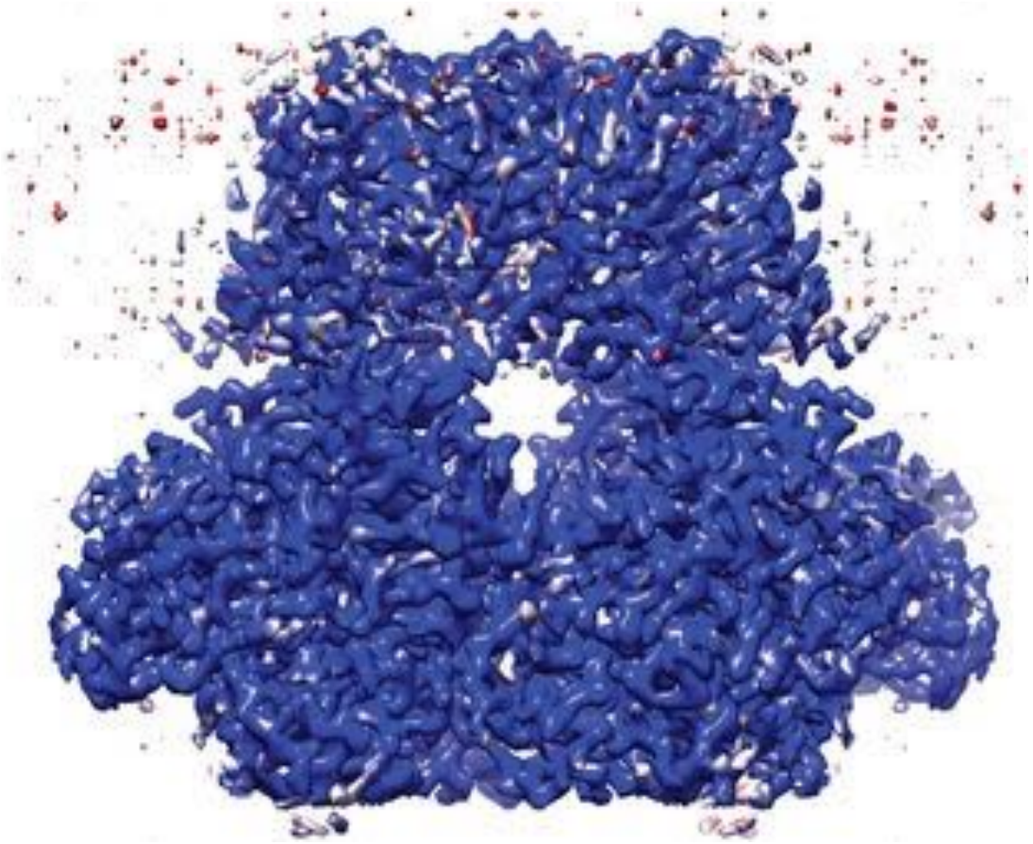
# Single-particle analysis (Part IV)
# Cryo-EM map interpretation

Rich Hite
Memorial Sloan Kettering Cancer Center
March 11, 2019

# You have a density map, now what?

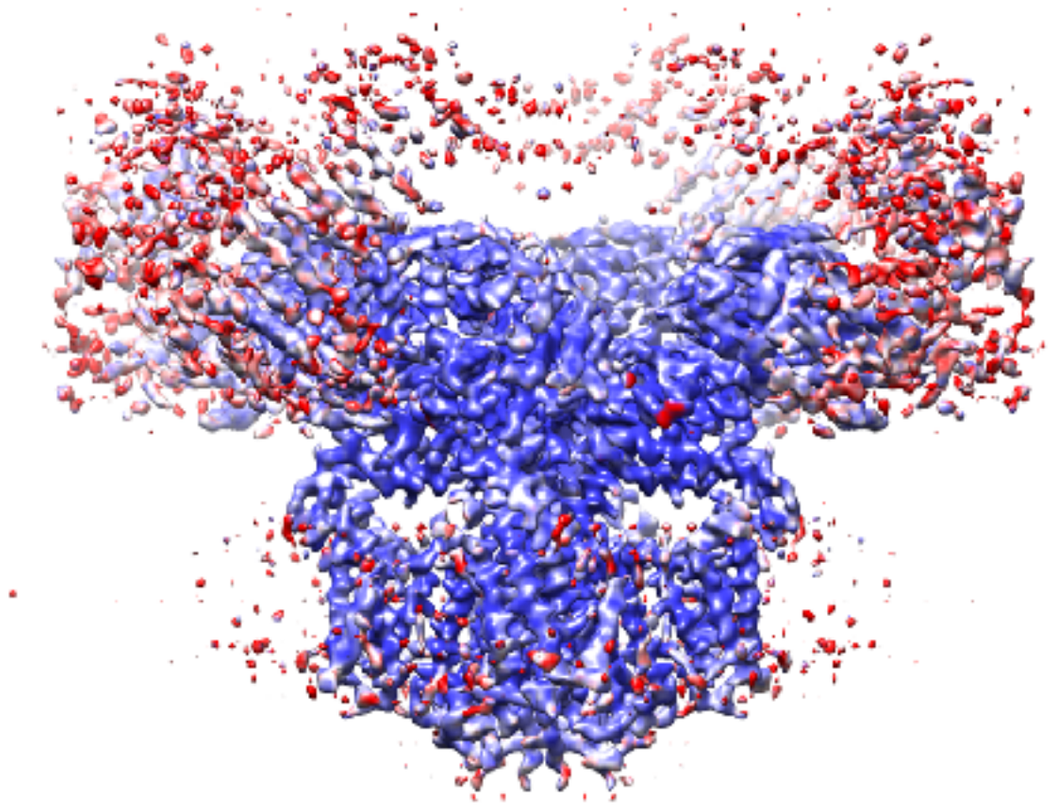
- First and most important question - What is your biological question and how can cryo-EM help you to answer it?
- Second, critically evaluate your map or maps!
  - Is the domain that you are interested in studying visible
  - If so, at what resolution?



Tao *et al*, Nature. (2017) 541:46-51

# You have a density map, now what?

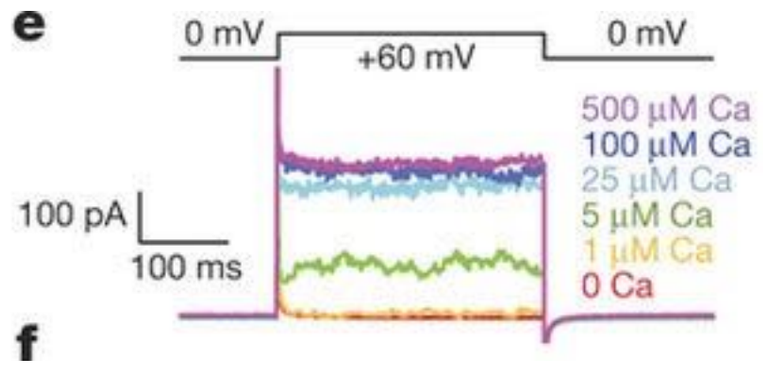
- First and most important question - What is your biological question and how can cryo-EM help you to answer it?
- It is essential to critically evaluate your map or maps to understand if they can help you to answer your biological question!
  - Is the domain that you are interested in studying visible?
  - If so, at what resolution?
  - Does the domain undergo conformational changes in different structures?

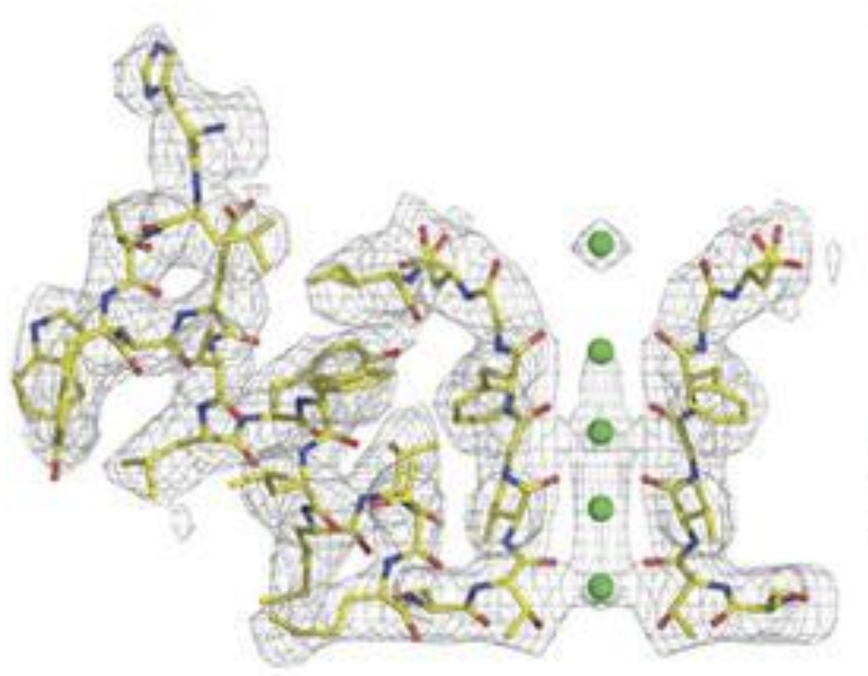


Paknejad and Hite, Nat. Struct. Mol. Biol. (2018) 25:660-668.

# You have a density map, now what?

- Second question - What additional evidence do you have to support the hypothesis that your density map faithfully represents your protein of interest?
  - Can you reconstitute your protein or complex in a functional assay?
  - Does the domain architecture agree with previous data?



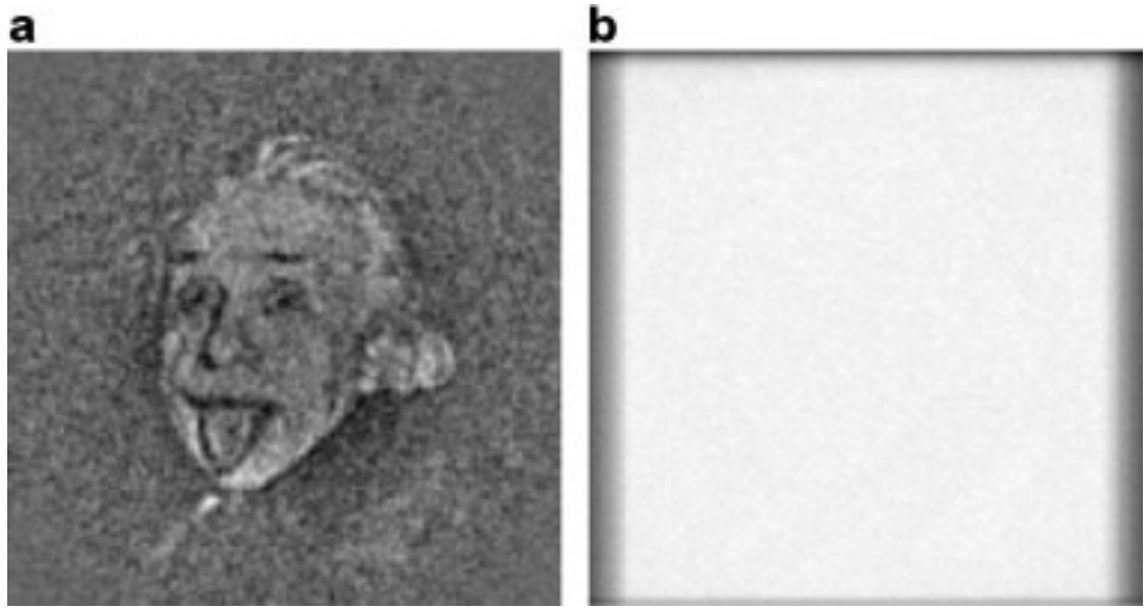



Tao *et al*, Nature. (2017) 541:46-51

# Map validation

Validate your map!

- What is the effect of initial model bias on your reconstruction?
  - Remember that you will always get back from the reconstruction algorithms the model that you provide at the start!
  - The use of masks will also lead to a bias, so make sure that you can get a similar reconstruction with and without the mask
- Beware of the Einstein-from-noise phenomenon




Shatsky *et al*, JSB. (2009) 166:67-78

# Map validation - continued

Initial models

- How reproducible are your initial models with different programs
- Can you generate an initial model without imposed symmetry?

