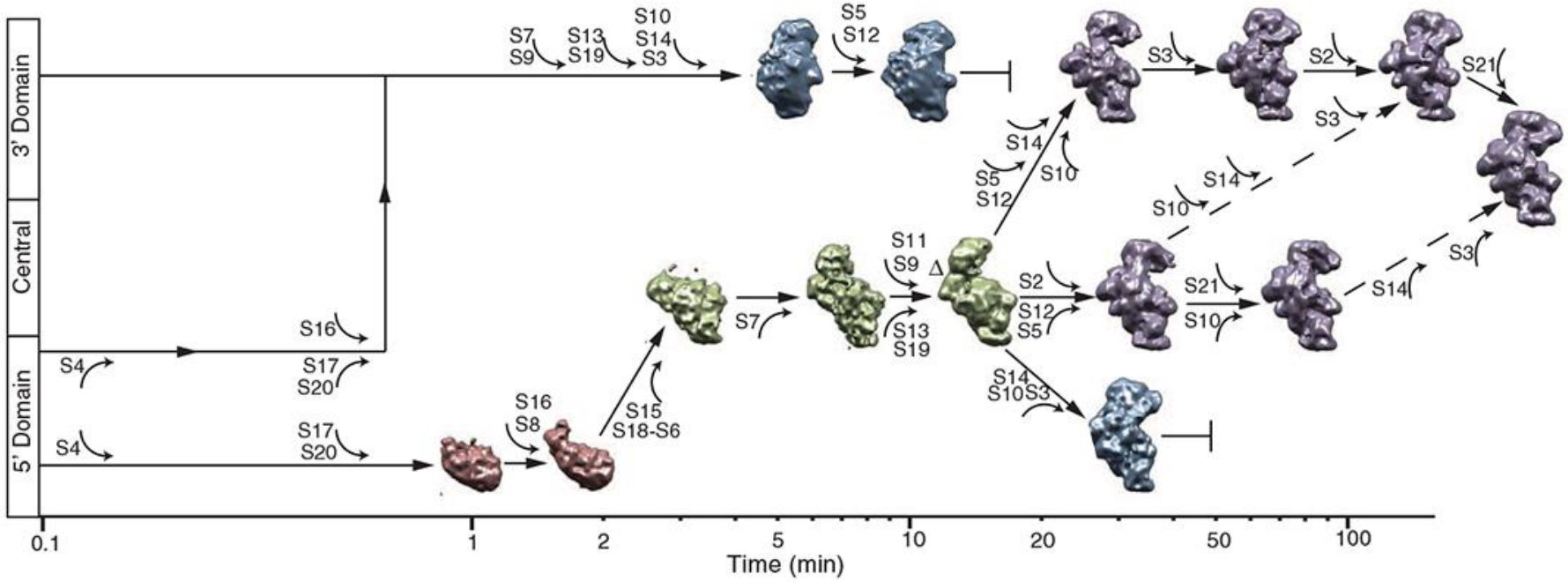
Berriman & Unwin (1994) *Ultramicroscopy* 56: 241-252
Unwin (1995) *Nature* 373: 37-43

Joachim Frank: Rapid mixing apparatus
Any news ???

