

# **Validation Methods**

**A blast from the past ...**

**NYSBC-NCCAT 2020  
Single-particle cryo-EM Course**

# The dark side of single-particle EM

The great thing about single-particle EM:  
Every data set and processing approach yields a 3D map !

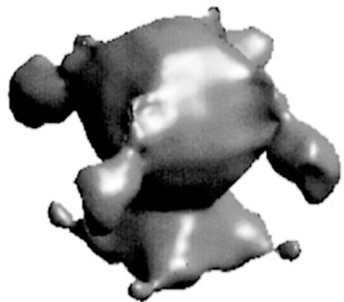
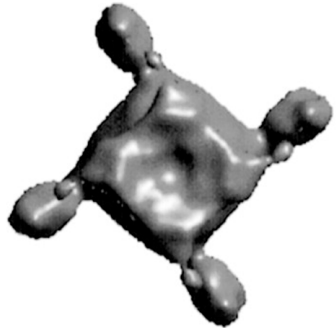
The bad thing about single-particle EM:  
Every data set and processing approach yields a 3D map !

But is it correct ???

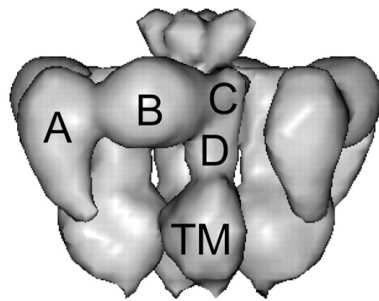
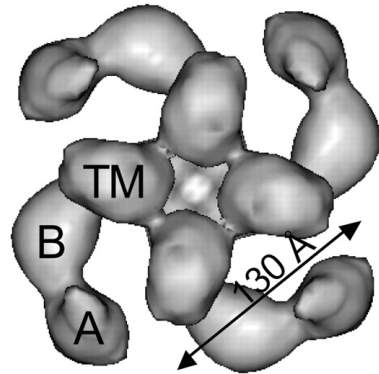


Particularly problematic  
for low-resolution maps

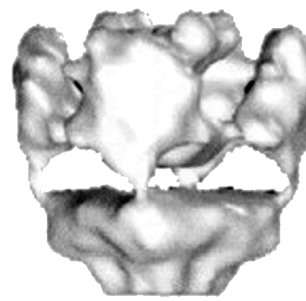
# The issue: Structures of the IP3 receptor as determined by single-particle EM



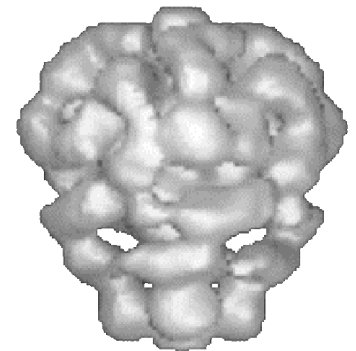
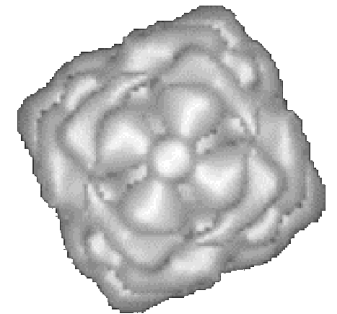
Jiang *et al.*,  
2002



Serysheva *et al.*,  
2003

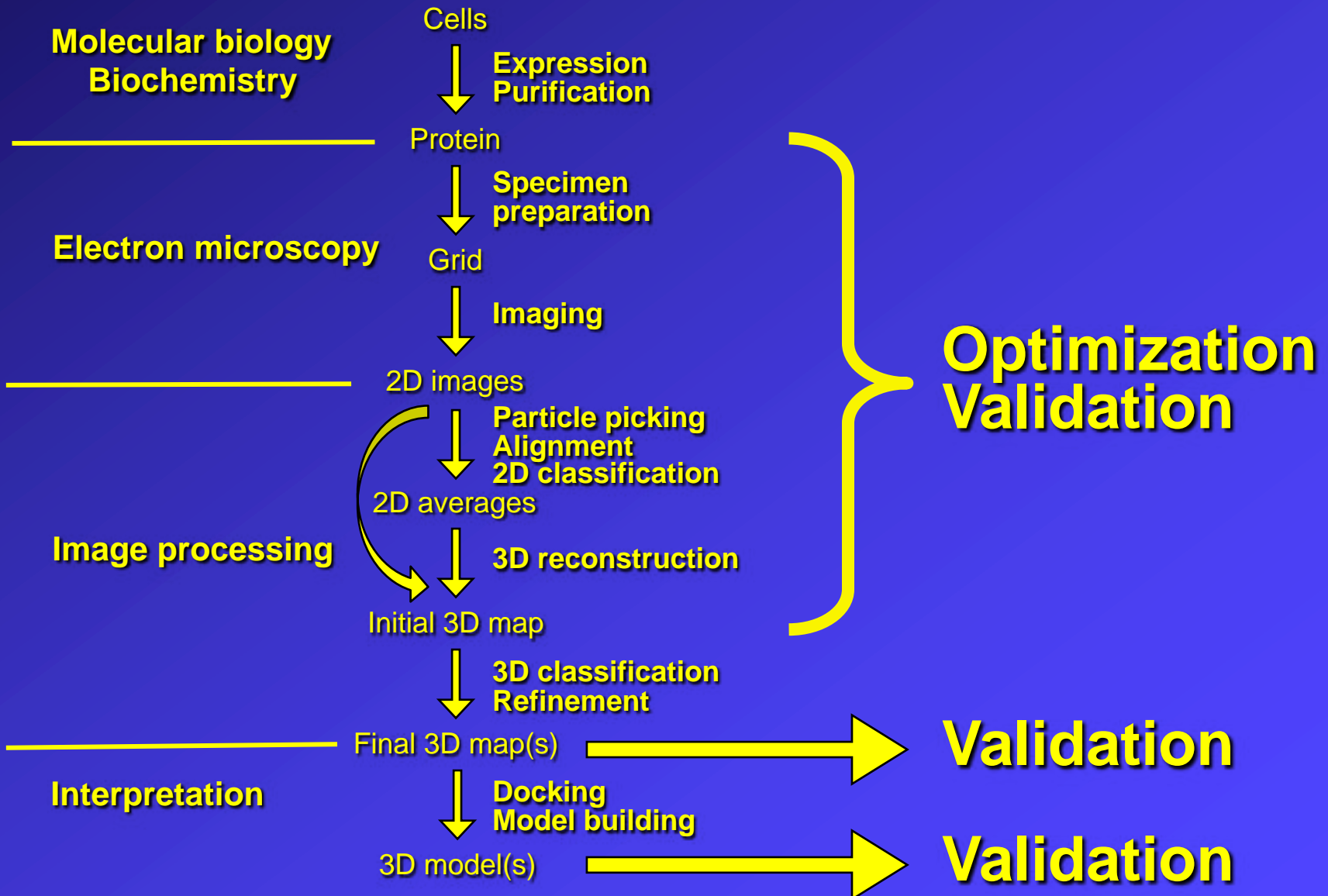


Jiang *et al.*,  
2003



Sato *et al.*,  
2004

# Structure determination by single-particle EM



# Structure determination by single-particle EM



## Potential issues:

### Heterogeneity

- Compositional
- Conformational
  - Discrete states
  - Continuous movement

### Effect of cross-linking



# Structure determination by single-particle EM

## Potential issues with samples

Before attempting structure determination –  
Understand and optimize your sample !

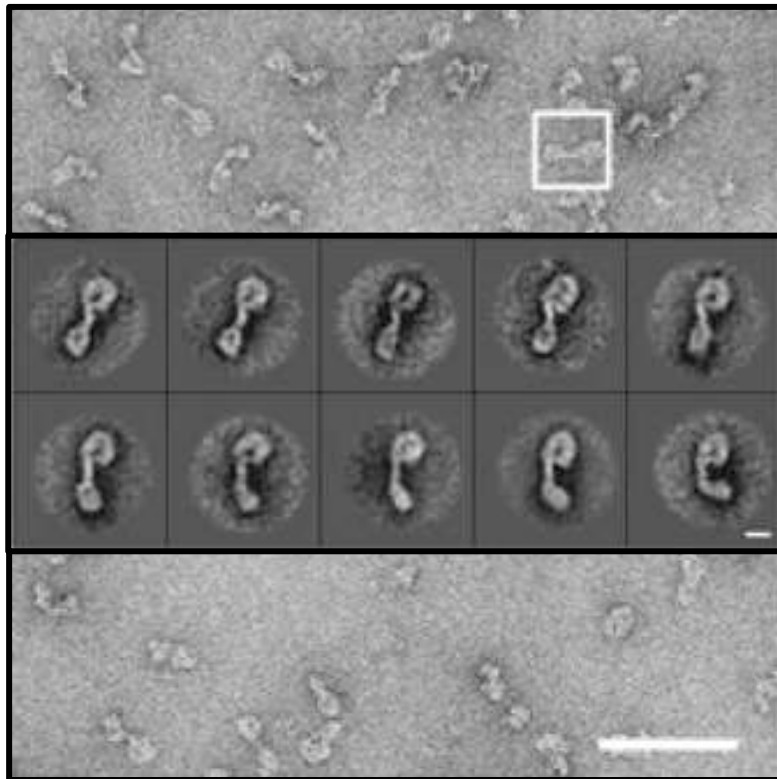
Prepare negatively stained specimens:  
Good contrast and preferred orientations  
→ Easy to assess heterogeneity

If particles look heterogeneous:  
Calculate class averages  
→ Assess type and degree of heterogeneity  
→ Minimize heterogeneity by any means possible

If chemical fixation was used:  
Look at unfixed sample to assess effect of cross-linking  
→ Assess whether structure of cross-linked sample is meaningful