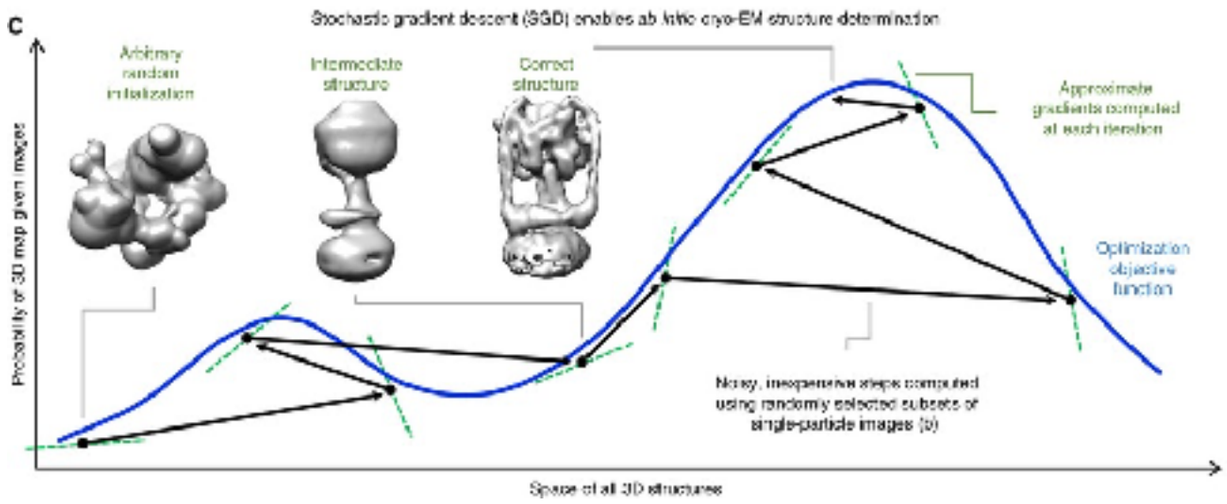
Symmetry

- How did you define the symmetry of your sample?
- Can you generate a similar reconstruction without imposing symmetry?

Internal (psuedo-)symmetry (i.e. NCS)

- Do domains that have similar architectures adopt similar structures?
- Can you average multiple non-symmetrical copies of a domain to improve the map locally?

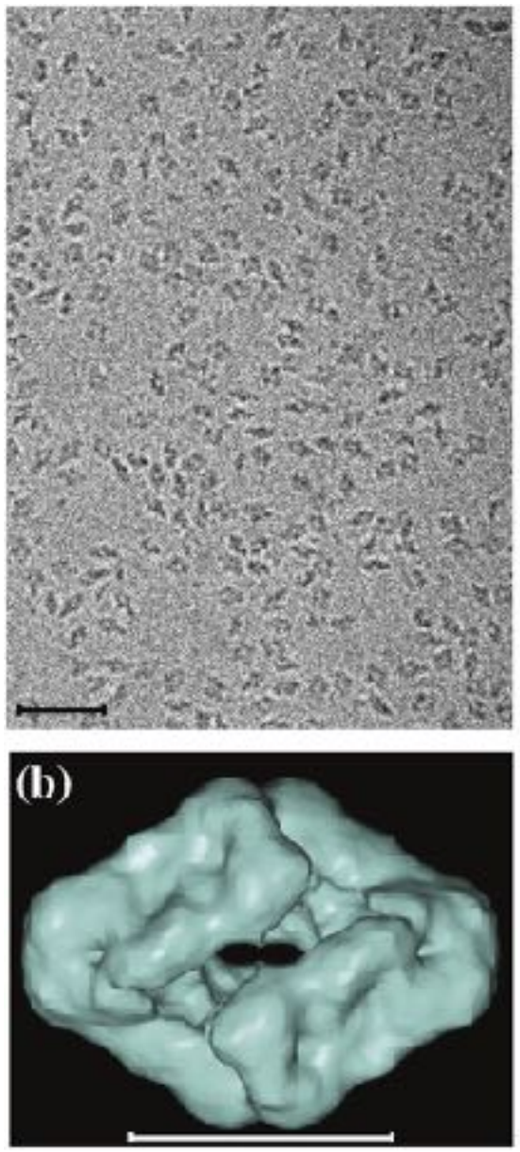
CryoSPARC

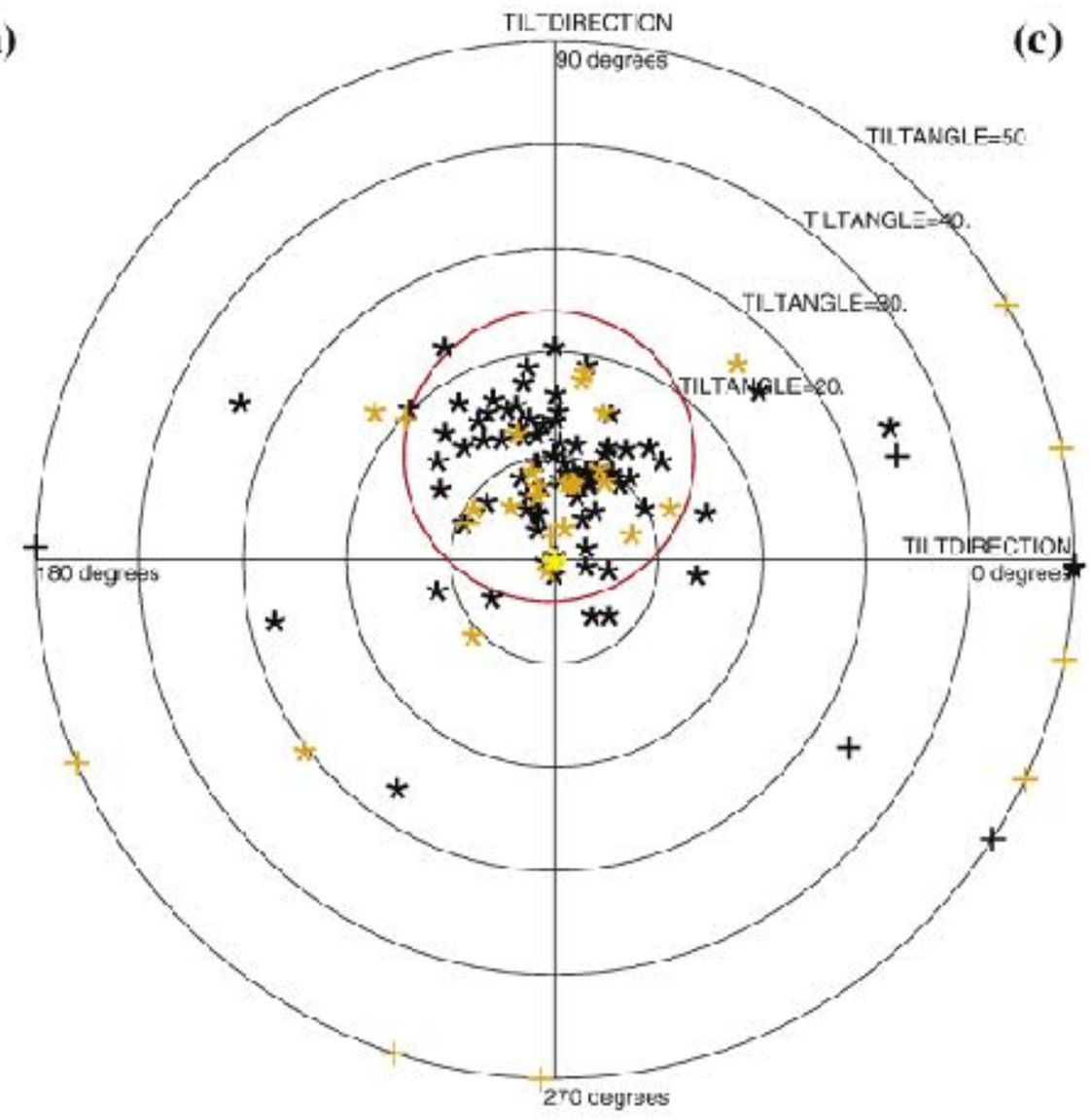



Punjani *et al*, Nat Methods. (2017) 14:290-296

# Tilt-pair validation

- Collect images of the same particles with and without tilting the stage
- Calculate orientation parameters by refinement
- Measure the difference between the tilted and untitled images to determine the accuracy of particle alignment