Time-resolved EM

Alternative approach 1:
large data sets and computational sorting

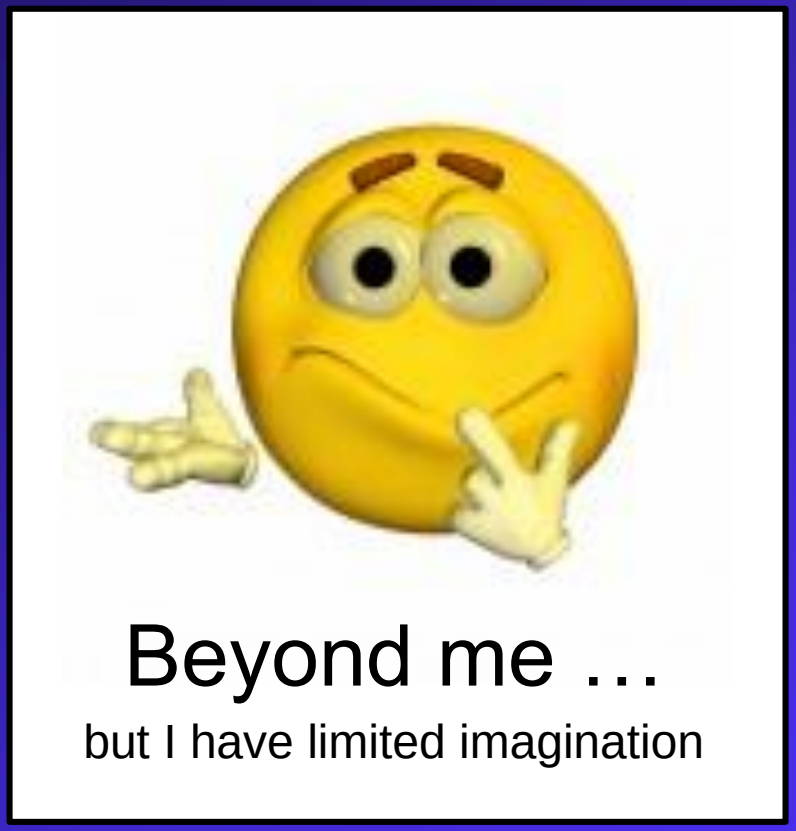
Visualizing ribosome biogenesis: parallel assembly pathways for the 30S subunit



Mulder *et al.* (2010) *Science* 330: 673-677

Time-resolved EM

Alternative approach 2:
in situ TEM using liquid specimen holders

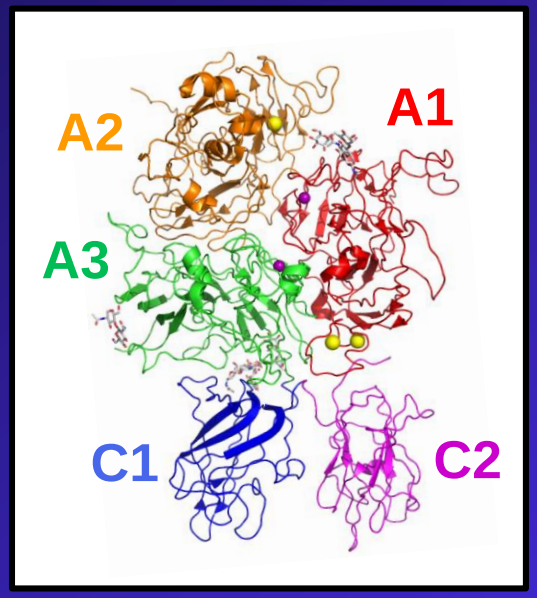


Shall we discuss ?