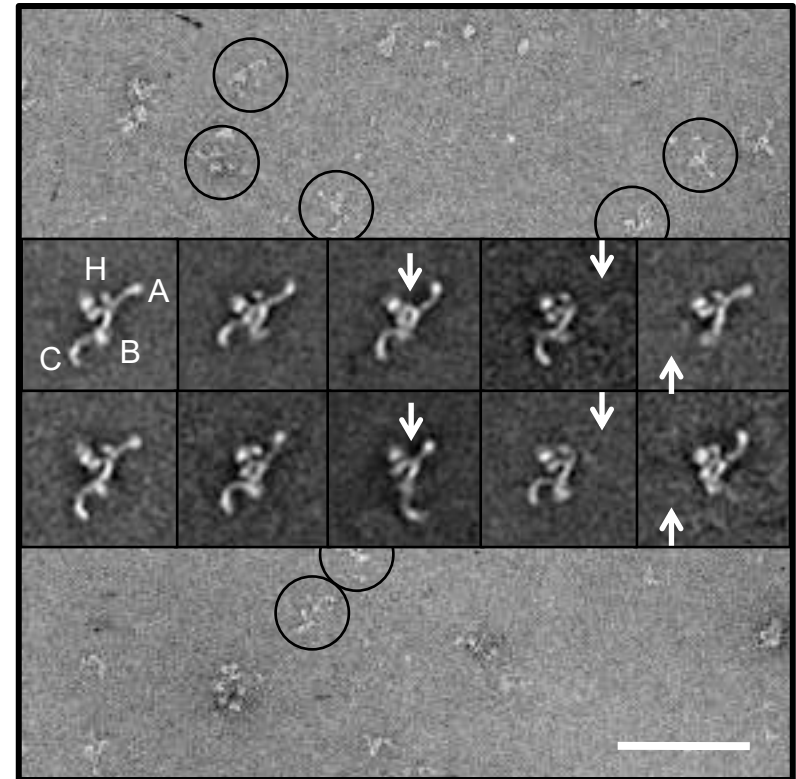
# Effect of cross-linking: The HOPS tethering complex

Cross-linked



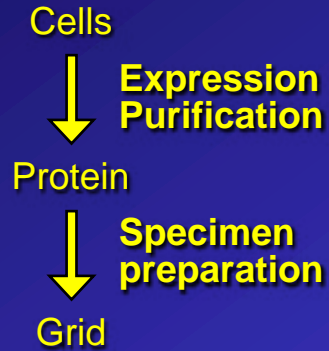
Bröcker *et al.* (2012)  
*PNAS* 109: 1991-1996

Native



Chou *et al.* (2016)  
*NSMB* 23: 761-763

# Structure determination by single-particle EM



## Potential issues:

- No particles
- Preferred orientations



# Structure determination by single-particle EM

## Potential issues with grids

No particles (particles bind to carbon and avoid holes)

- Increase protein concentration
- Double blotting
- Use thin support film (carbon or graphene oxide)
- Use different grids, e.g., PEG-treated or gold grids

Preferred orientation (particles align at air/water interface)

Lack of views will result in:

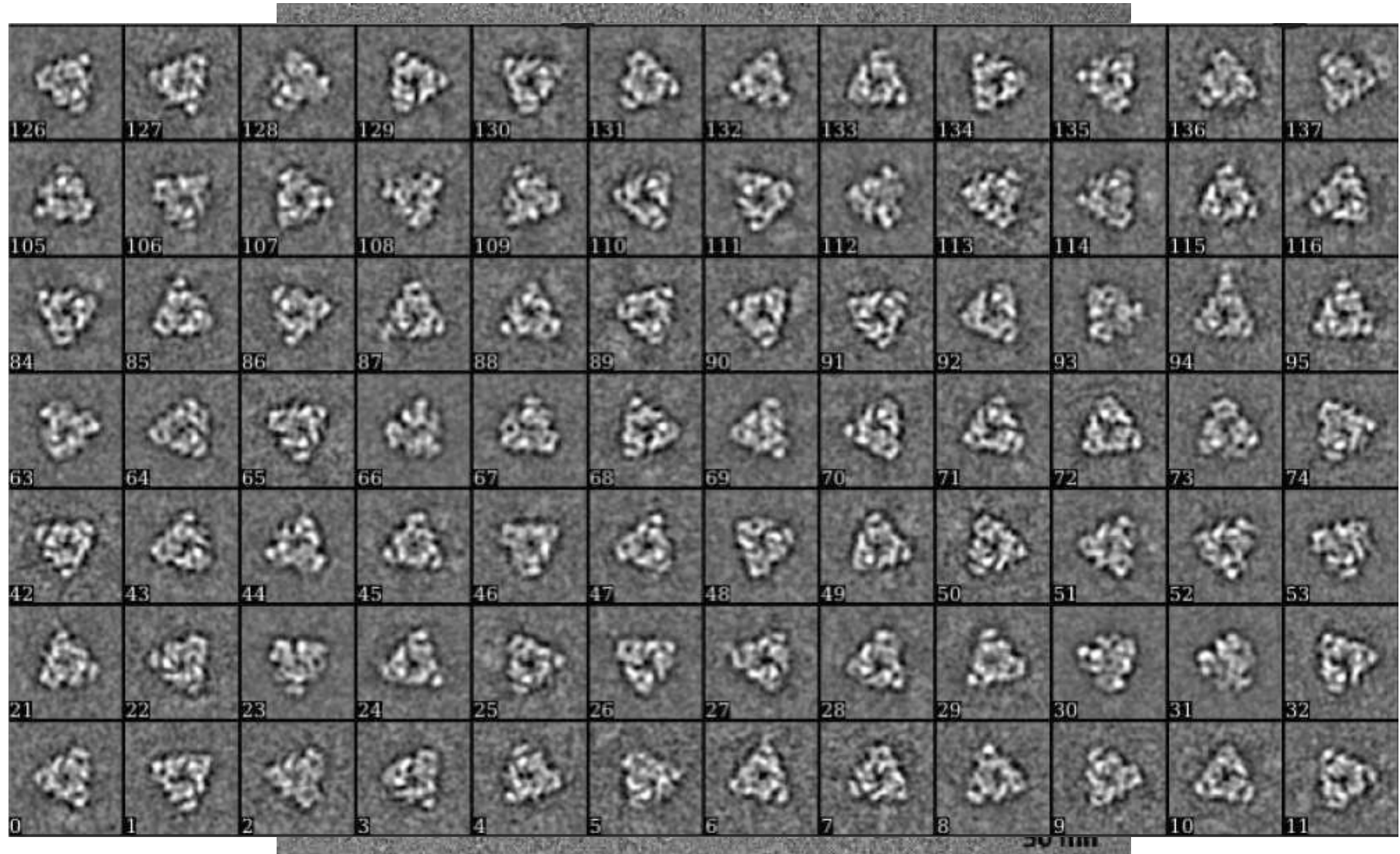
- non-isotropic resolution of the density map
- can potentially lead to an incorrect density map
- Use low concentration of detergent (changes surface tension)
- Use thin carbon film (commonly used for ribosome samples)
- Use gold grids

Different sample preparation approach (e.g., Spotiton)

Collect images from tilted specimens

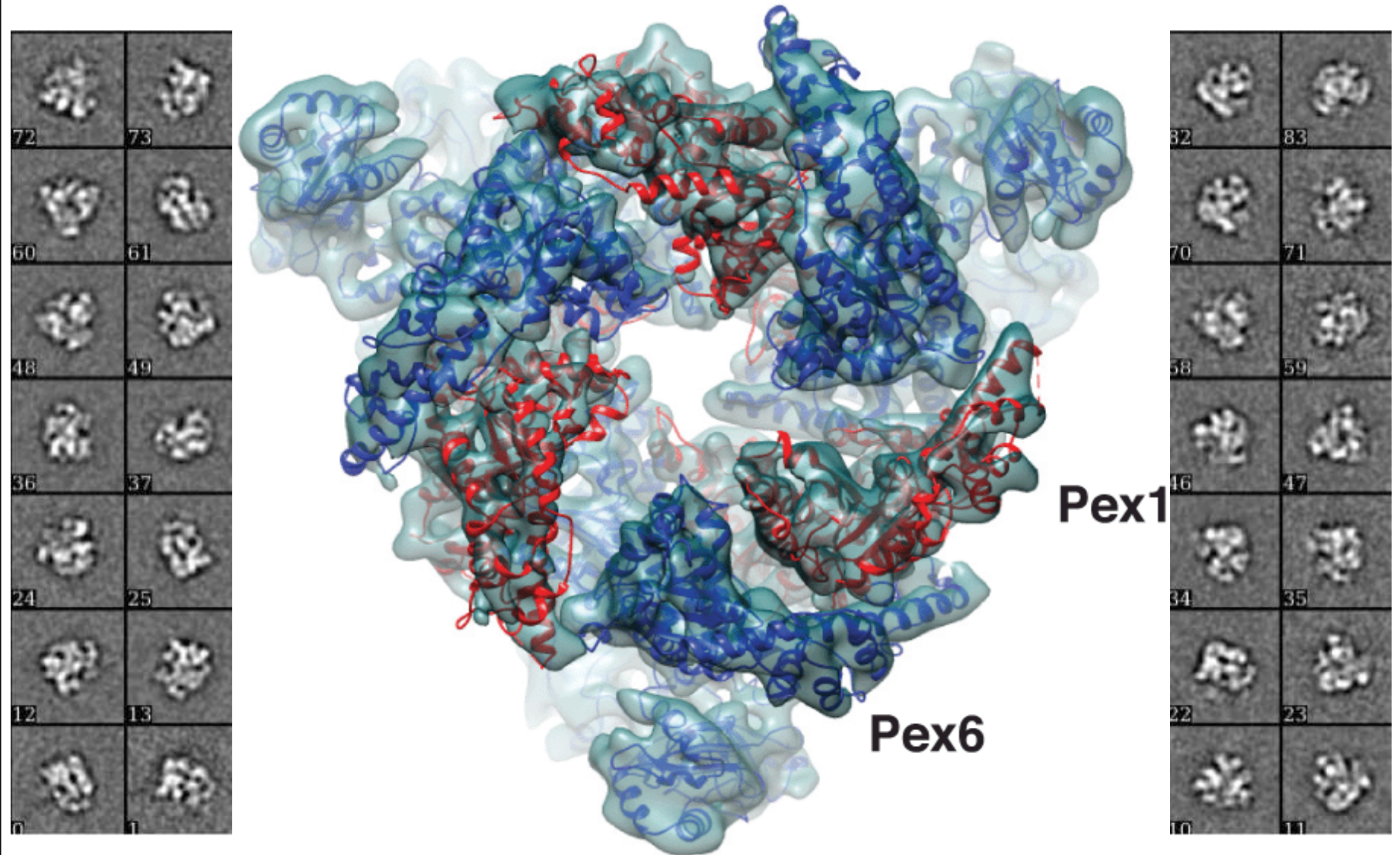
# Preferred orientations: Pex1/6 complex