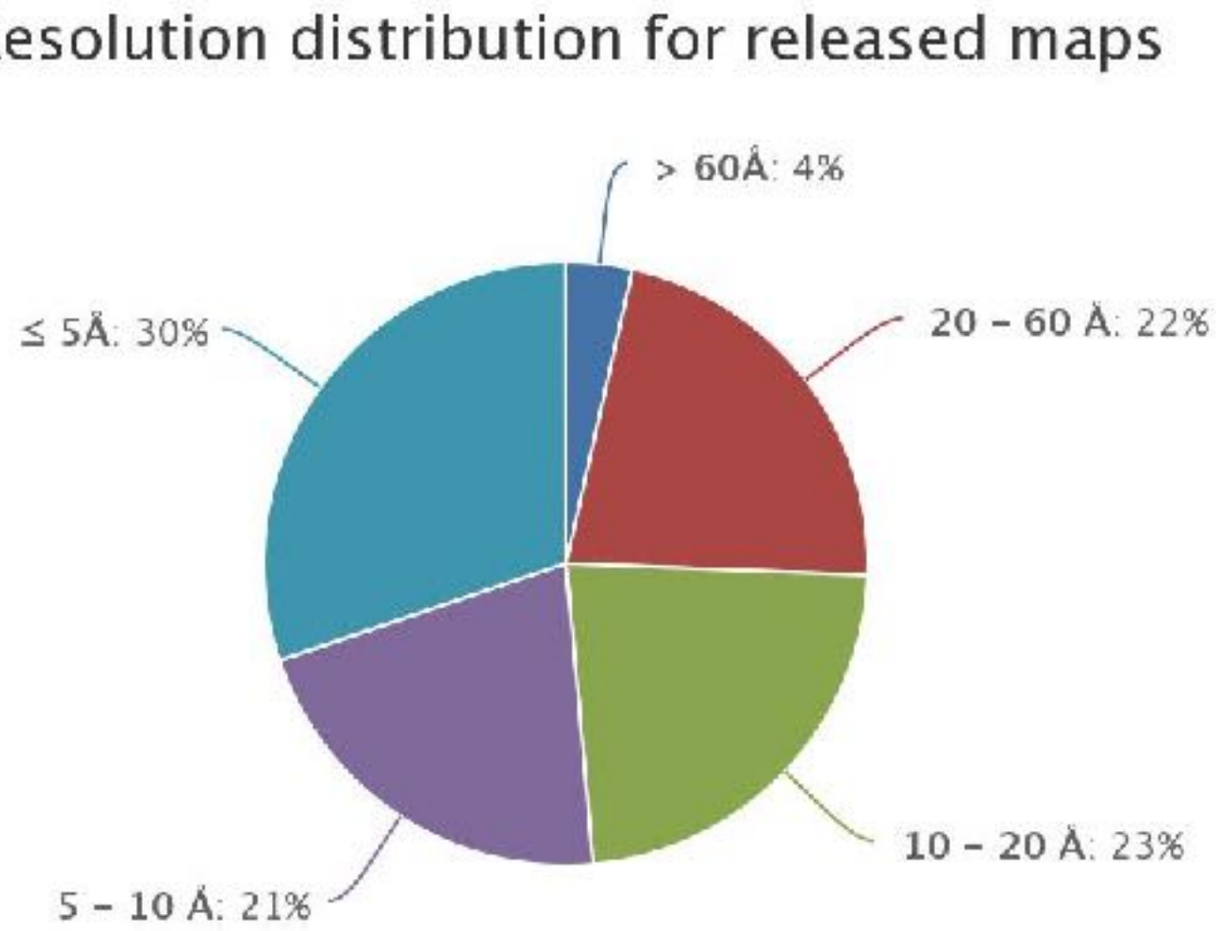



Henderson *et al*, JMB. (2011) 413:1028-104

# Resolution - what does it mean?

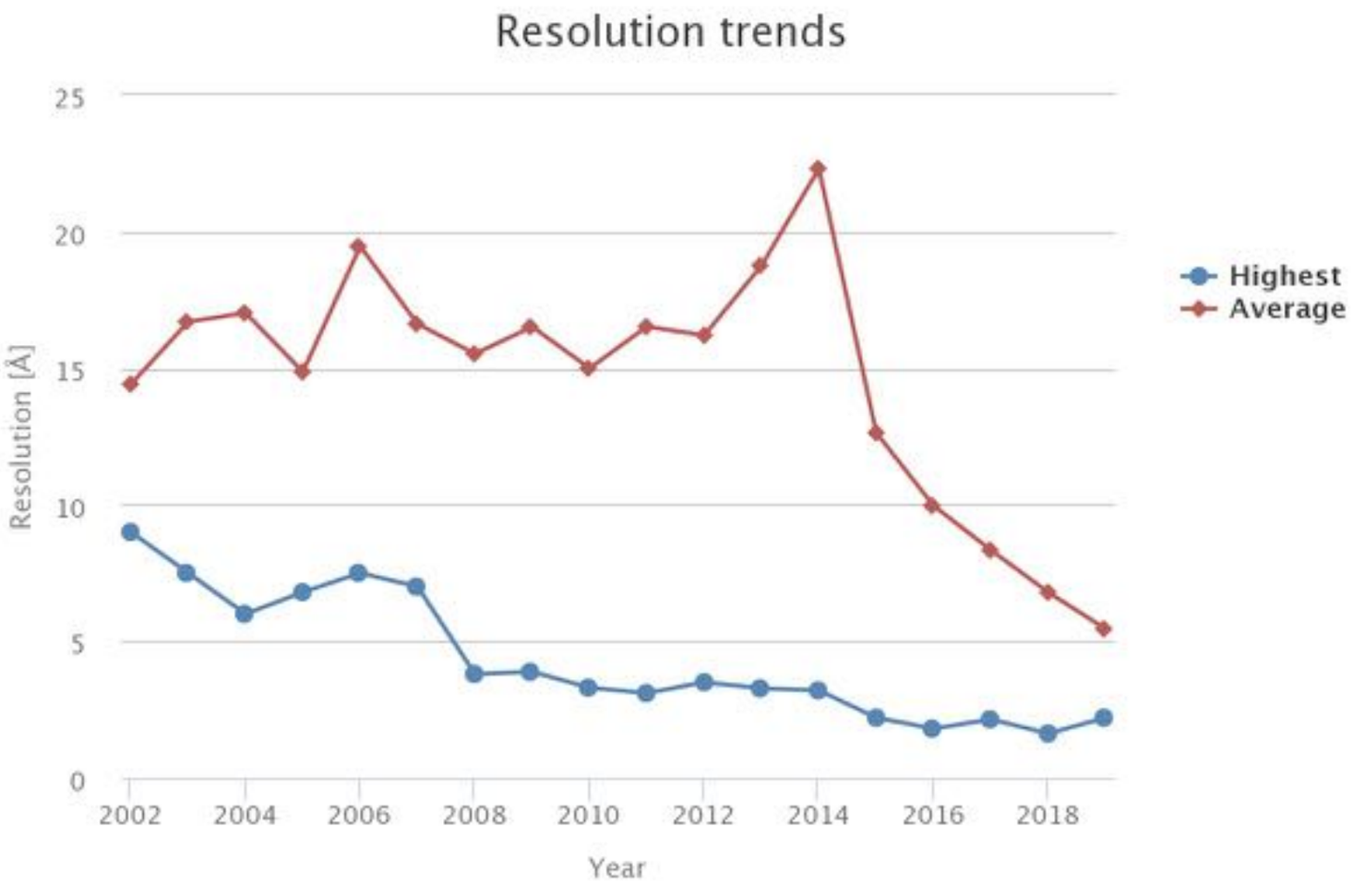
- What is the meaning of 4 Å map? A 3 Å map?
- How does the resolution influence our interpretation?
- What do we gain from higher resolution reconstructions?
- How can we improve the resolution of reconstruction?




EMDB 2019.03.08

# Resolution - what does it mean?

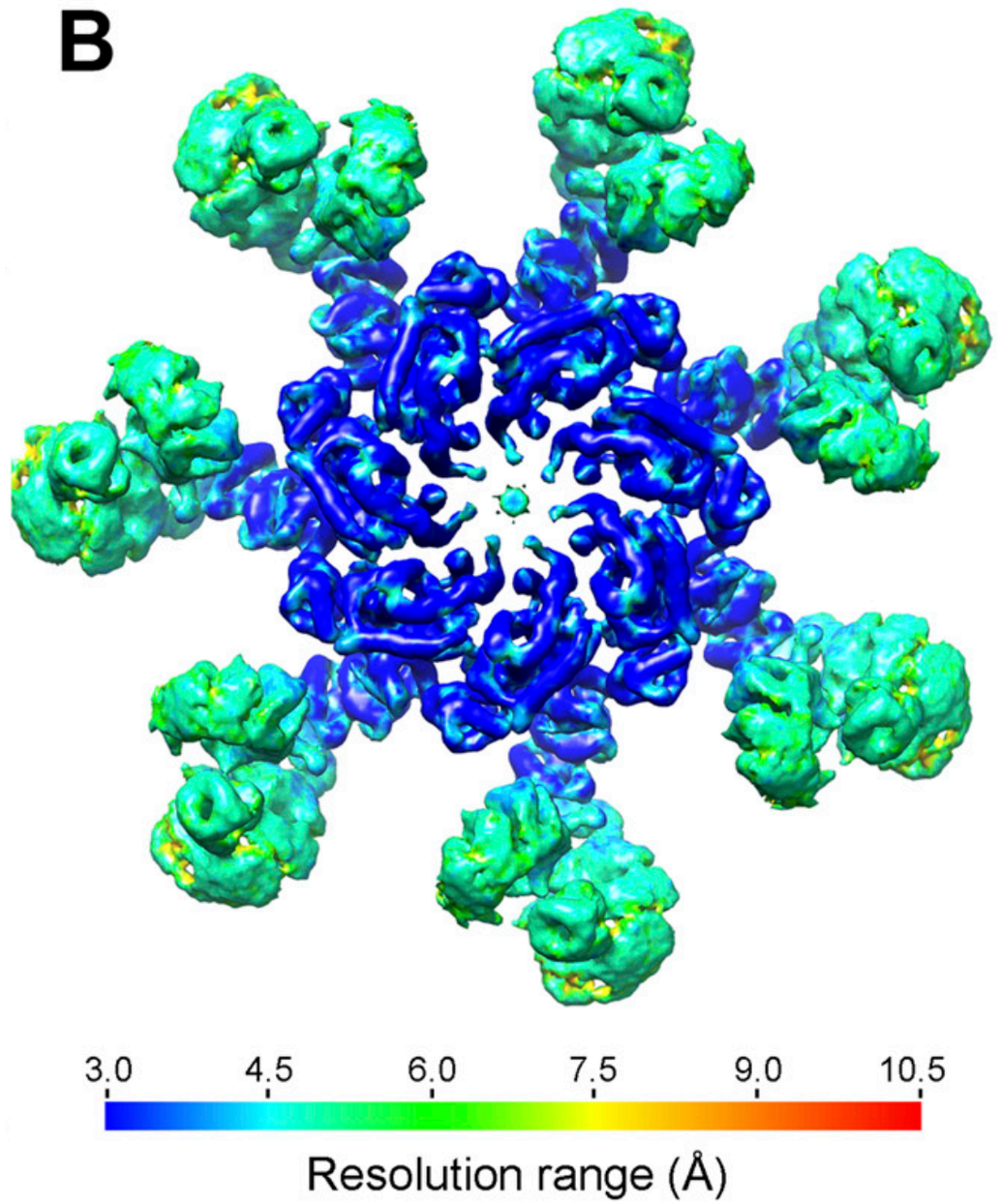
- What is the meaning of 4 Å map? A 3 Å map?
- How does the resolution influence our interpretation?
- What do we gain from higher resolution reconstructions?
- How can we improve the resolution of reconstruction?




EMDB 2019.03.08

# Resolution - what does it mean?

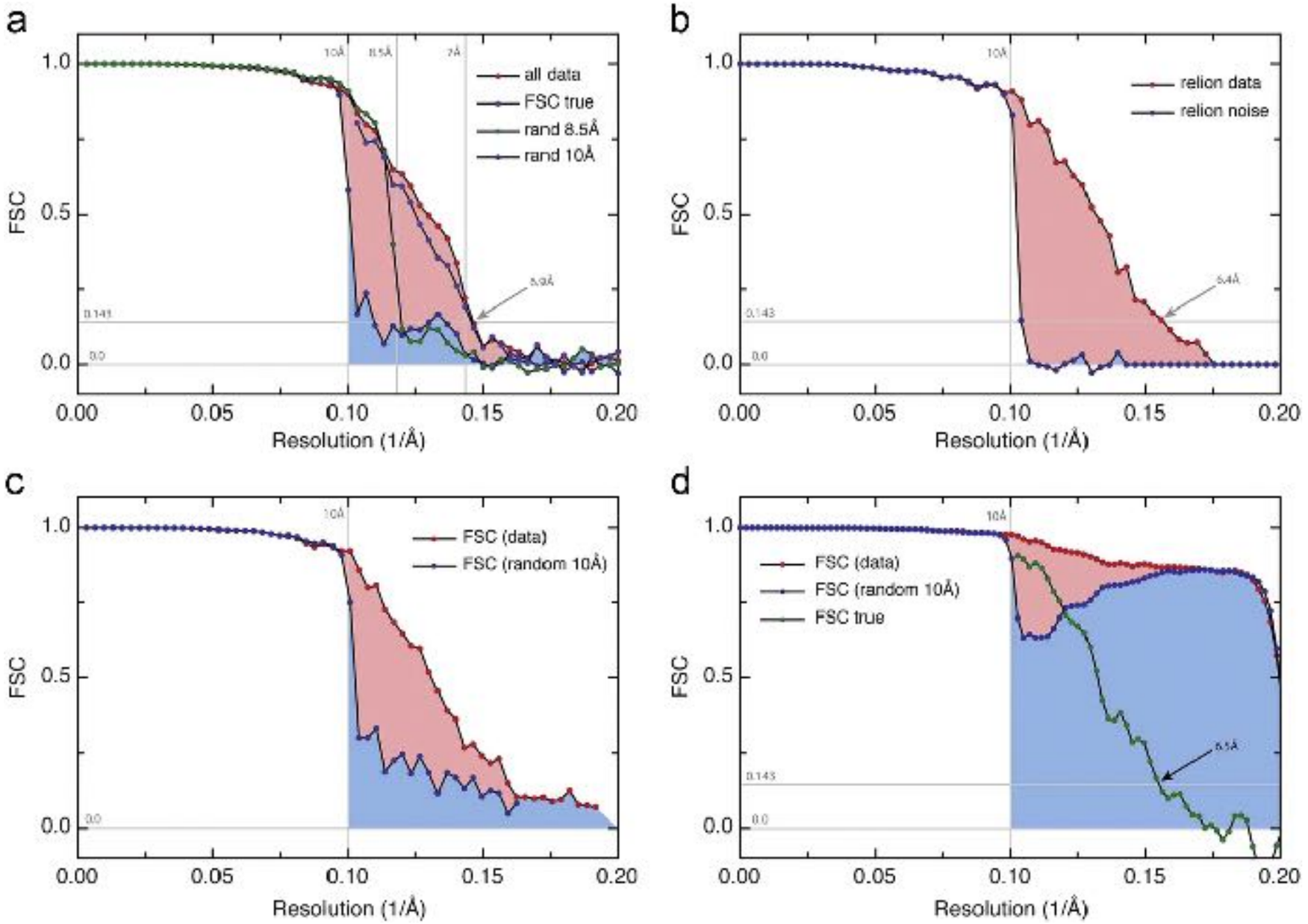
- Overall resolution does not equal local resolution
- Quality can vary greatly within a map and care should be taken to not over interpret poorly ordered domains
- Local resolution estimates can be performed
  - ResMap http://resmap.sourceforge.net
  - Blocalres (package in Bsoft) https://lsbr.niams.nih.gov/bsoft/programs/blocres.html
- Density slices can also be extremely informative for evaluating local map quality




Li *et al*, PNAS. (2017) 114:1542-1547
Kucukelbir *et al*, Nat. Methods. (2014) 11:63-65
Heymann and Belnap, JSB. (2007) 157:3-18