Specimen preparation, a (painful) example

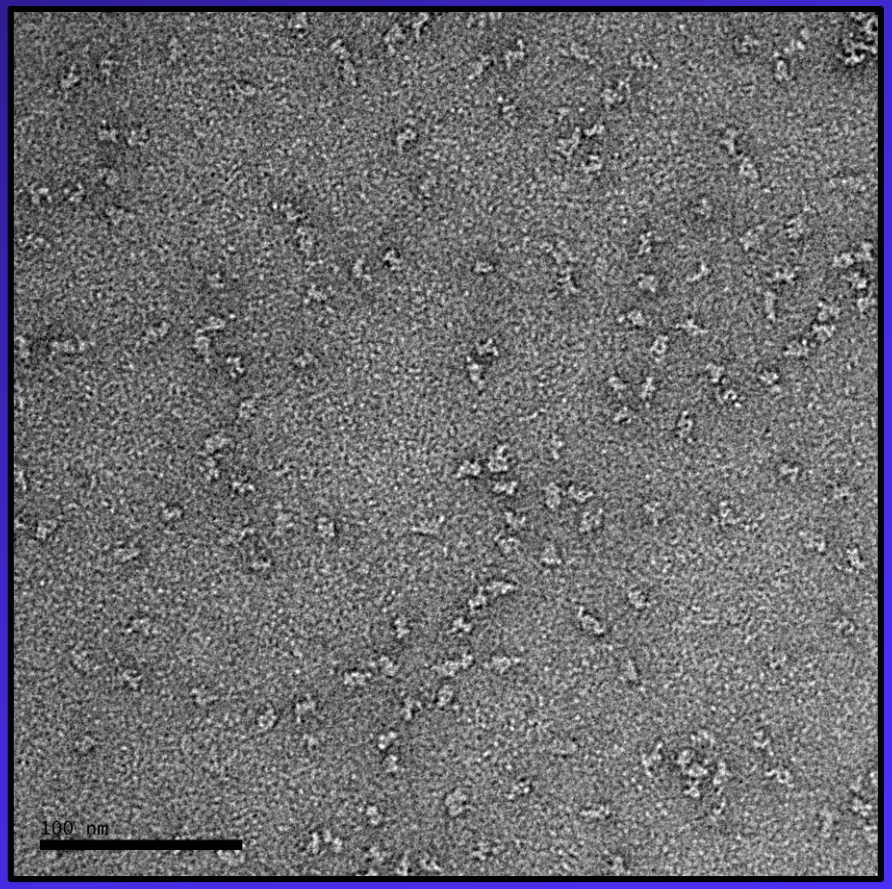
First negative-stain EM analysis

FVIII
~ 160 kDa



VWF D'D3 fragment
~ 50 kDa, dimeric

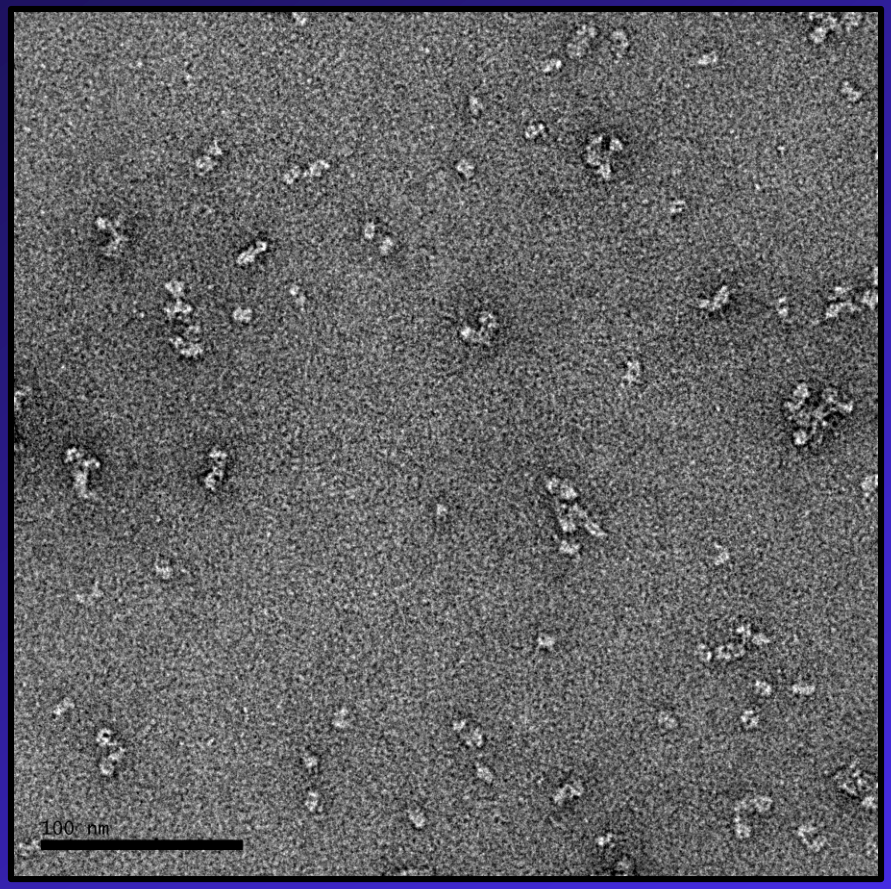