## Without detergent

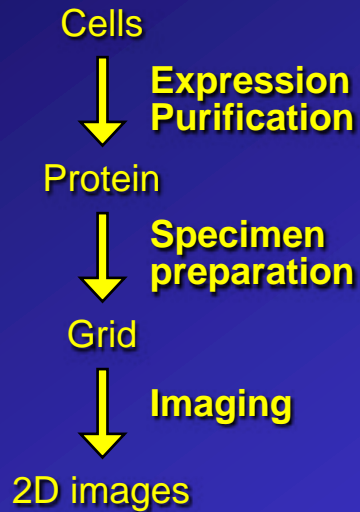




# Preferred orientations: Pex1/6 complex With detergent



# Structure determination by single-particle EM



## Potential issues:

- Low contrast
- Beam damage



# Structure determination by single-particle EM

## Potential issues with images

Poor electron scattering  
→ high electron dose



Beam sensitivity  
→ low electron dose

→ Poor SNR can be fixed  
by averaging

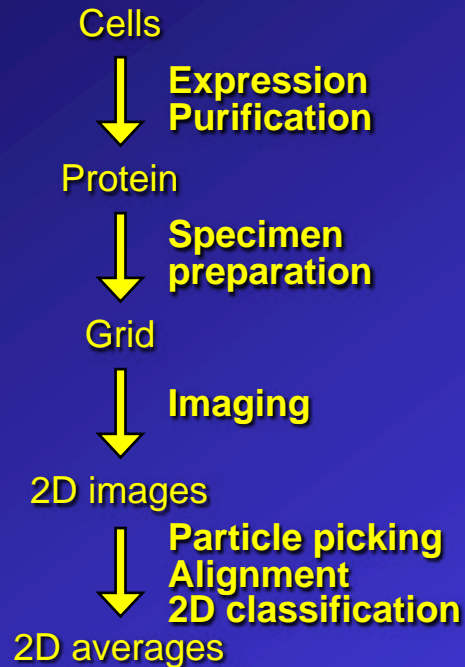
→ Loss of information  
cannot be fixed

→ Electron micrographs recorded with low electron doses  
→ Particles hard to see, especially small ones

## Problem fixed by DDD cameras

- Collect long movies
- Add frames with resolution filter  
(removes damaged high-resolution information  
retains low-resolution information for good SNR)

# Structure determination by single-particle EM



## Potential issues:

### Particle picking:

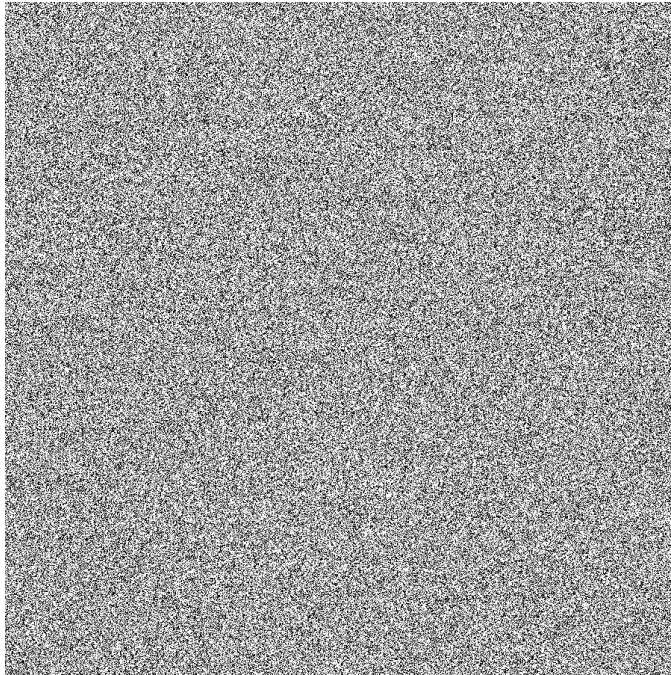
- Model/reference bias

### 2D classification:

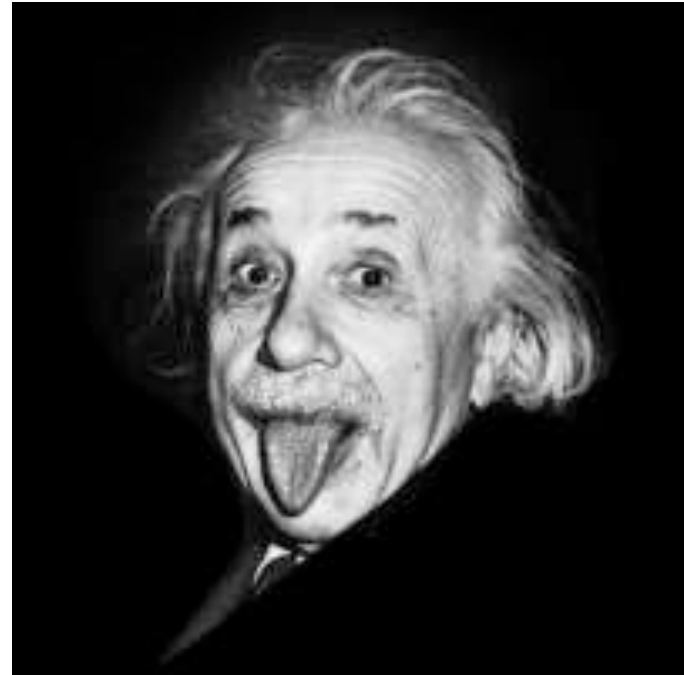
- Model/reference bias
- Number of classes
- Heterogeneous classes
- Disappearing classes

# Structure determination by single-particle EM

## Potential issues with particle picking



1,000 images containing  
pure white noise



Reference:  
Albert Einstein

Shatsky *et al.* (2009) *J. Struct. Biol.* 166: 67-78

Henderson (2013) *Proc. Natl. Acad. Sci. USA* 110: 18037-18041



# Structure determination by single-particle EM

## Potential issues with particle picking



Model/reference bias

Average of 1,000 images containing pure white noise after alignment to an image of Albert Einstein

→ Einstein from noise

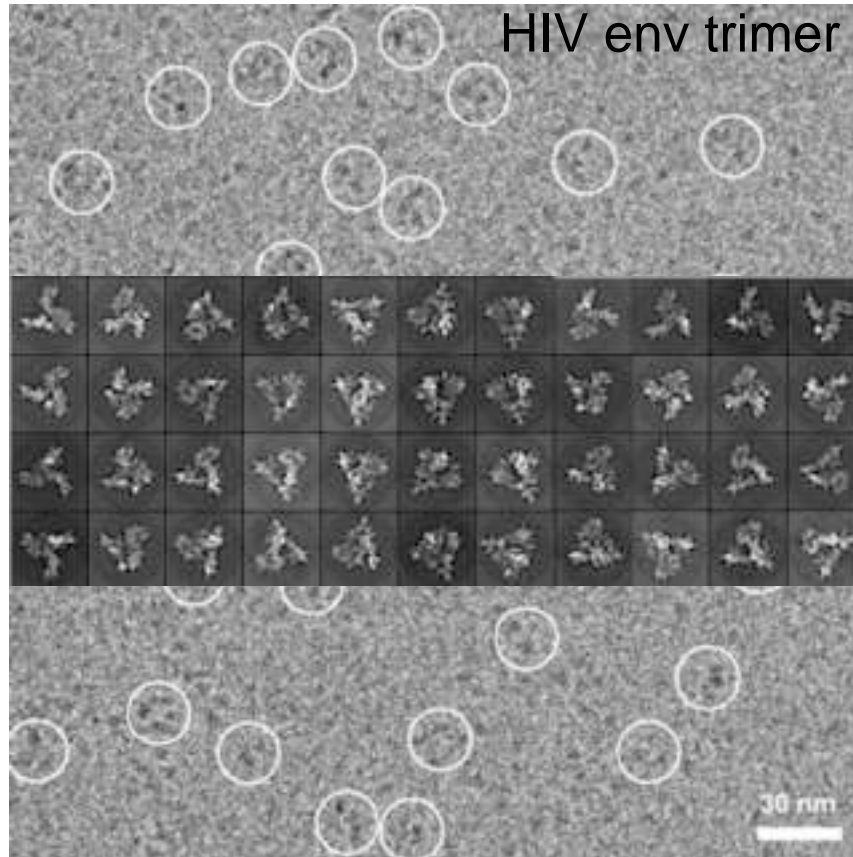
Shatsky *et al.* (2009) *J. Struct. Biol.* 166: 67-78

Henderson (2013) *Proc. Natl. Acad. Sci. USA* 110: 18037-18041

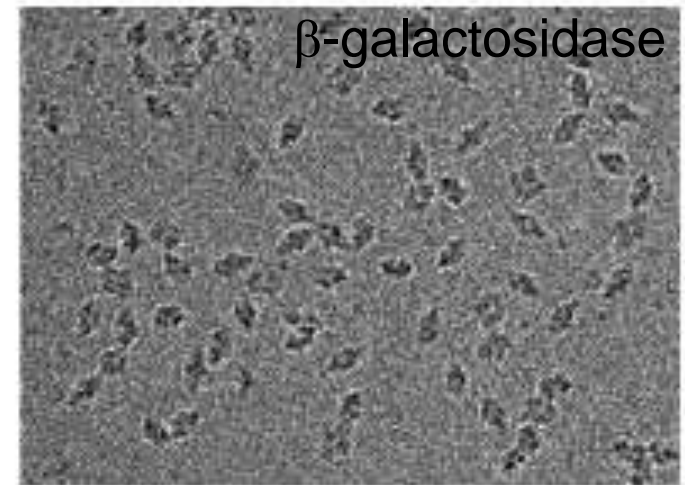
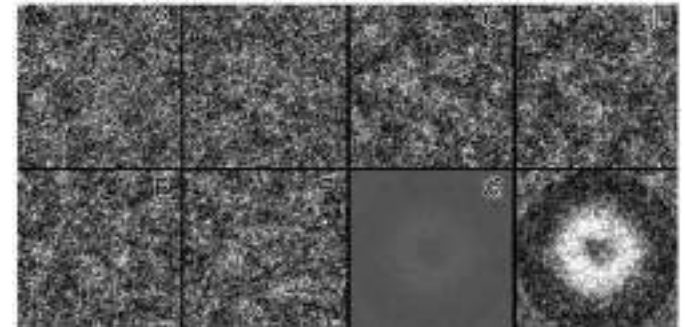