# High-resolution noise substitution

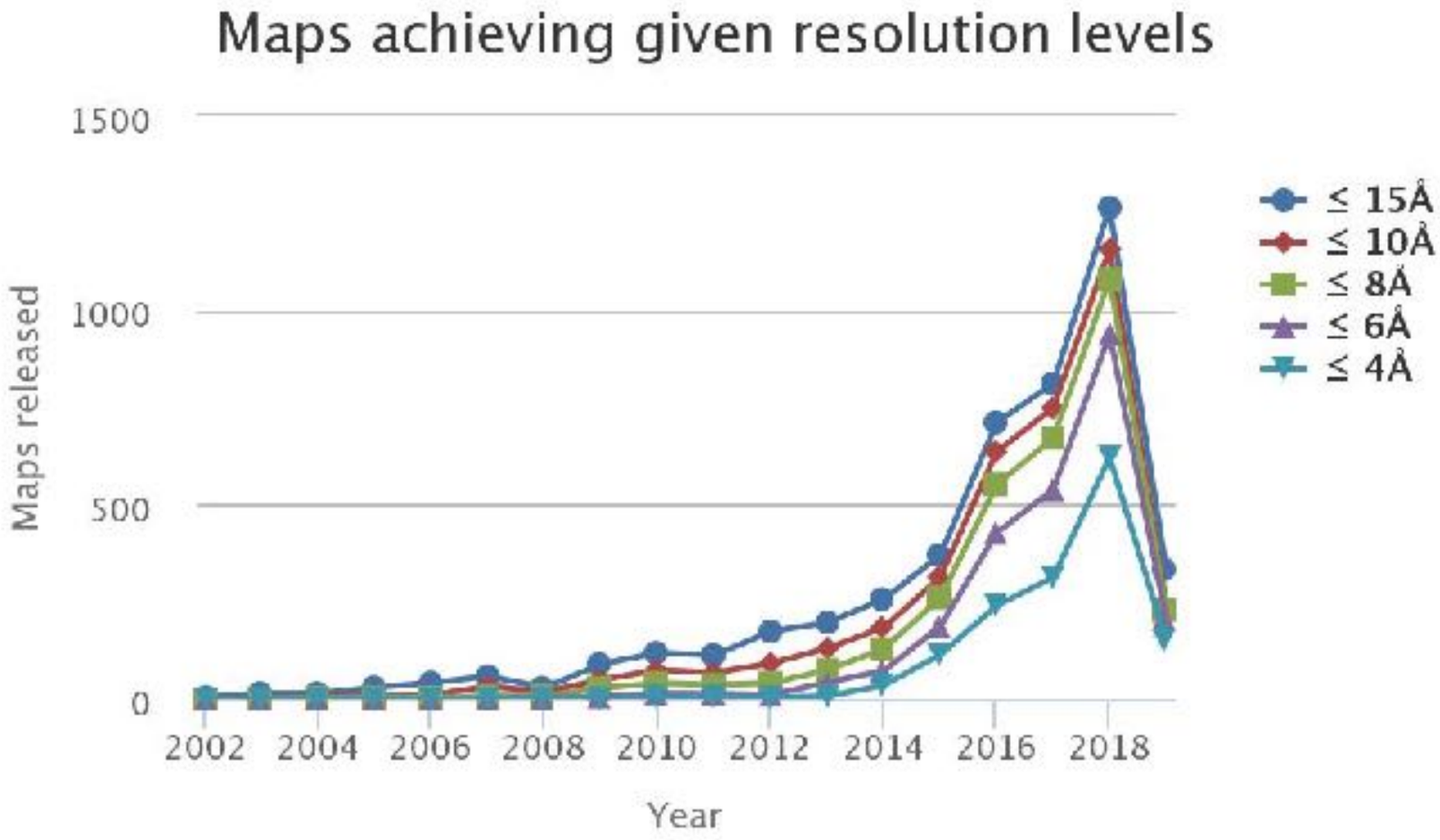
- Method to determine the effect of masking upon FSC correlations
- Replace signal beyond a particular resolution with noise
- Calculate FSC for map with signal and with noise
- $FSC_{true} = (FSC_{data} - FSC_{noise}) / (1 - FSC_{noise})$




Chen *et al*, Ultramicroscopy. (2013) 135:24-35

# Resolution - what does it mean?

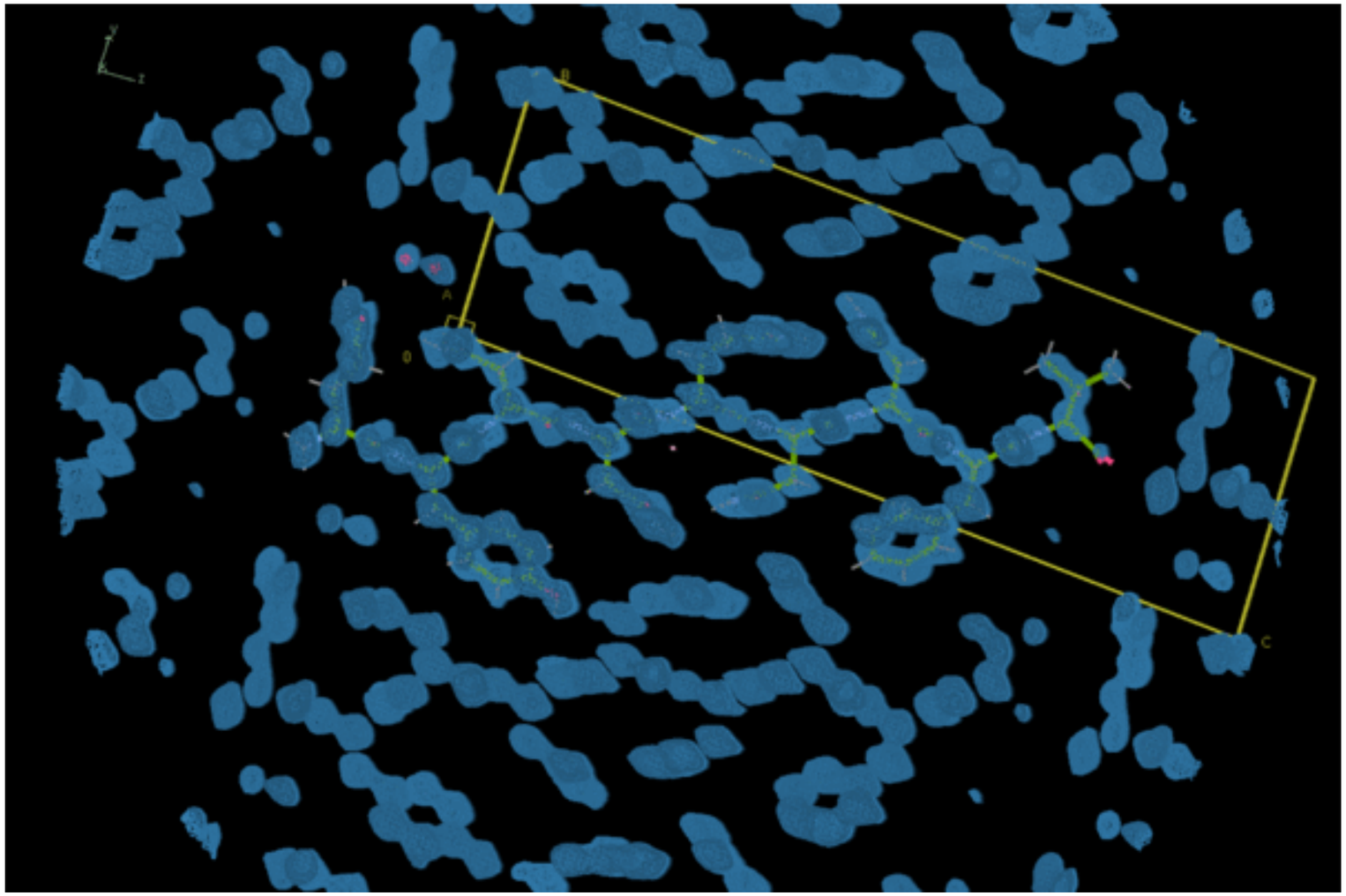
- What is the meaning of 4 Å map? A 3 Å map?
- How does the resolution influence our interpretation?
- What do we gain from higher resolution reconstructions?
- How can we improve the resolution of reconstruction?




EMDB 2019.03.08

# Resolution - what does it mean?

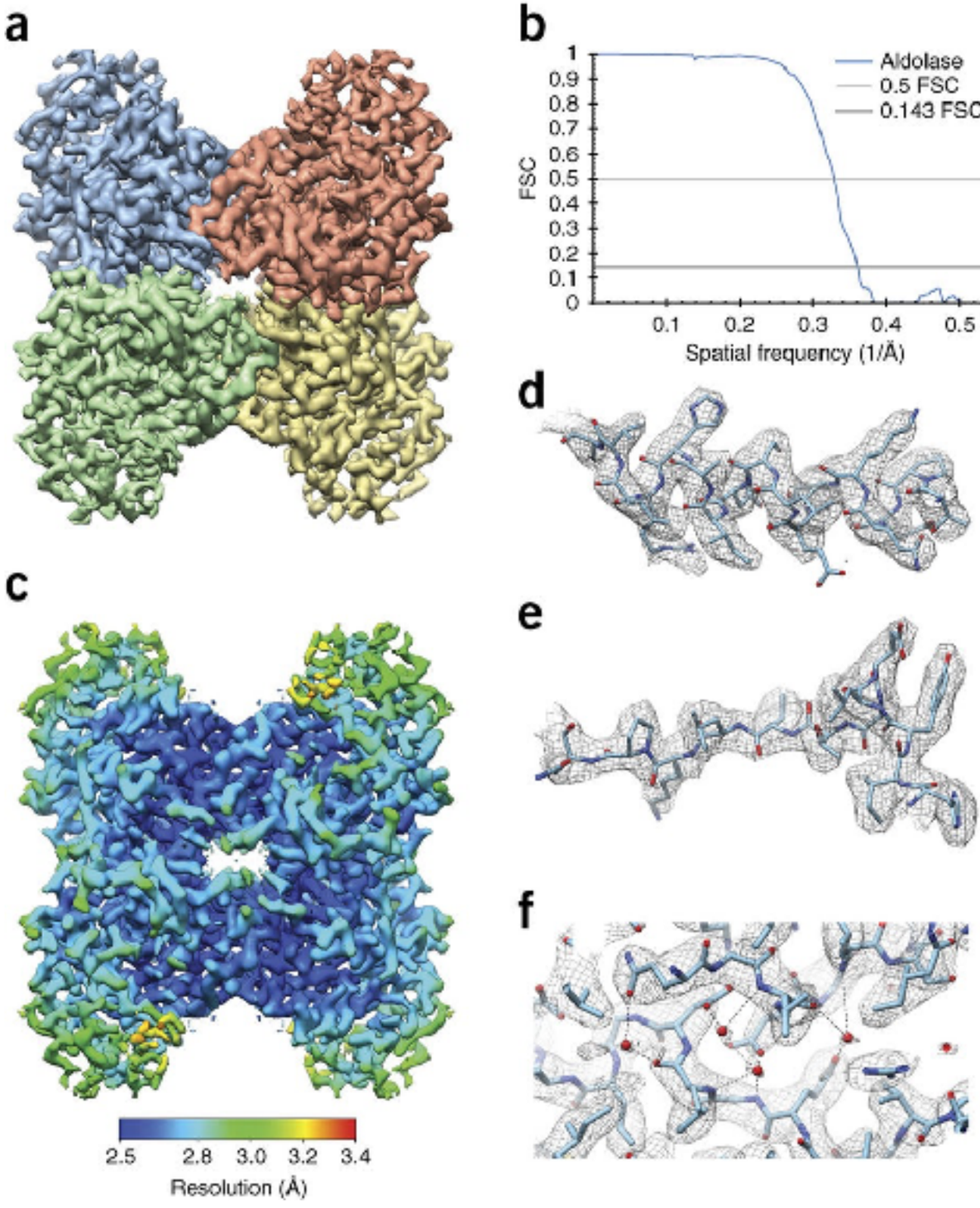
Sub-1 Å microED of bank vole prion protein 168-176 QYNNQNNFV



EMDB 7017
Gallagher-Jones M et al, NSMB **25** 131-134 (201

# Resolution - what does it mean?

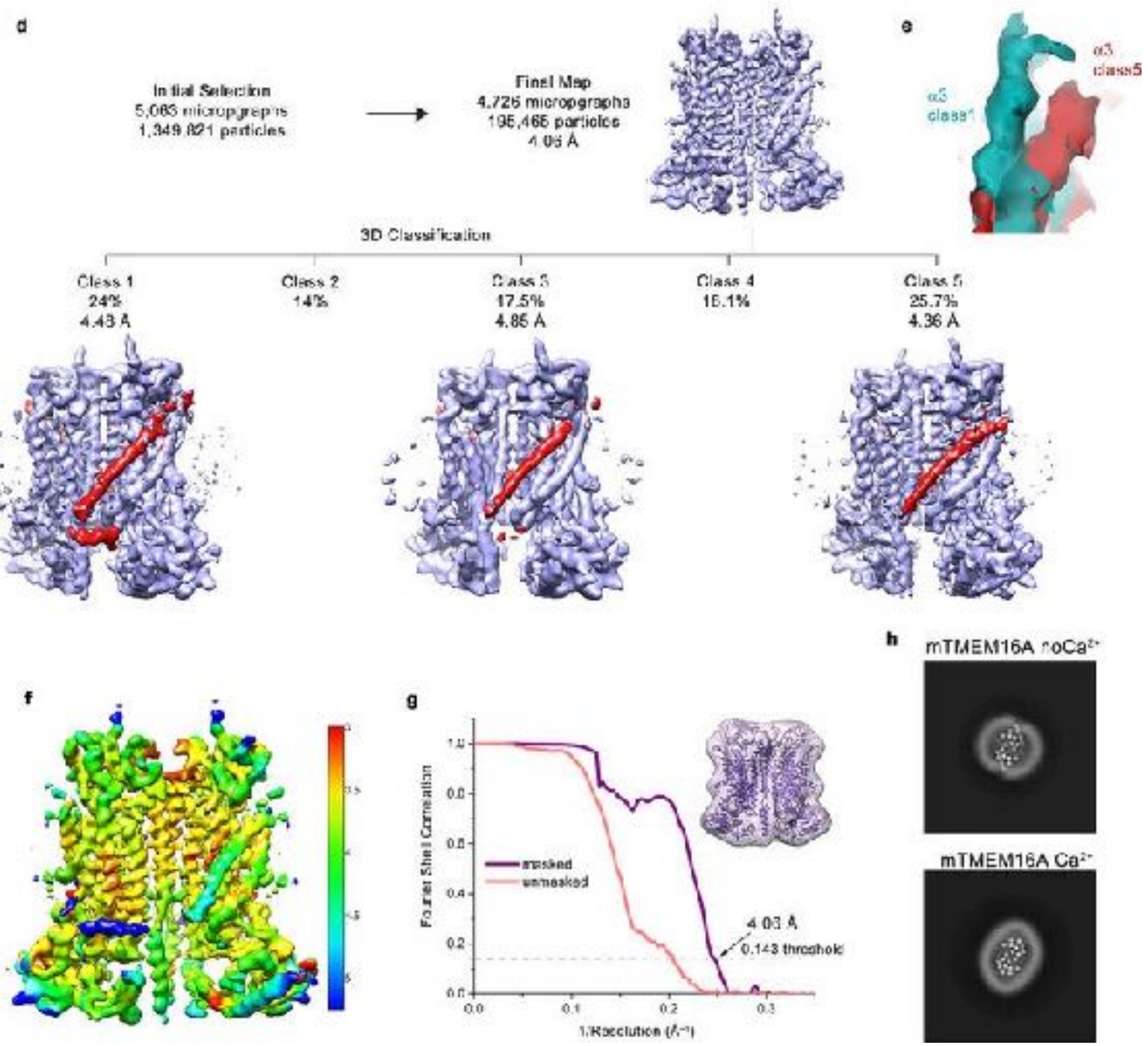
2.6 Å single particle reconstruction of aldolase from 200 keV