Complex
~ 420 kDa



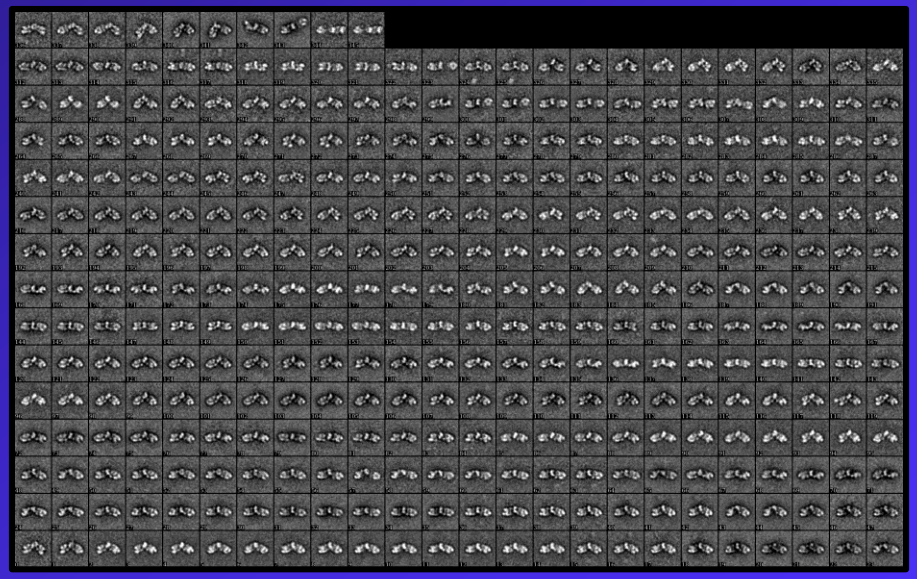
→ dimer did not form
→ bad FVIII preparation

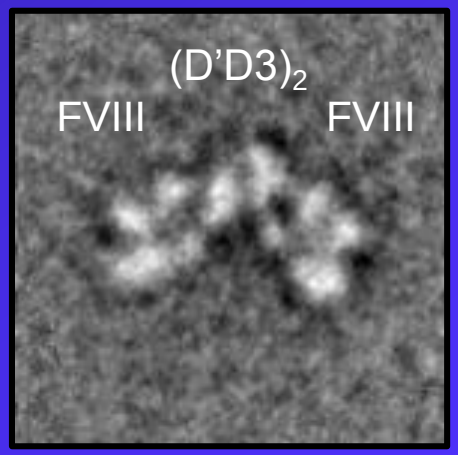
Specimen preparation, a (painful) example