# Structure determination by single-particle EM

## Potential issues with particle picking



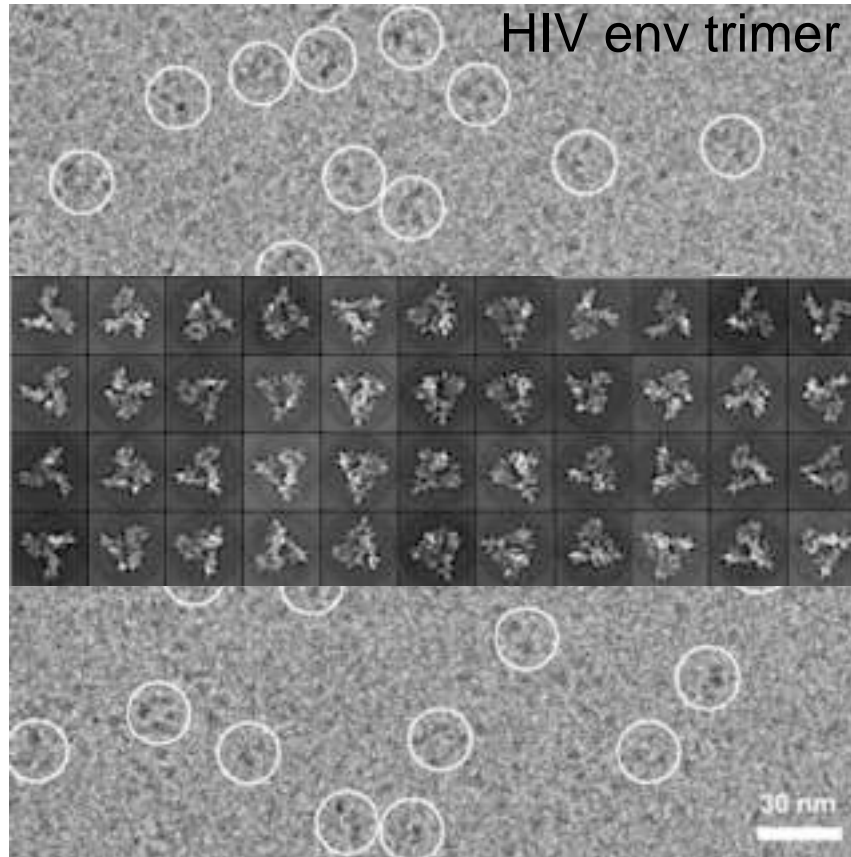
Mao *et al.* (2013)  
*PNAS* 110: 12438-12443



Henderson (2013)  
*PNAS* 110: 18037-18041

# Structure determination by single-particle EM

## Potential issues with particle picking



Mao *et al.* (2013)  
*PNAS* 110: 12438-12443

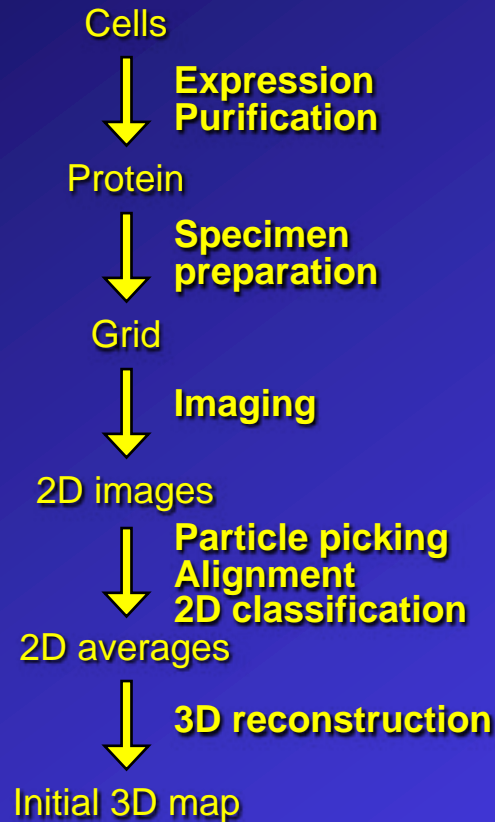
Using template matching to pick particles from very noisy images is dangerous

→ Averages will end up looking like templates used for particle picking

→ Better to first pick images without templates and use resulting averages as templates for re-picking

Henderson (2013)  
*PNAS* 110: 18037-18041

# Structure determination by single-particle EM



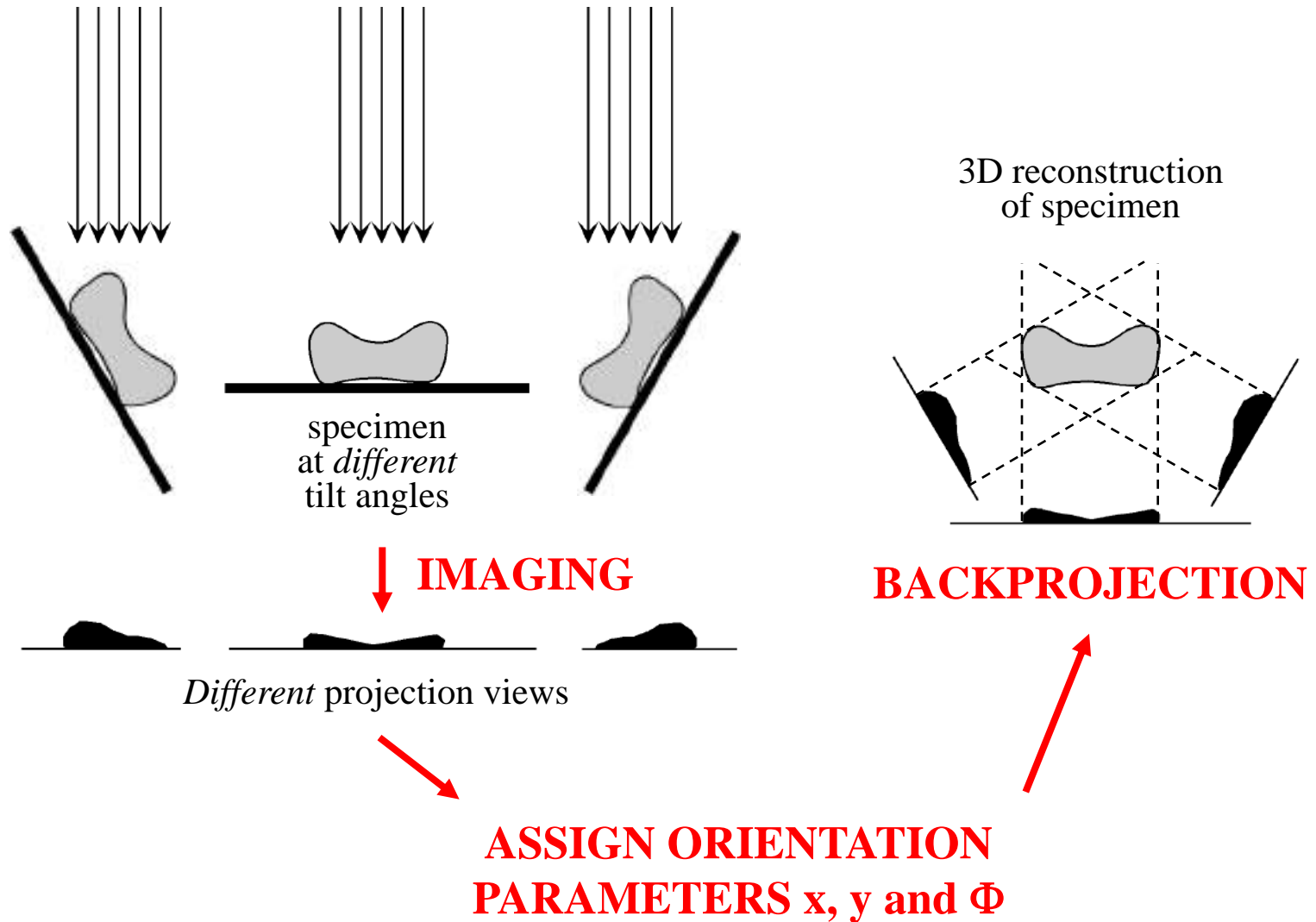
## Potential issues:

Incorrect map

Because of:

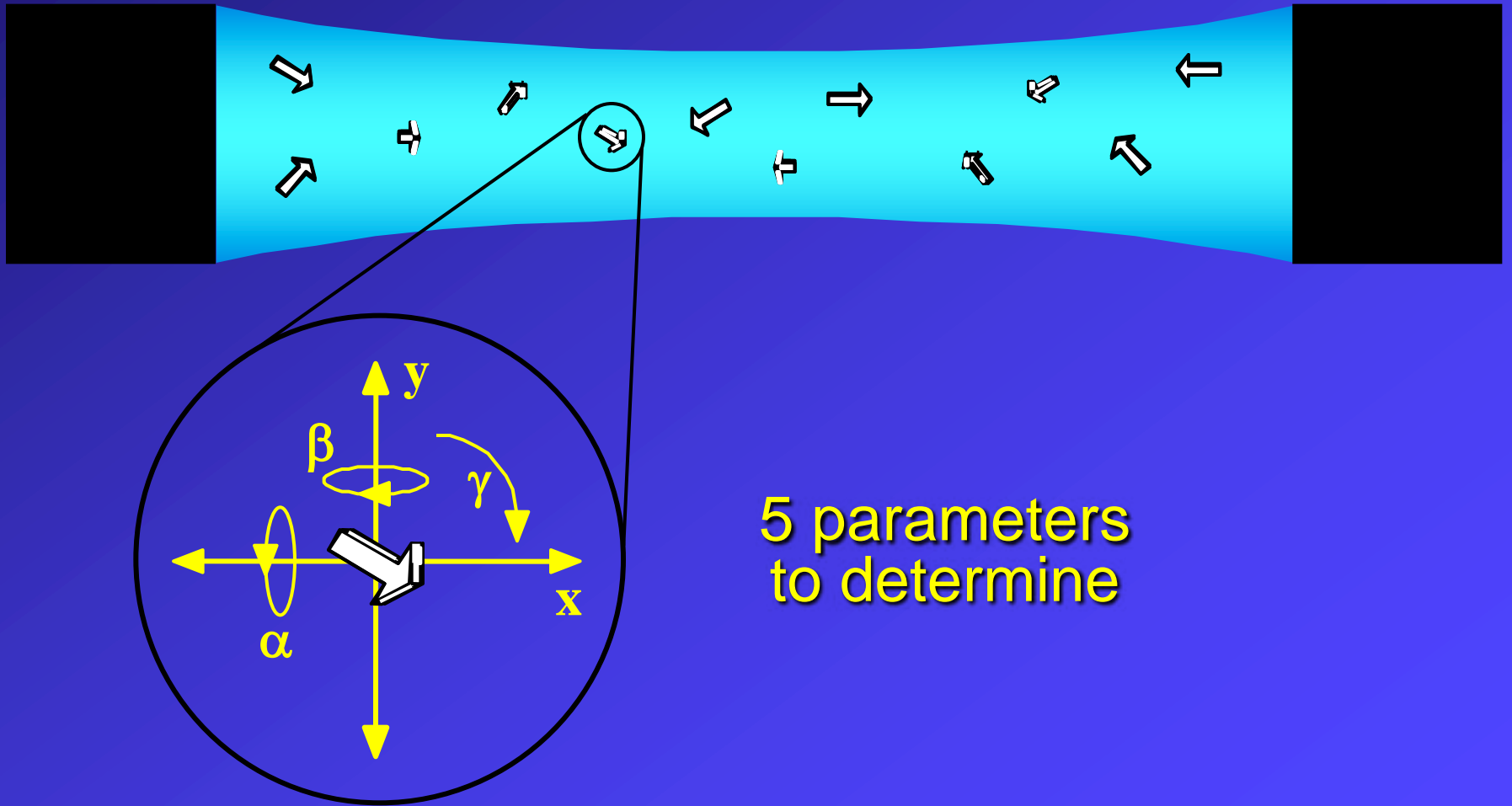
- Heterogeneous sample
- Missing views
- Incorrect solution