EMDB 8743
Herzik et al., *Nat. Methods* **14** 1075-1078 (2017

# Resolution - what does it mean?

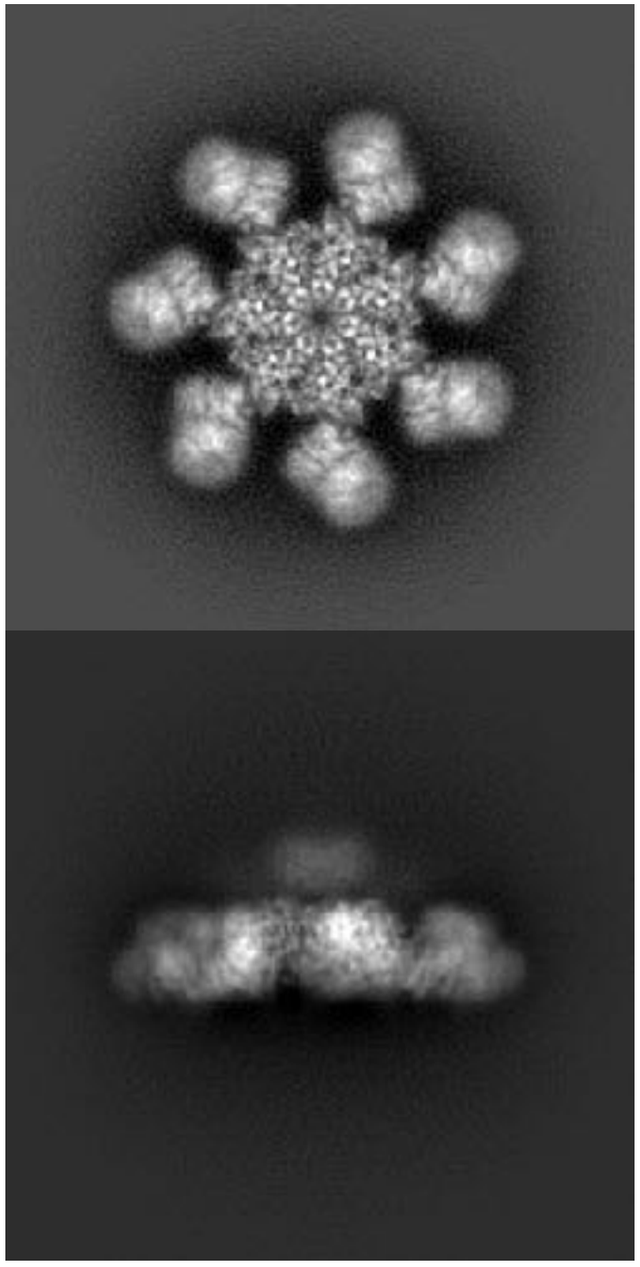
4.0 Å single particle reconstruction of TMEM16A (ion channel)




EMDB 3860
Paulino et al., *Nature* **552** 421-425 (2017)

# Resolution - what does it mean?

- Overall resolution does not equal local resolution
- Quality can vary greatly within a map and care should be taken to not over interpret poorly ordered domains
- Local resolution estimates can be performed
  - ResMap http://resmap.sourceforge.net
  - Blocalres (package in Bsoft) https://lsbr.niams.nih.gov/bsoft/programs/blocres.html
- Density slices can also be extremely informative for evaluating local map quality
- In the end, evaluate your maps features before making any claims

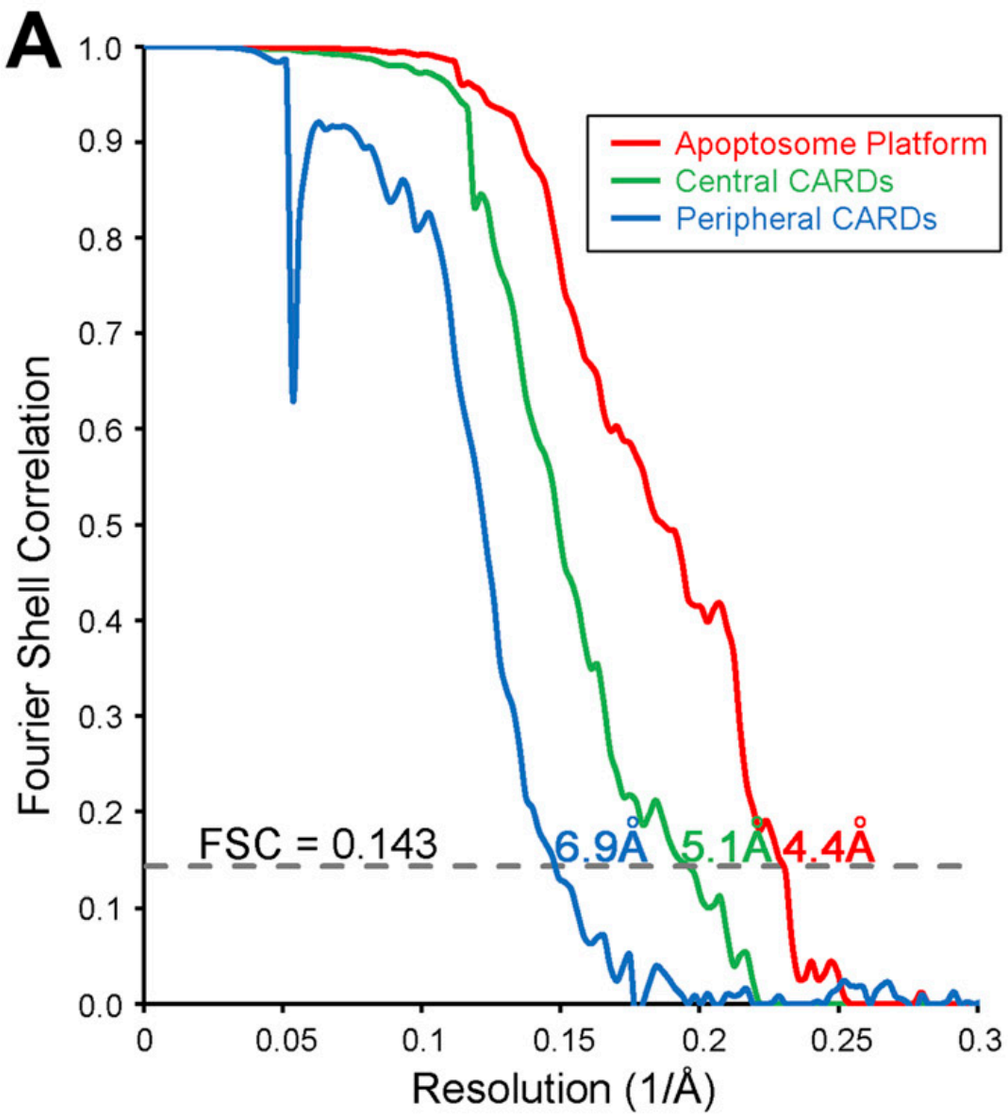


EMD-6690
4.4 Å

Li *et al*, PNAS. (2017) 114:1542-1547
Kucukelbir *et al*, Nat. Methods. (2014) 11:63-65
Heymann and Belnap, JSB. (2007) 157:3-18

# Interpretation of local maps

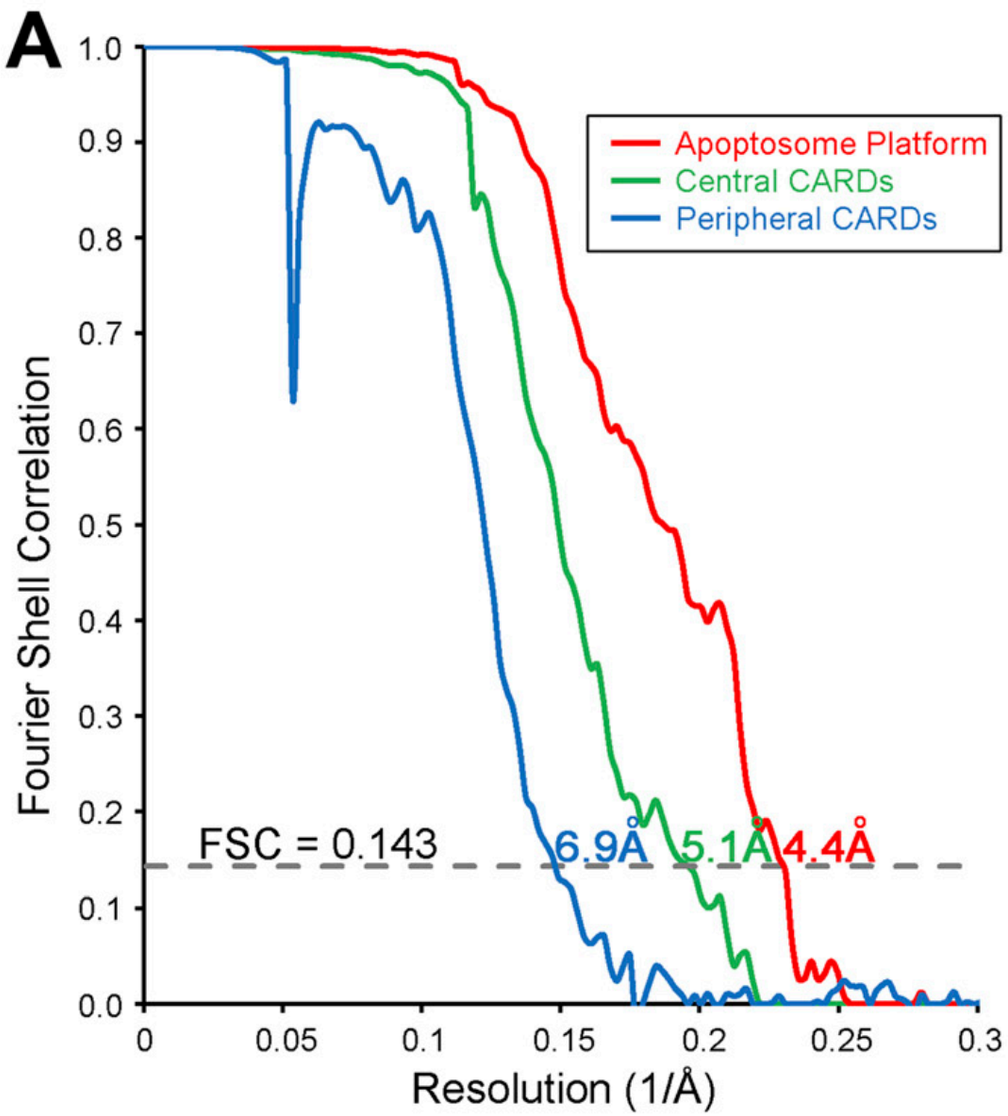
- If your specimen is large enough, masks can be used allow different domains to be refined independently
- FSC calculations performed using these different maps can allow mean resolution estimations of the individual domains
- These different resolution estimates can guide you in interpreting your maps
- The map itself is always the final guide!
- Carefully evaluate your density to learn as much as possible about your specimen