Second negative-stain EM analysis



→ dimers formed





ISAC averages

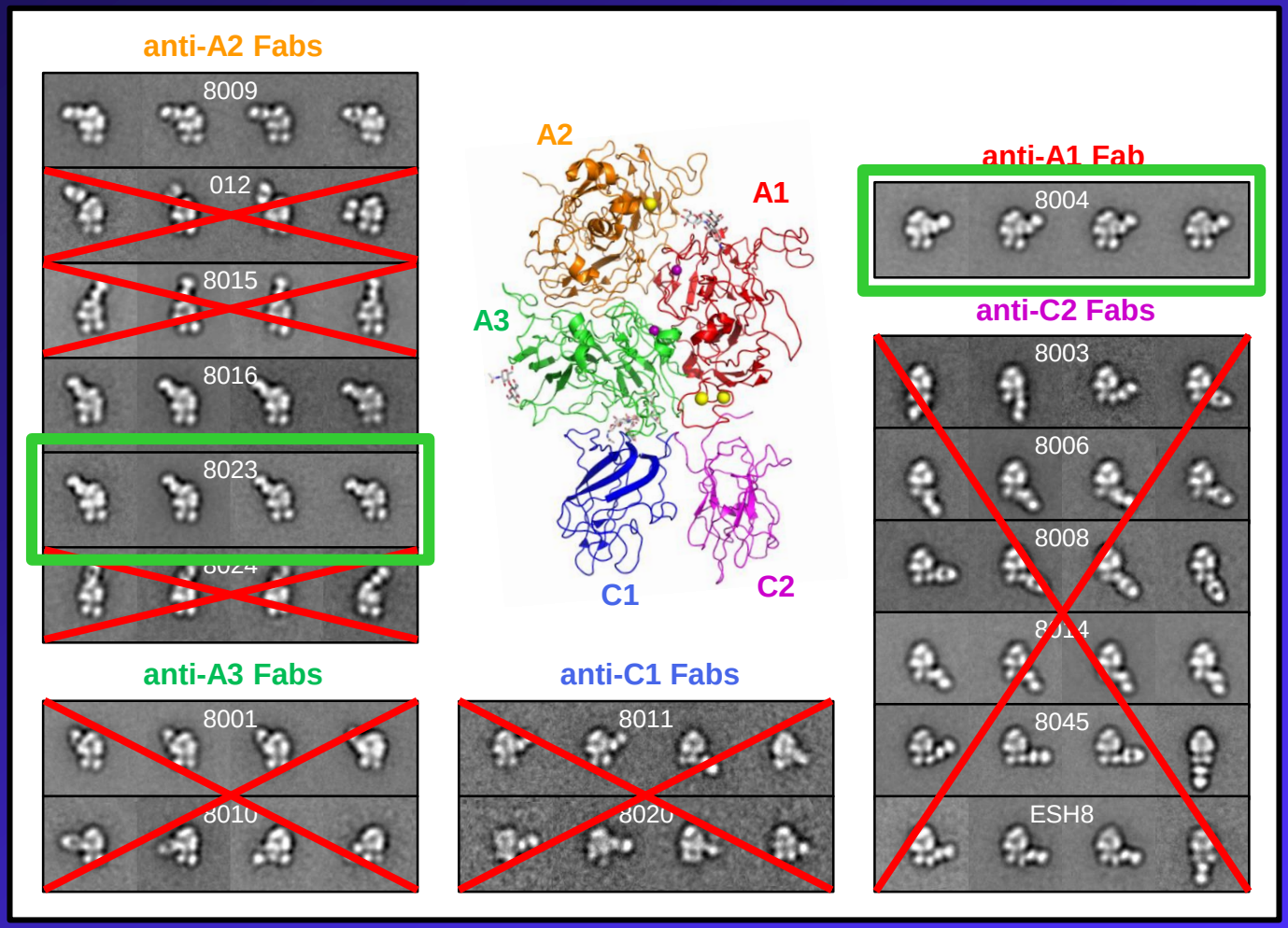
→ dimer too flexible

→ have to work with monomeric D'D3

→ too small for cryo (~ 210 kDa)

Specimen preparation, a (painful) example

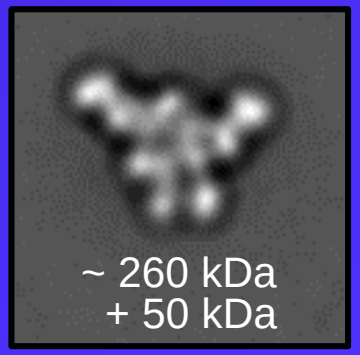
Fab analysis



Fabs appear to bind to the wrong domain

Fabs interfere with blood clotting (Bethesda assay)