# Random conical tilt reconstruction





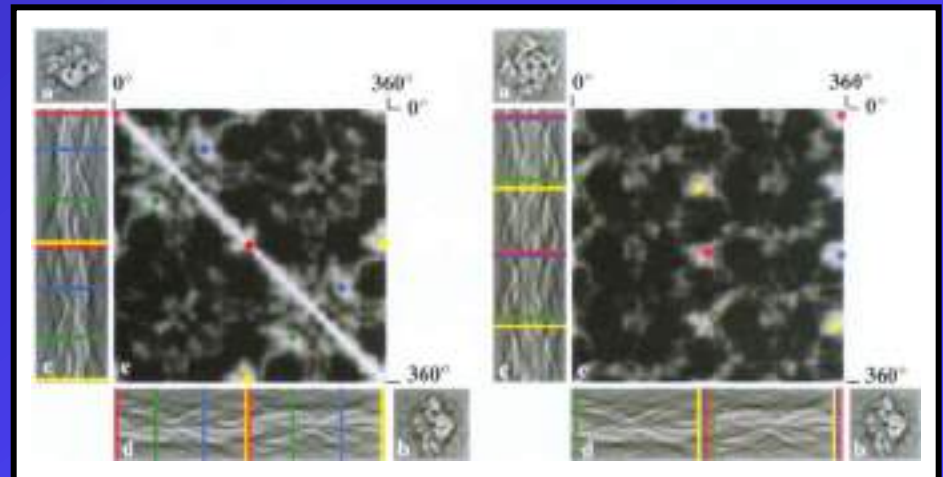
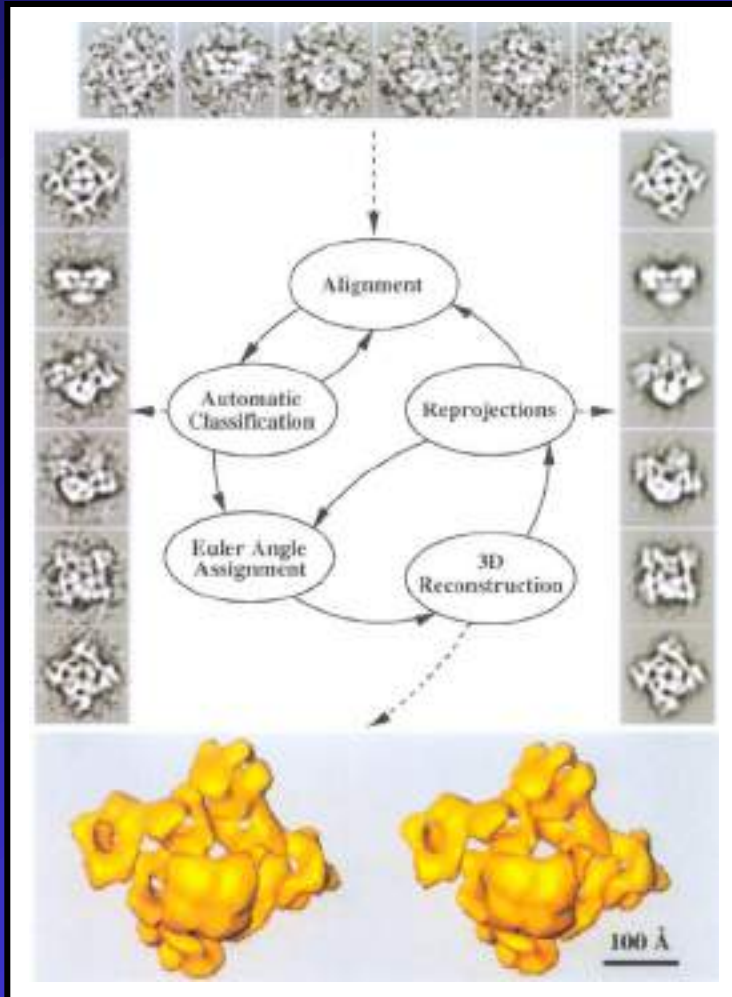
# Single particles in ice



# Angular reconstitution

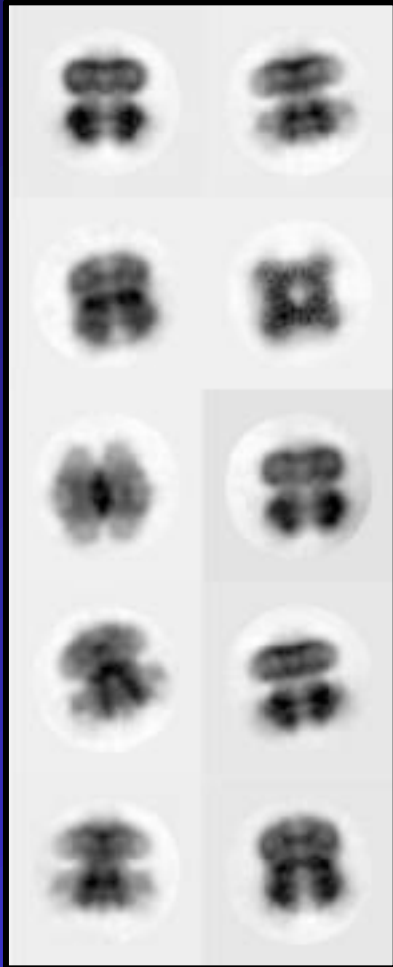
van Heel, 1987

1. choose 3 projection images that are perpendicular views of the particle (anchor set)
2. add in further projections and keep refining