Li *et al*, PNAS. (2017) 114:1542-1547

# Model building in EM maps

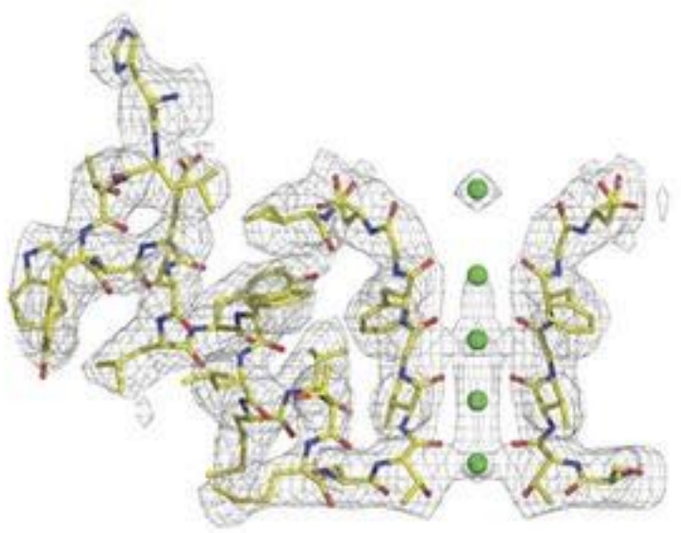
- With well-defined side chains and backbones, *de novo* atomic model building may be possible
- At slightly lower resolution, model building may only be possible with the use of homology models or other computational tools such as rosetta
- At lower resolutions, modeling is limited to docking of crystal structures
- At all resolution be careful when making specific conclusions based upon side-chain interactions - only describe what your density actually shows

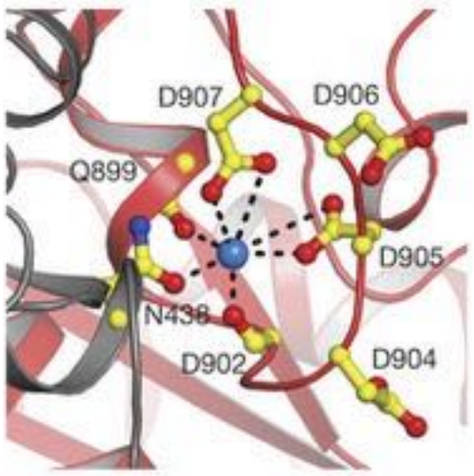


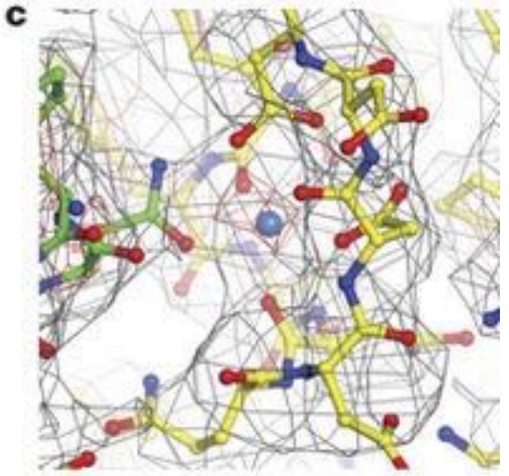

Li *et al*, PNAS. (2017) 114:1542-1547

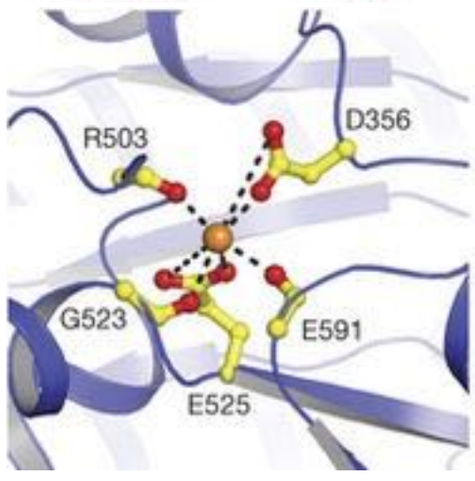
# Model building in EM maps

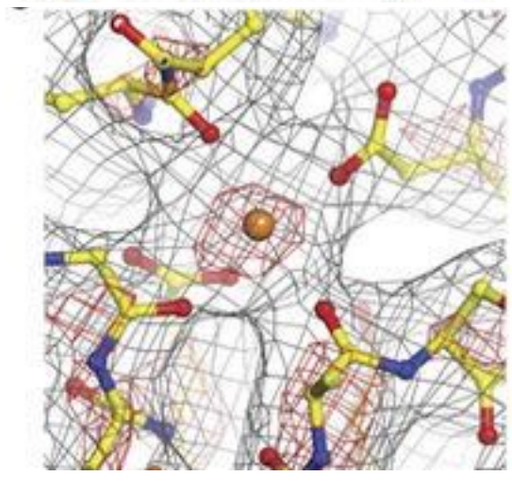
- With well-defined side chains and backbones, *de novo* atomic model building may be possible
- At slightly lower resolution, model building may only be possible with the use of homology models or other computational tools such as rosetta
- At lower resolutions, modeling is limited to docking of crystal structures
- At all resolution be careful when making specific conclusions based upon side-chain interactions - only describe what your density actually shows





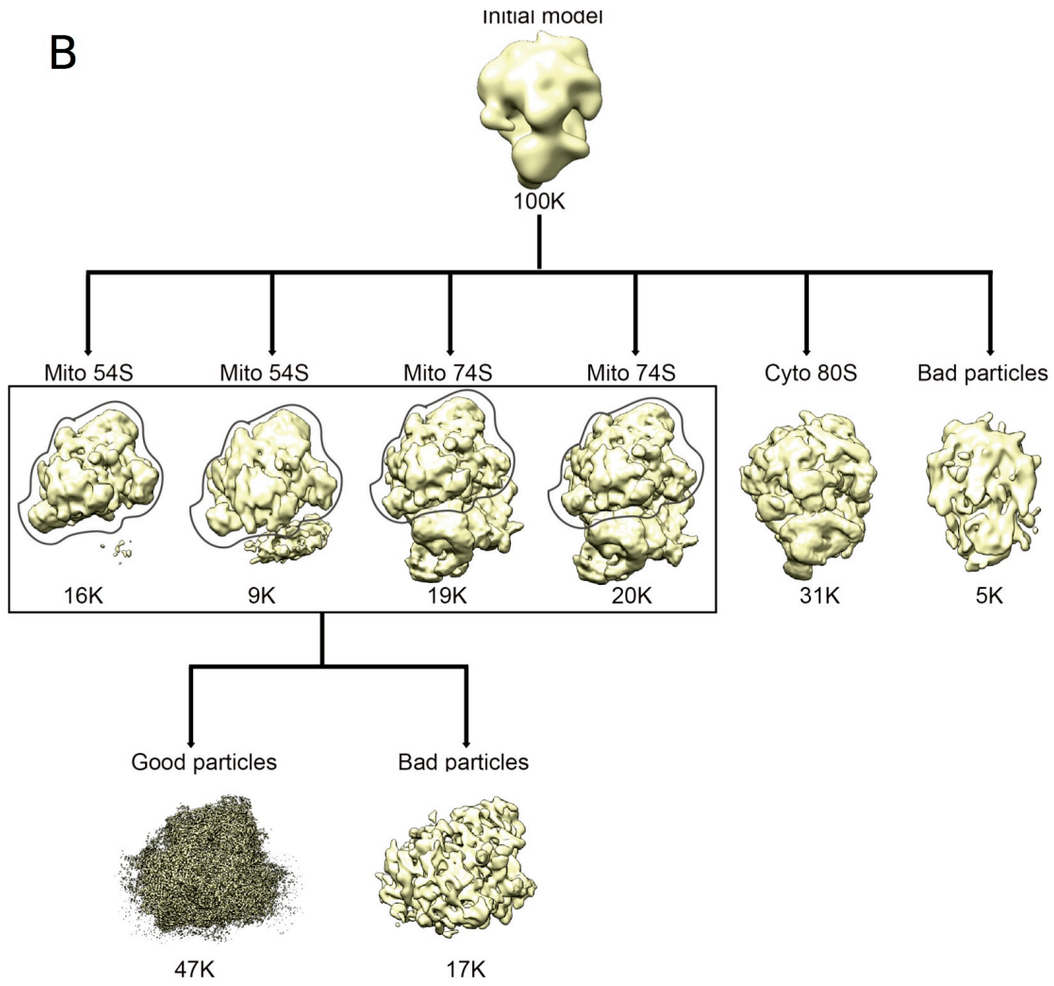



Tao *et al*, Nature. (2017) 541:46-51

# Sample heterogeneity

- There are multiple source of heterogeneity in sample preparation
  - Compositional heterogeneity - mixture of different components or stoichiometries
  - Structural heterogeneity - domains of the specimen can adopt multiple conformations
  - In some cases, both types of heterogeneity exist within a single sample
- These will degrade the resolution of reconstructions, but also provide insights into function of the specimen



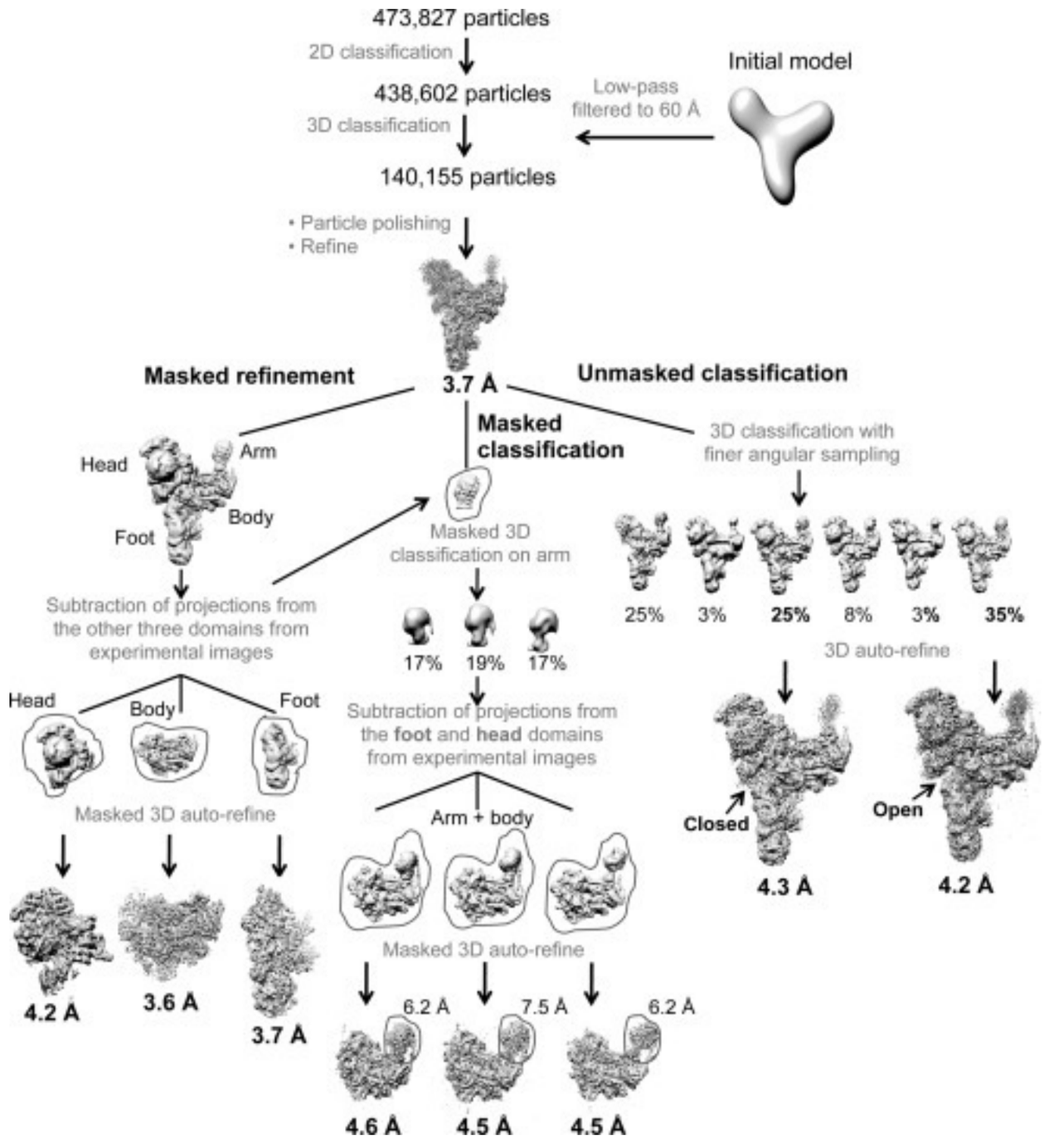

Amunts *et al*, Science. (2014) 343:1485-1489

# Overcoming heterogeneity - biochemistry

- Optimizing biochemistry can often help to alleviate heterogeneity and is generally the best place start to improve sample quality
  - Improvements in sample purification can reduce compositional heterogeneity by obtaining a more uniform starting sample
  - Structural heterogeneity can be minimized by altering purification conditions (i.e. presence of activating or inhibiting ligands, different pH or salt conditions)
  - Construct alterations can also reduce sample heterogeneity by removing flexible domains
- In some cases chemical cross-linking can helpful to reduce flexibility
  - Testing cross-linking reagents with different lengths and varying the concentration can be helpful to optimize conditions
  - However, it is essential that the chemically cross-linked structure be validated with a non-cross-structure to demonstrate the the cross-linking does not introduce artifactual protein-protein interactions