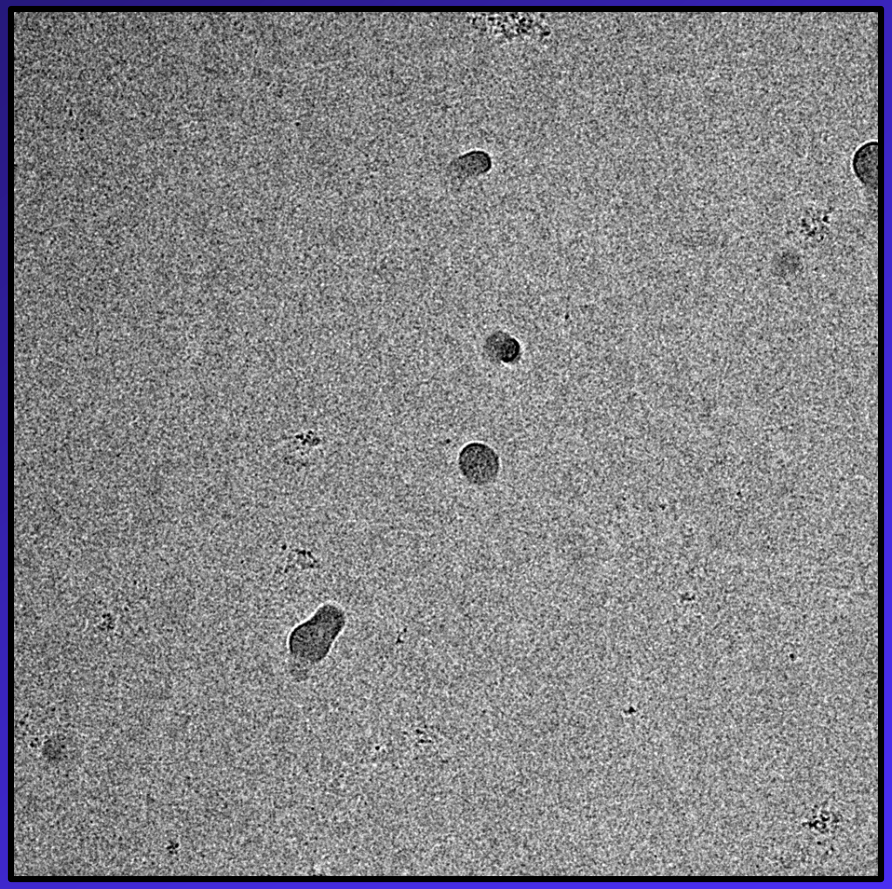
Fabs introduce structural heterogeneity



All tested Fabs have low off rates

Specimen preparation, a (painful) example

First cryo-EM attempt of FVIII-D'D3-Fabs complex



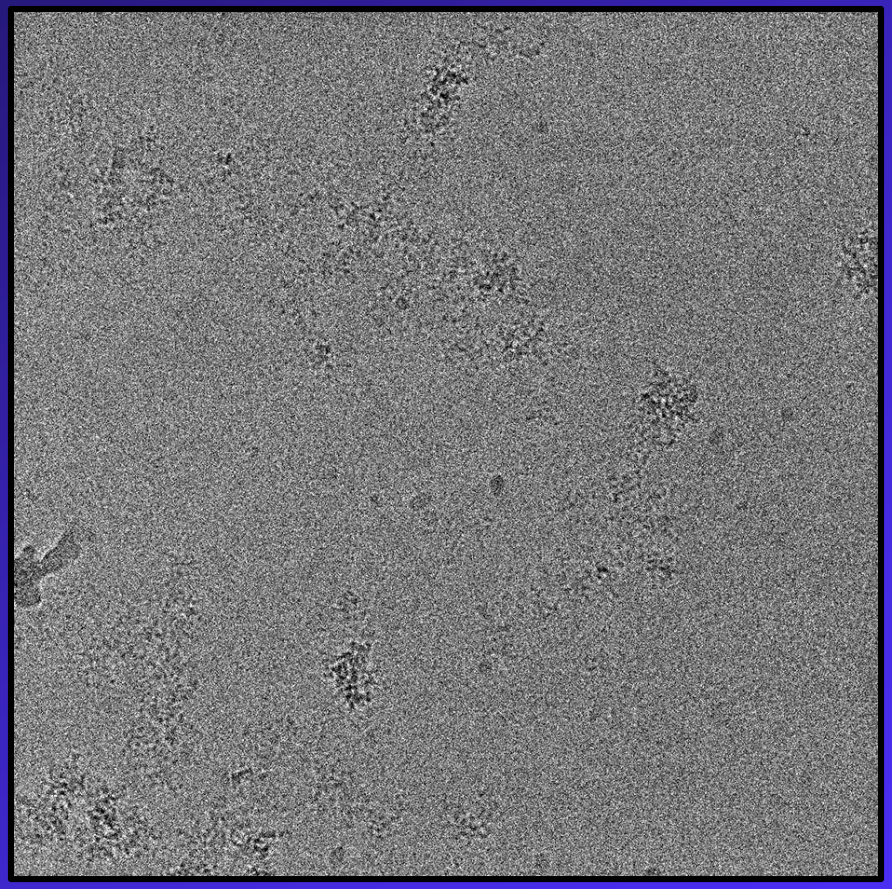
→ too dilute

→ complex cannot be concentrated

→ individual subunits have to be concentrated before complex formation

Specimen preparation, a (painful) example

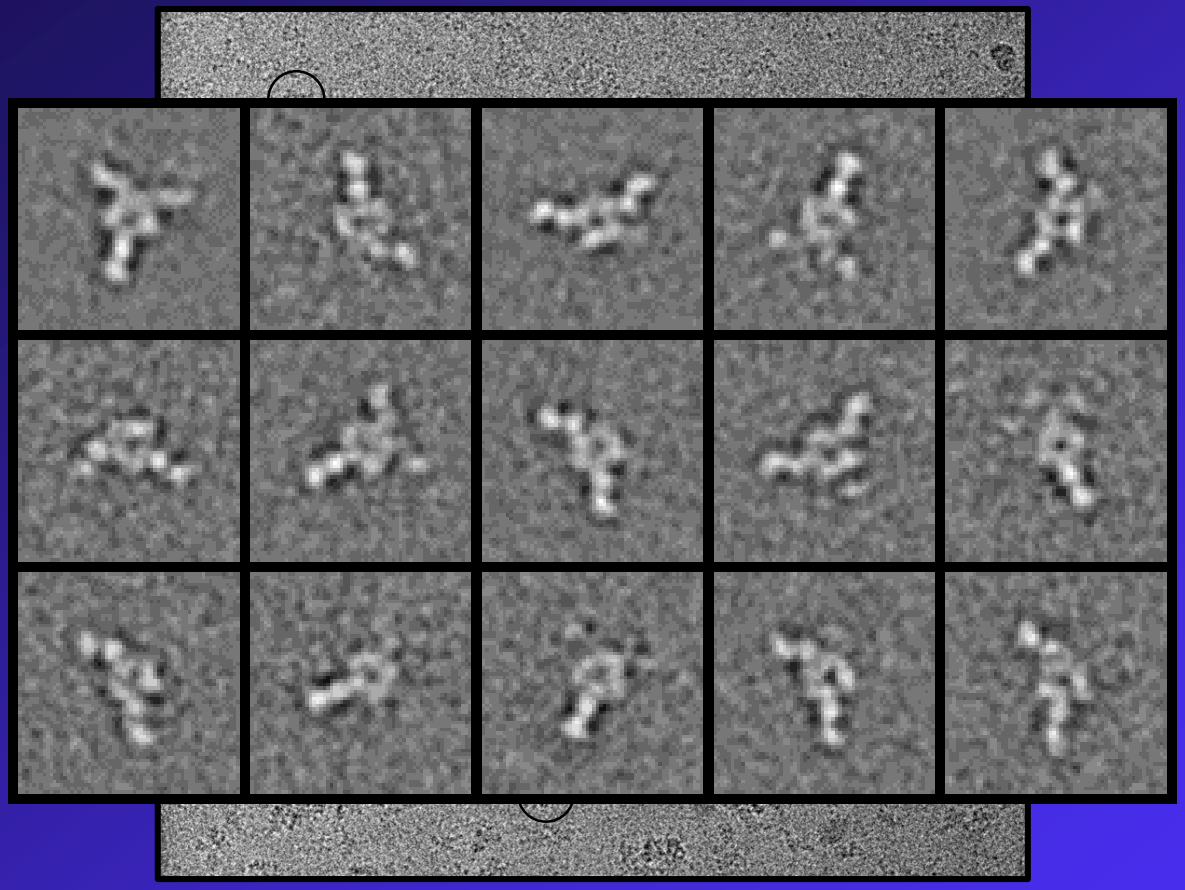
Second cryo-EM attempt of FVIII-D'D3-Fabs complex



→ complexes aggregate
→ optimize buffer
→ too much salt and detergent dissociate complex

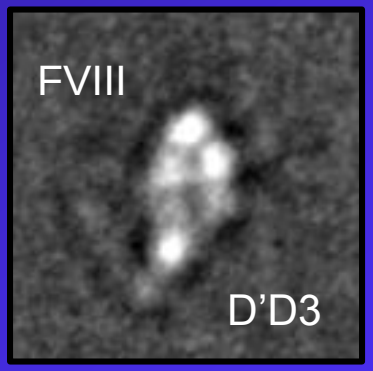
Specimen preparation, a (painful) example

Third cryo-EM attempt of FVIII-D'D3-Fabs complex



→ complexes somewhat separated
→ requires many images to be collected
→ averages look somewhat promising

BUT:



→ try to sort out heterogeneity computationally

→ try to minimize heterogeneity by cross-linking

Questions

Most specimens are still not ready for atomic resolution.

I could not agree more.

What are the specific and general problems ?