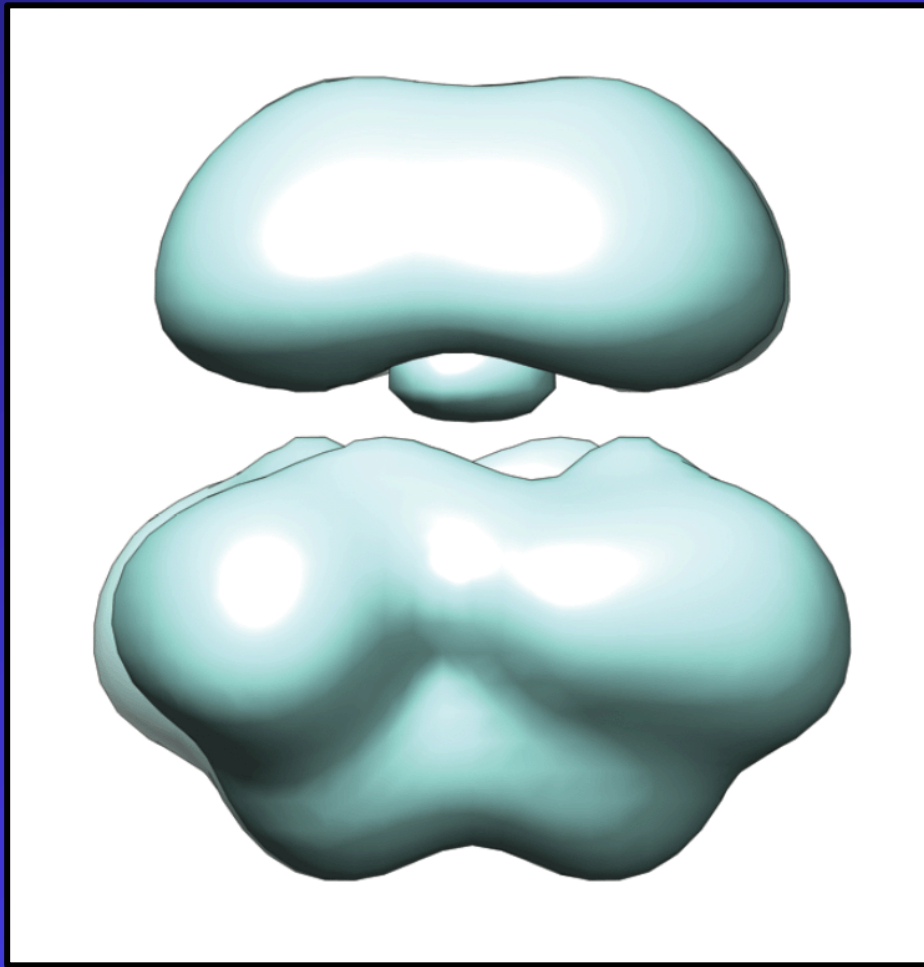


Serysheva *et al.*, 1995

# Chicken Slo2.2 in the absence of $\text{Na}^+$



Class averages



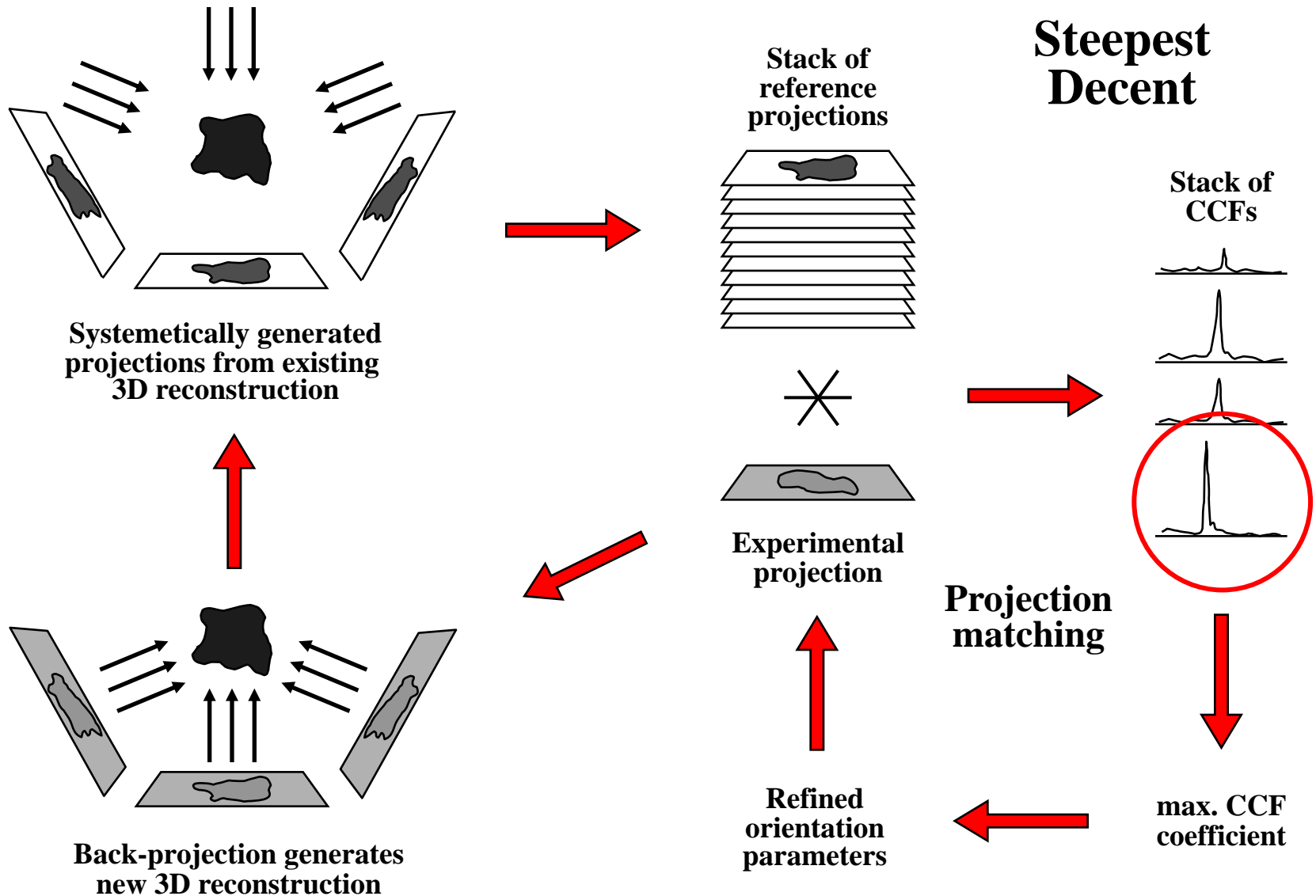
Initial model (obtained with VIPER)

VIPER

**Stochastic  
Hill Climbing**

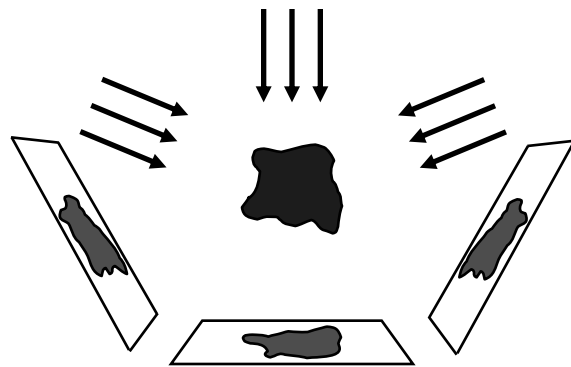
(initially introduced  
in program SIMPLE)

# Angular refinement





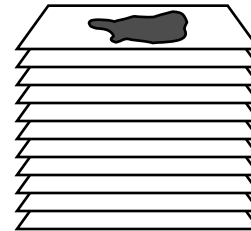
# Angular refinement



Systemetically generated  
projections from existing  
3D reconstruction



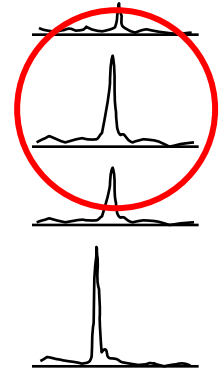
Stack of  
reference  
projections



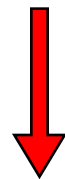
Experimental  
projection

## Stochastic Hill Climbing

Stack of  
CCFs



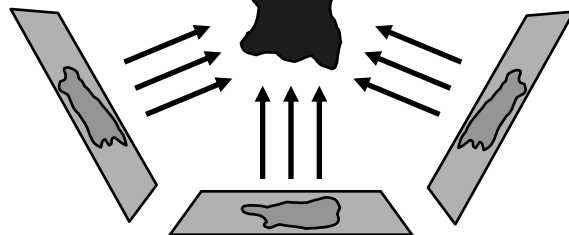
Projection  
matching



FIRST  
max. CCF  
coefficient

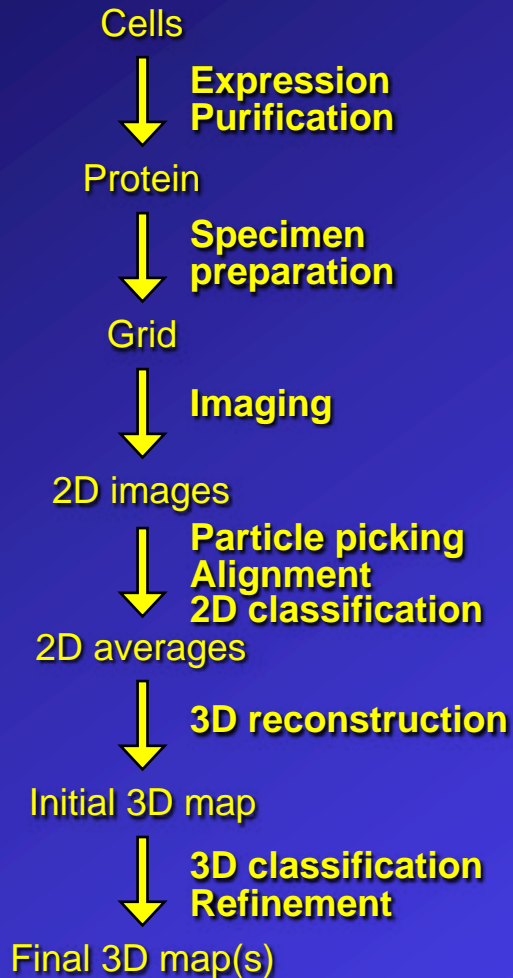


Refined  
orientation  
parameters



Back-projection generates  
new 3D reconstruction

# Structure determination by single-particle EM



## Potential issues:

Reference bias

Overfitting

Resolution assessment

# Structure determination by single-particle EM

## Potential issues with density map



Model/reference bias

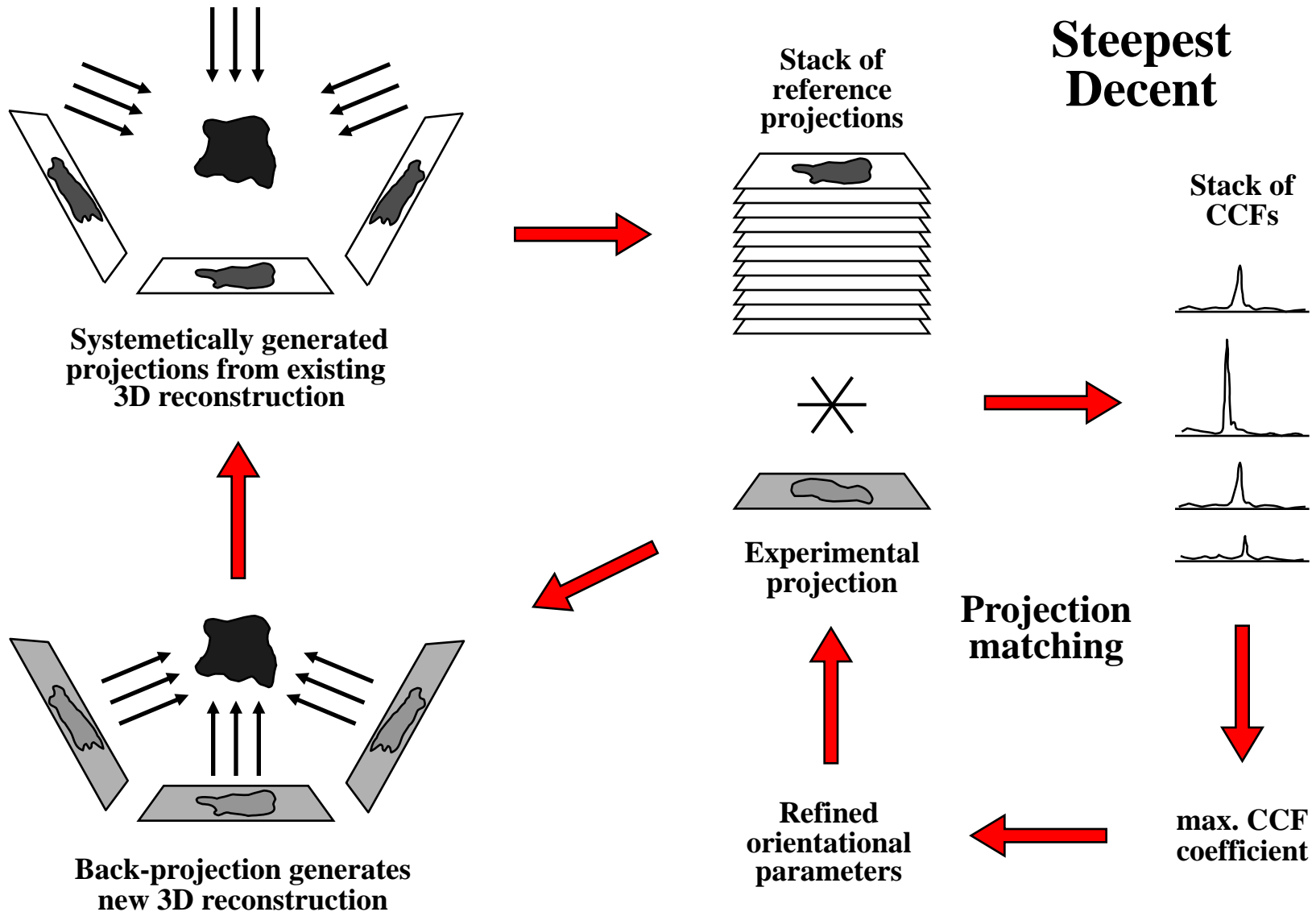
Average of 1,000 images containing pure white noise after alignment to an image of Albert Einstein