# Overcoming heterogeneity - computation

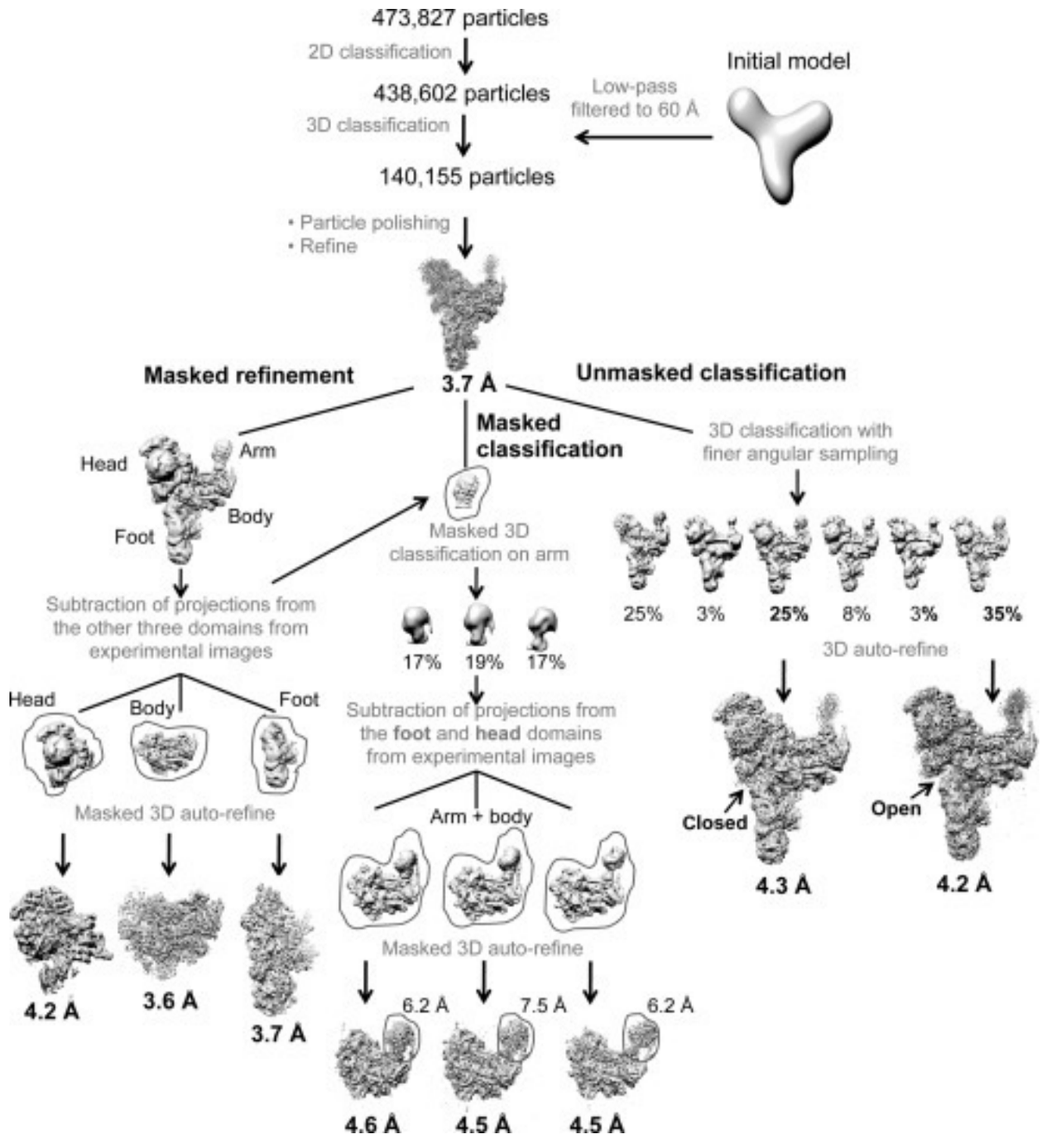
- Heterogeneity may be unavoidable for some samples and must be dealt with computationally after image acquisition
- There are now several different software packages that sort and classify particles, allowing one to create “pure” subsets of the particles images
- The simplest approach is classify based upon the entire molecule, which works well with large conformational differences




Scheres, Meth Enz. (2016) 579:125-157

# Overcoming heterogeneity - computation

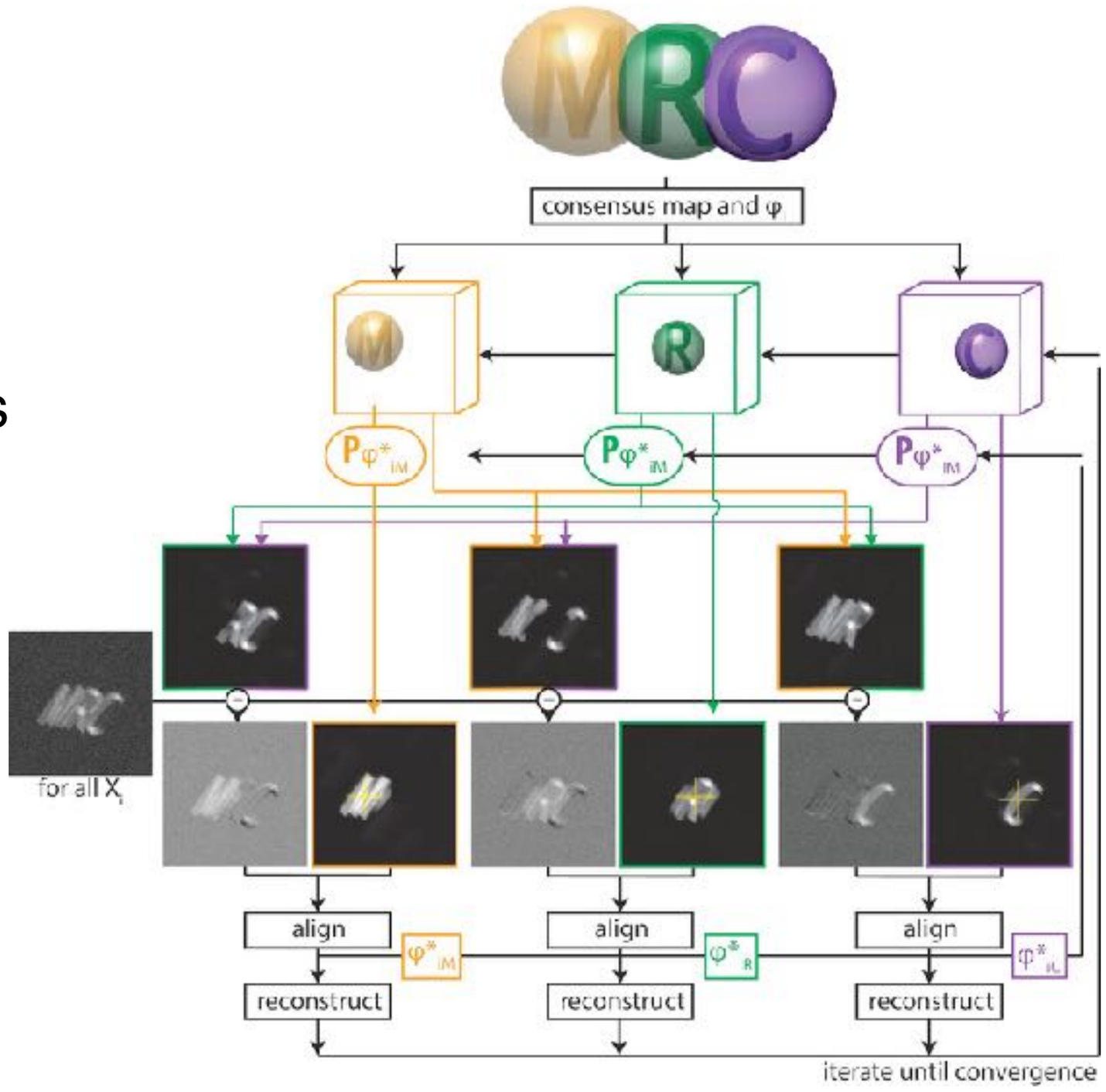
- Classification can be enhanced through the use of masks
- A mask can be placed around the region of interest - allowing independent sorting of different domains
- This multi-classification approach is particularly powerful for samples that have multiple different types of movements
- Another modification to classification is the use of background subtraction prior to classification to reduce the signal of constant domains during classification




Scheres, Meth Enz. (2016) 579:125-157

# Overcoming heterogeneity - computation

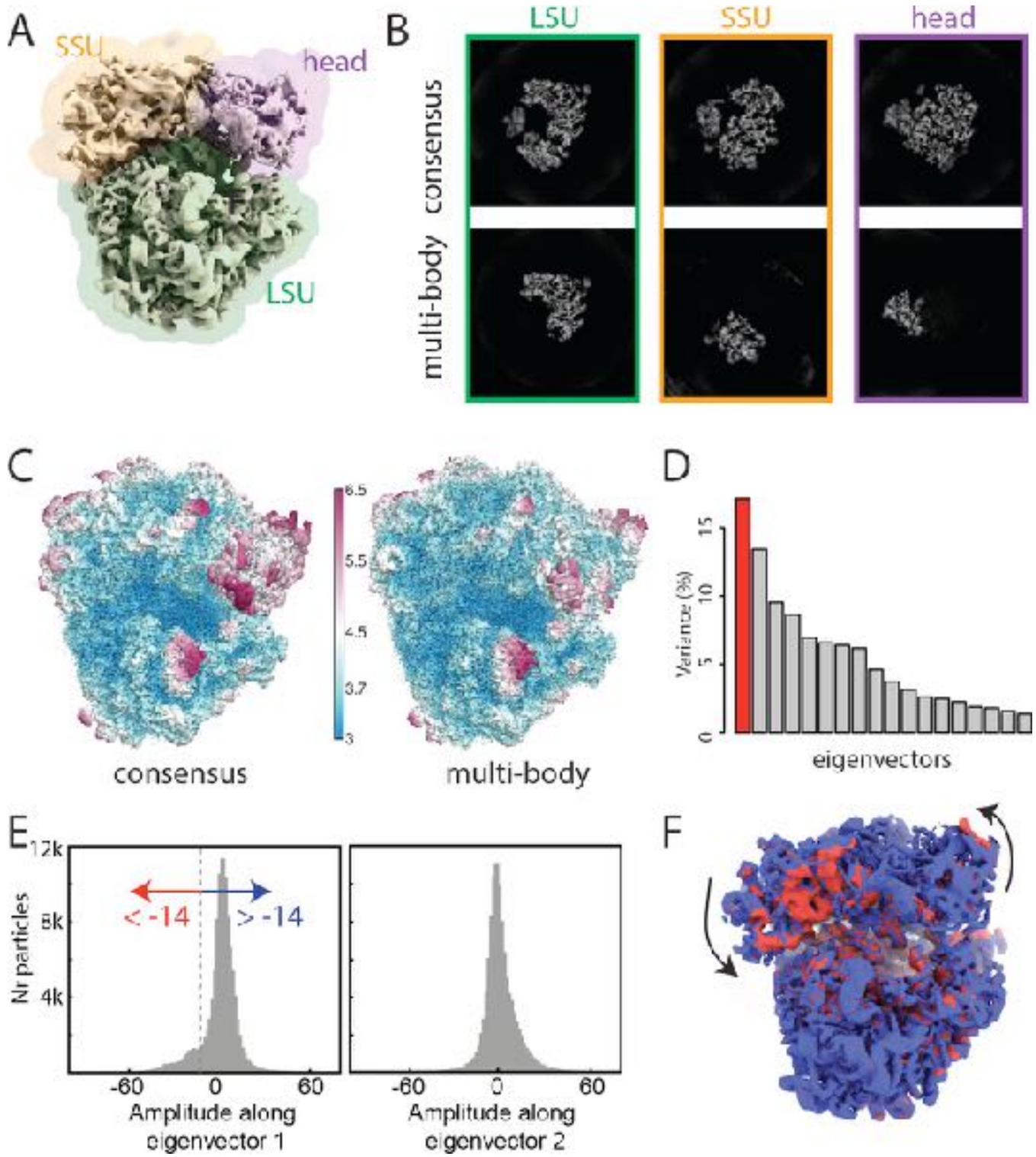
- Relion has an automated procedure to apply masks based upon distinct flexible domains
- This approach is known as multi-body refinement
- It also determines the vectors of movement allowing an understanding of conformational dynamics across protein complexes