Combination of size, shape, symmetry and heterogeneity of protein.

What can be done about them ?

*Optimize biochemistry, prepare optimal grids, collect perfect data, use the best data processing strategy, and then **hope** for high resolution.*

If heterogeneity is limiting the resolution of the 3D maps, it is okay –
Learn some interesting biology ! → ***May be even more interesting than high resolution …***

Which approaches have been tried in the past ?
How successful have they been ?
Which approaches look like the most promising ?

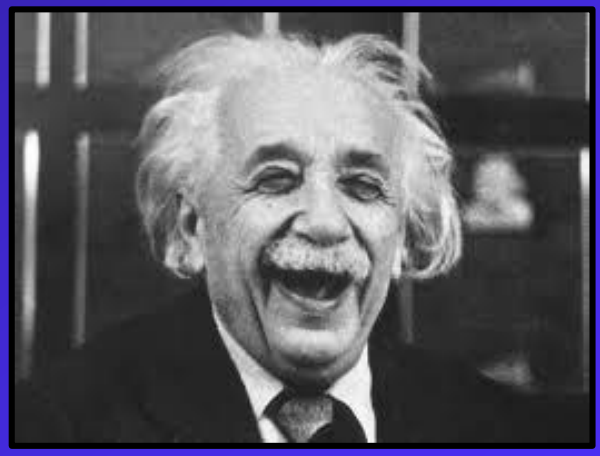
– EM data and image processing have become much better and are still getting better
– Automation & computer power allows collection & processing of increasingly larger data sets
→ resolution will become increasingly better for increasingly more difficult samples
→ sample homogeneity will become increasingly less important (but will always make it easier)
– In terms of biochemistry:
screen homologs (especially extremophiles), optimize buffer, cross-linking & Fab labeling

Questions

Given that many small and heterogeneous samples may only be suitable for examination in negative stain or at low resolution, how do we make sure that the general scientific community (and ours, too!) understands that not every-thing is getting to atomic resolution ?

Beyond me …

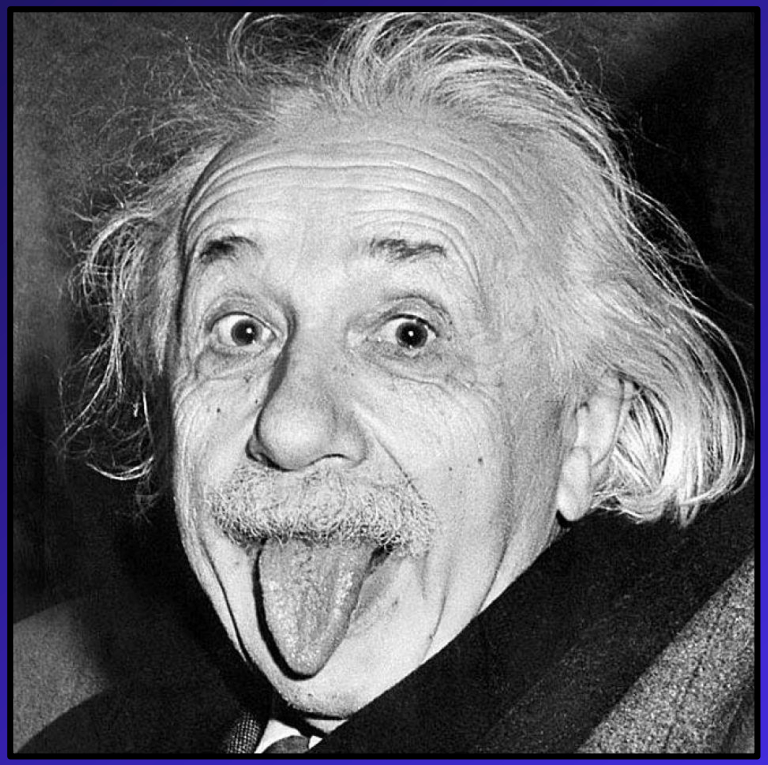
but I have limited imagination



”Two things are infinite:
the universe and human stupidity;
and I’m not sure about the universe.”
— Albert Einstein

”If you can’t explain it to a six year old,
you don’t understand it yourself.”
— Albert Einstein

That's it !



Good luck !