→ Einstein from noise

Shatsky *et al.* (2009) *J. Struct. Biol.* 166: 67-78

Henderson (2013) *Proc. Natl. Acad. Sci. USA* 110: 18037-18041

# Angular refinement





# Structure determination by single-particle EM

## Potential issues with density map



Model/reference bias

Average of 1,000 images containing pure white noise after alignment to an image of Albert Einstein

→ Einstein from noise

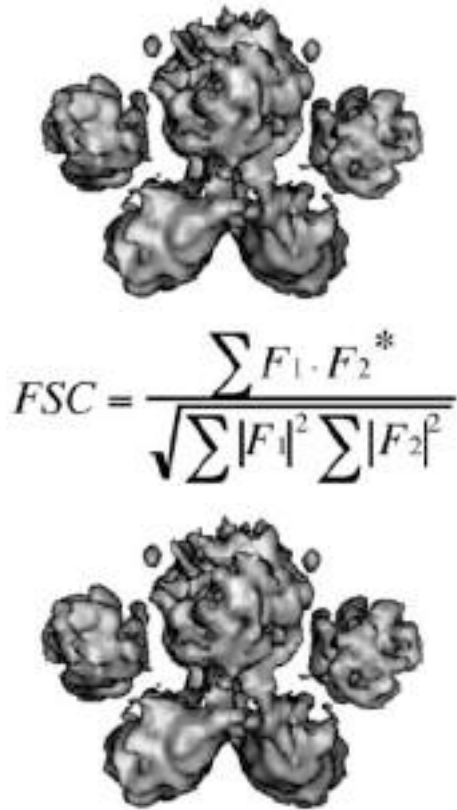
Over-fitting results in spurious high-resolution features due to alignment of noise

Shatsky *et al.* (2009) *J. Struct. Biol.* 166: 67-78

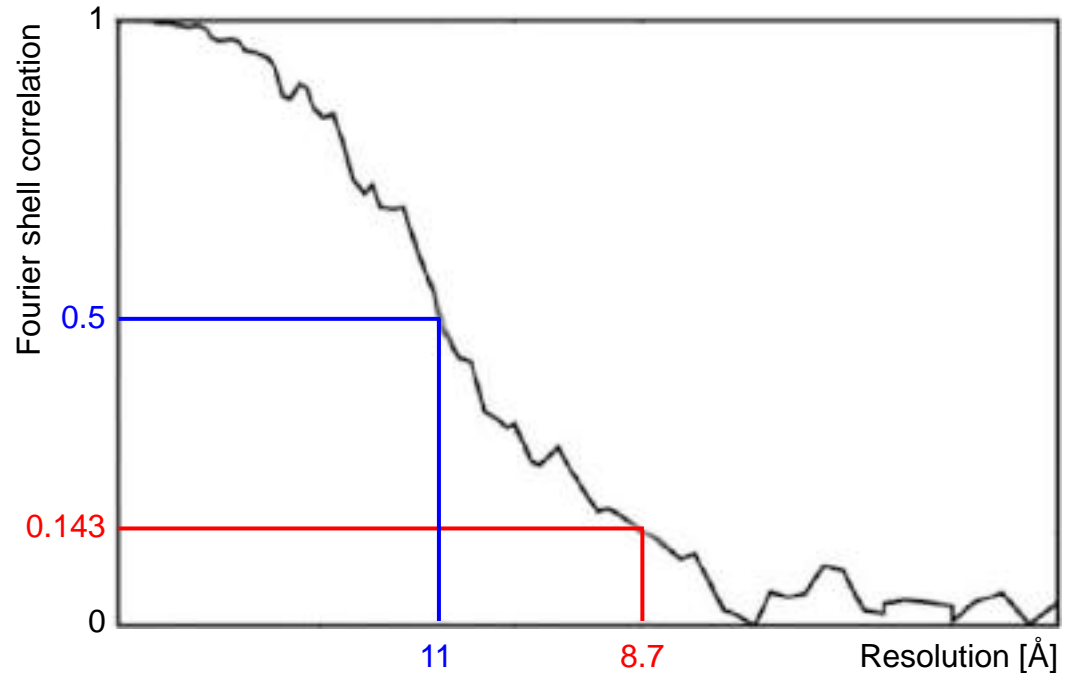
Henderson (2013) *Proc. Natl. Acad. Sci. USA* 110: 18037-18041

# Structure determination by single-particle EM

## Resolution assessment



Maps have to be independent !



FSC = 0.5      Signal = Noise

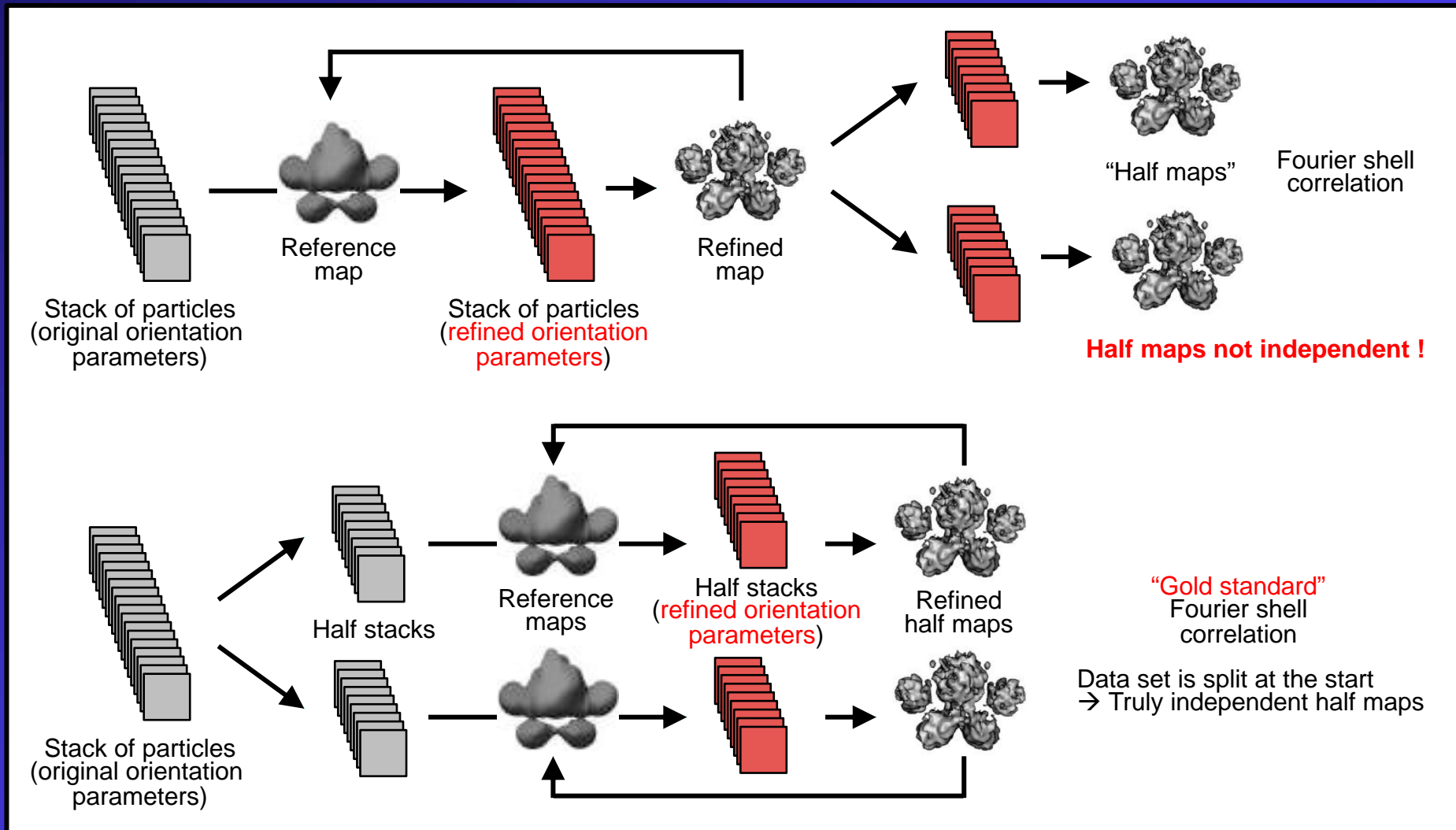
Böttcher *et al.* (1997) *Nature* 386: 88-91

FSC = 0.143      Phase error = 60°

Rosenthal & Henderson (2003) *J. Mol. Biol.* 333: 721-745

# Structure determination by single-particle EM

## Resolution assessment



# Structure determination by single-particle EM

## Resolution assessment

What should be resolved ?

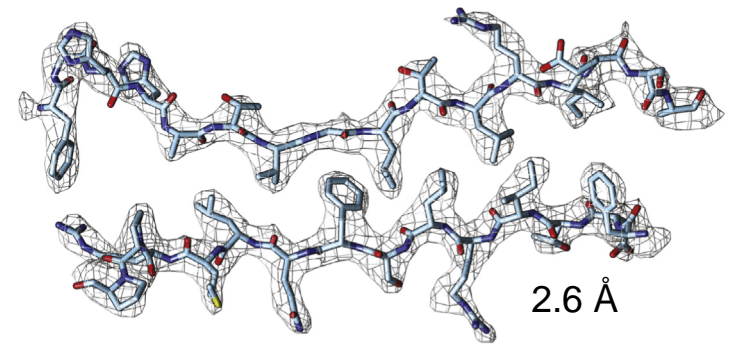
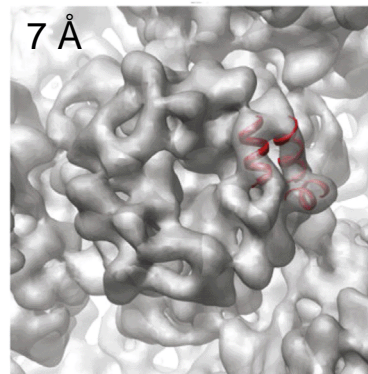
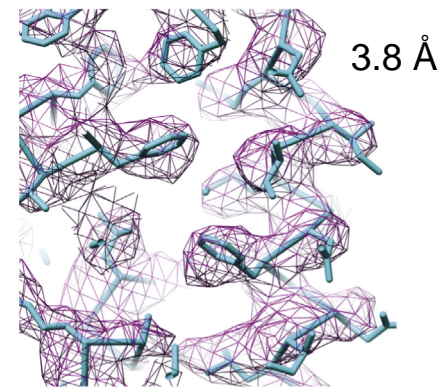
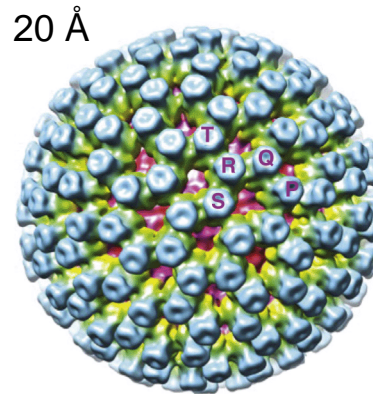
> 20 Å  
protein envelope