Nakane et al, eLife. (2018) e36861

# Overcoming heterogeneity - computation

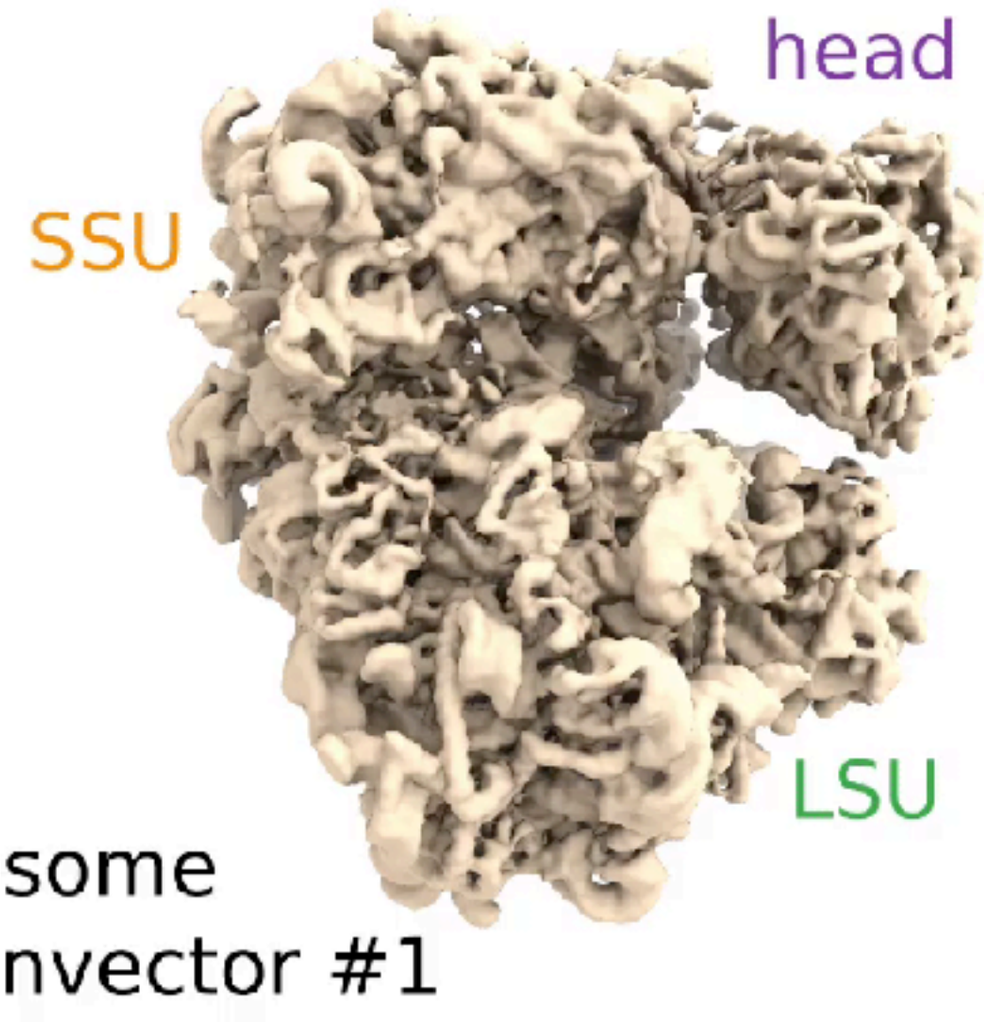
- Relion has an automated procedure to apply masks based upon distinct flexible domains
- This approach is known as multi-body refinement
- It also determines the vectors of movement allowing an understanding of conformational dynamics across protein complexes




Nakane et al, eLife. (2018) e36861

# Overcoming heterogeneity - computation

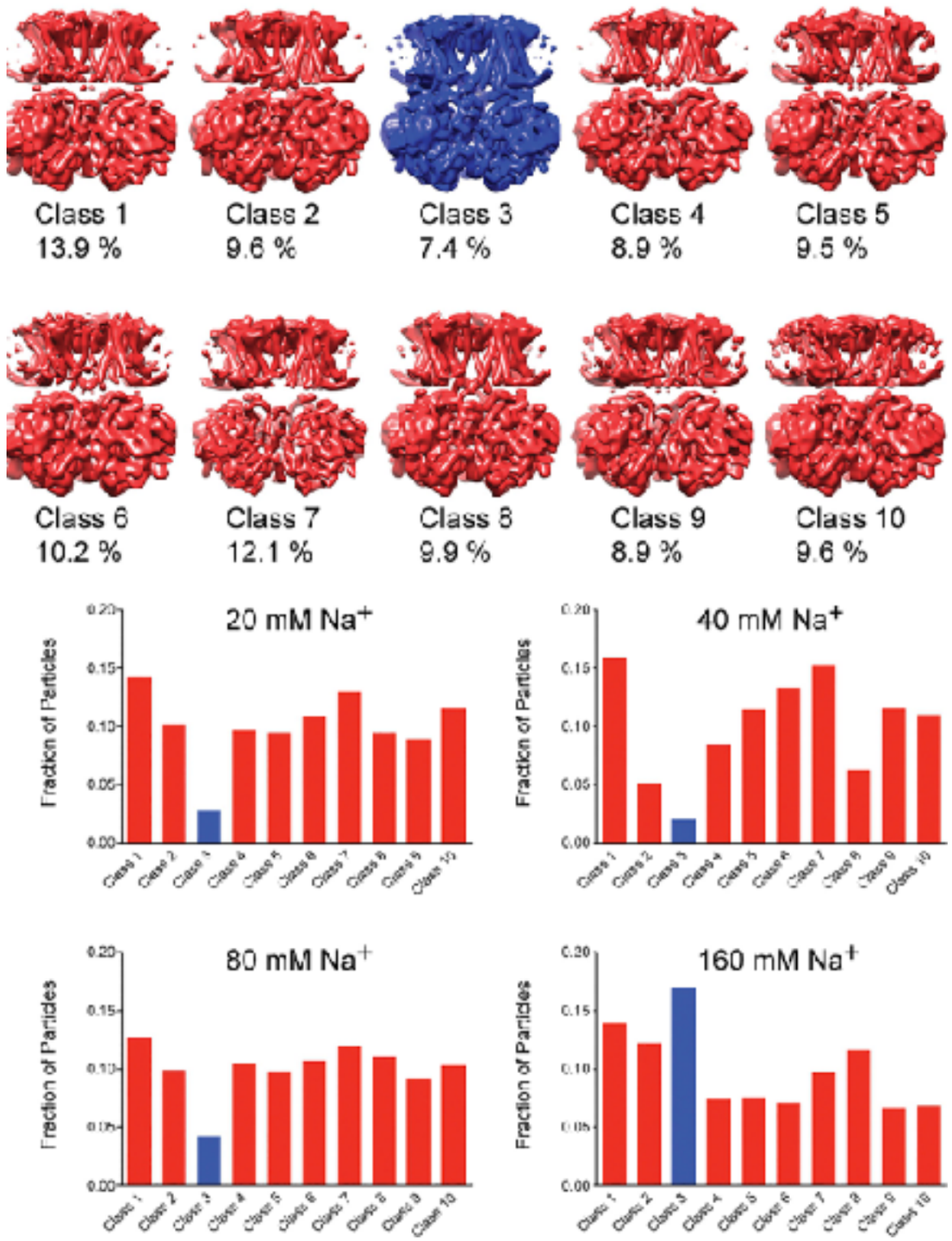
- Relion has an automated procedure to apply masks based upon distinct flexible domains
- This approach is known as multi-body refinement
- It also determines the vectors of movement allowing an understanding of conformational dynamics across protein complexes




Nakane et al, eLife. (2018) e36861

# Benefits of heterogeneity?

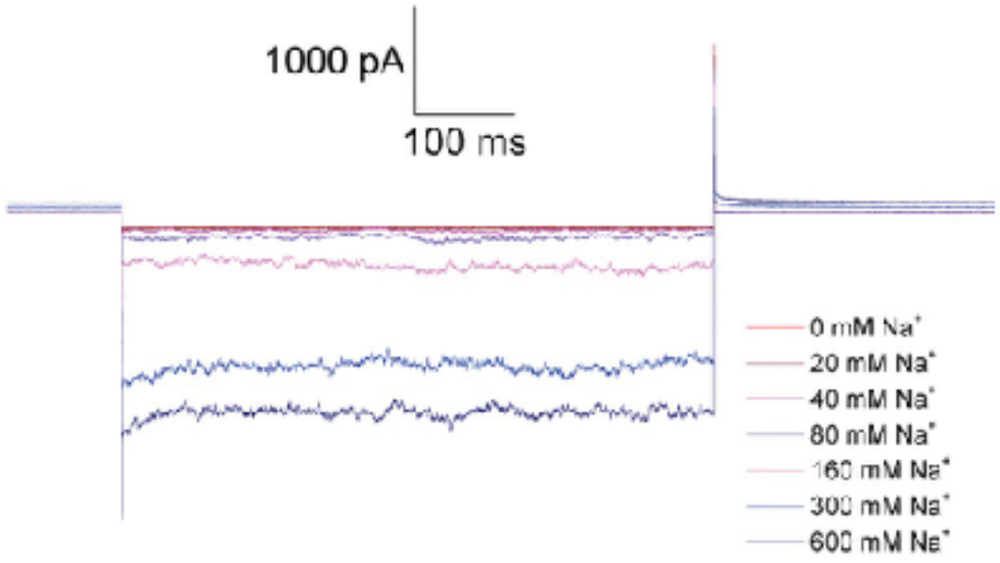
- How can you use heterogeneity to better understand the biology of your samples?
- Does your heterogeneity correlate with functional changes?
- Always test to ensure that your representative density map is actually representative of your sample, and not merely some small portion of the particles that generate a high-resolution structure?
  - If the map does result from a very small fraction of particles, try to understand why?
  - Can you test activity to see if that makes sense biologically?

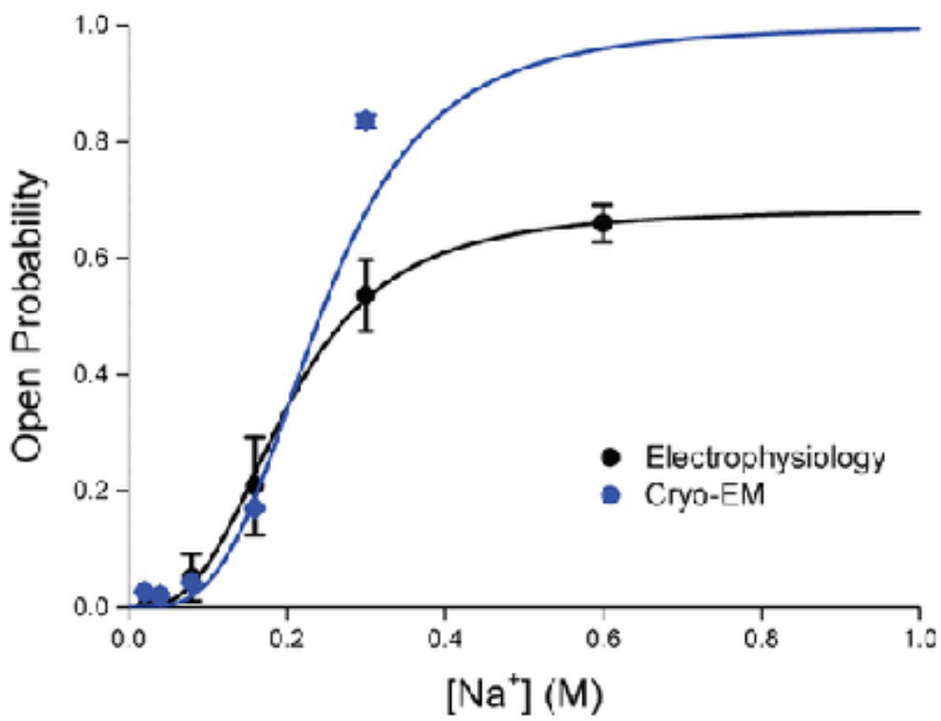



Hite and MacKinnon, Cell. (2017) 168:390-399

# Benefits of heterogeneity?

- How can you use heterogeneity to better understand the biology of your samples?
- Does your heterogeneity correlate with functional changes?
- Always test to ensure that your representative density map is actually representative of your sample, and not merely some small portion of the particles that generate a high-resolution structure?
  - If the map does result from a very small fraction of particles, try to understand why?
  - Can you test activity to see if that makes sense biologically?




Hite and MacKinnon, Cell. (2017) 168:390-399