~ 9-10 Å  
 $\alpha$ -helices

< 4.8 Å  
 $\beta$ -sheets

~ 4 Å  
bulky side chains

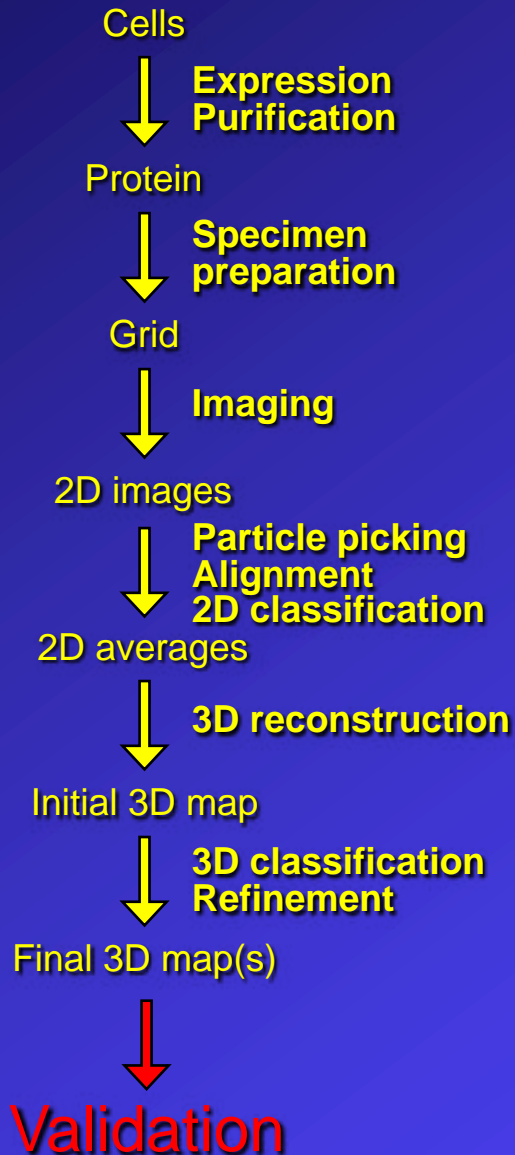
Rotavirus double-layered particle



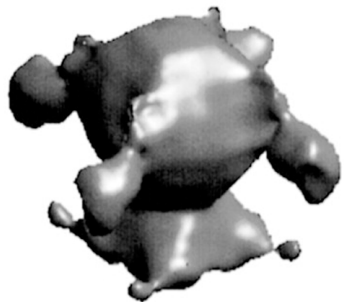
Rosenthal & Rubinstein (2015) *Curr. Opin. Struct. Biol.* 34: 135-144



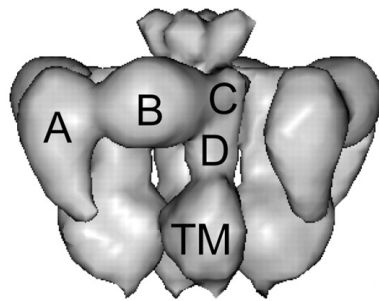
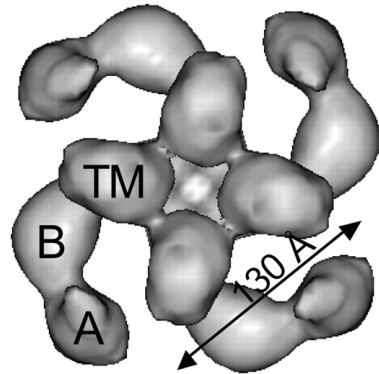
# Structure determination by single-particle EM



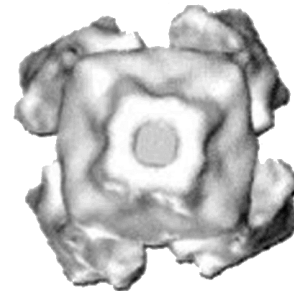
# The issue: Structures of the IP3 receptor as determined by single-particle EM



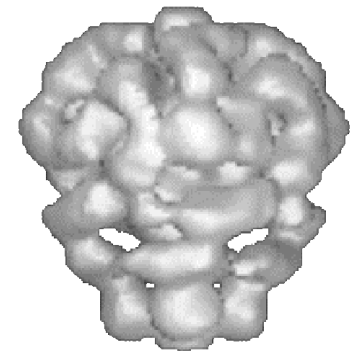
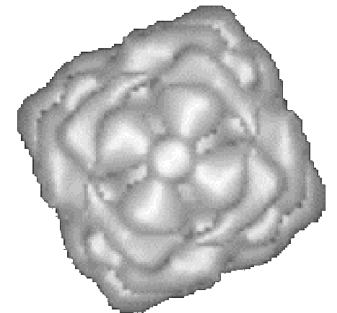
Jiang *et al.*,  
2002



Serysheva *et al.*,  
2003



Jiang *et al.*,  
2003



Sato *et al.*,  
2004

# Map validation

Meeting of experts in 2010 to come up with standards for map validation

Outcome summarized in 2012:

Structure  
Meeting Review



## Outcome of the First Electron Microscopy Validation Task Force Meeting

Richard Henderson,<sup>1</sup> Andrej Sali,<sup>2</sup> Matthew L. Baker,<sup>3</sup> Bridget Carragher,<sup>4</sup> Batsal Devkota,<sup>5</sup> Kenneth H. Downing,<sup>6</sup> Edward H. Egelman,<sup>7</sup> Zukang Feng,<sup>5</sup> Joachim Frank,<sup>8,9</sup> Nikolaus Grigorieff,<sup>10</sup> Wen Jiang,<sup>11</sup> Steven J. Ludtke,<sup>3</sup> Ohad Medalia,<sup>12,21</sup> Pawel A. Penczek,<sup>13</sup> Peter B. Rosenthal,<sup>14</sup> Michael G. Rossmann,<sup>15</sup> Michael F. Schmid,<sup>3</sup> Gunnar F. Schröder,<sup>16</sup> Alasdair C. Steven,<sup>17</sup> David L. Stokes,<sup>18</sup> John D. Westbrook,<sup>5</sup> Willy Wriggers,<sup>19</sup> Huanwang Yang,<sup>5</sup> Jasmine Young,<sup>5</sup> Helen M. Berman,<sup>5</sup> Wah Chiu,<sup>3</sup> Gerard J. Kleywegt,<sup>20</sup> and Catherine L. Lawson<sup>5,\*</sup>

Henderson *et al.* (2012) *Structure* 20: 205-214

# Map validation

- Compare reference-free averages with projections

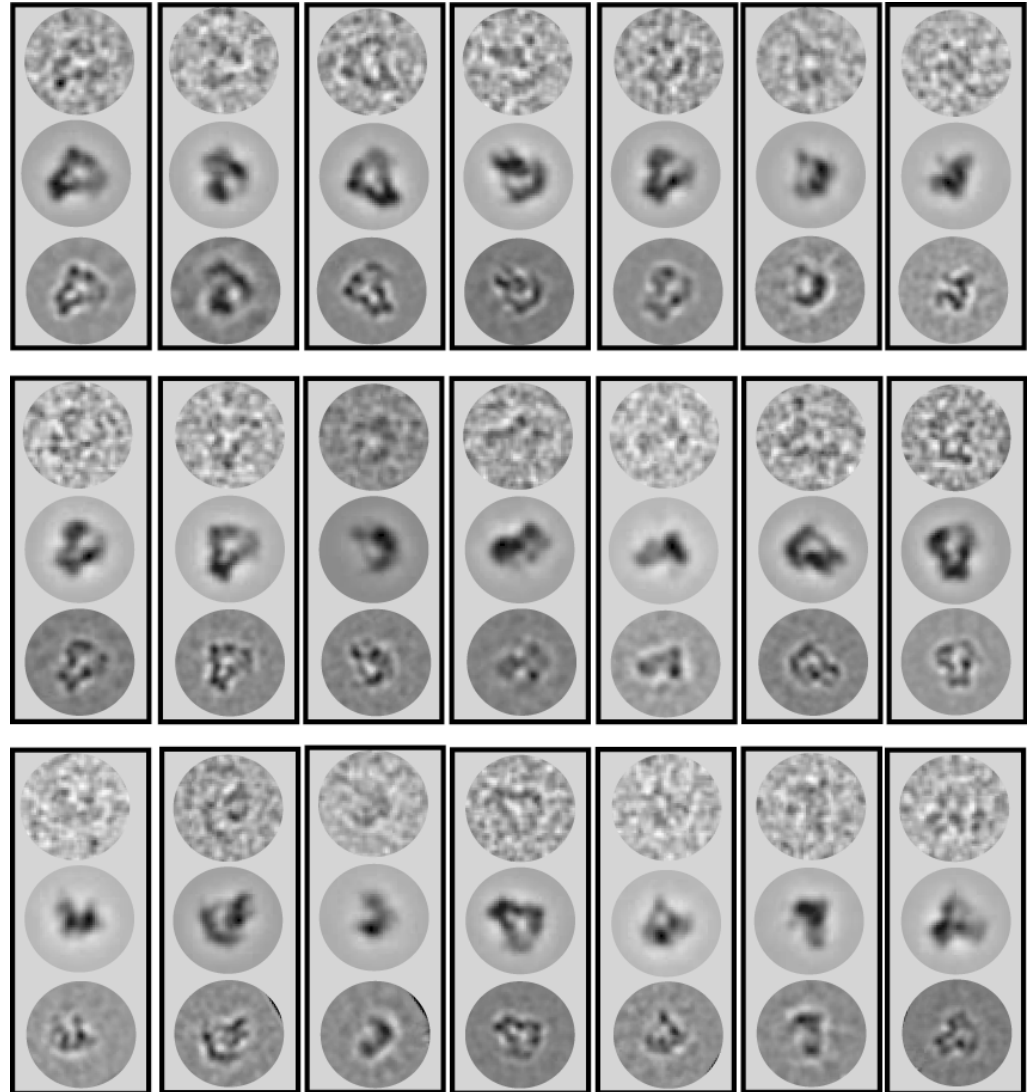
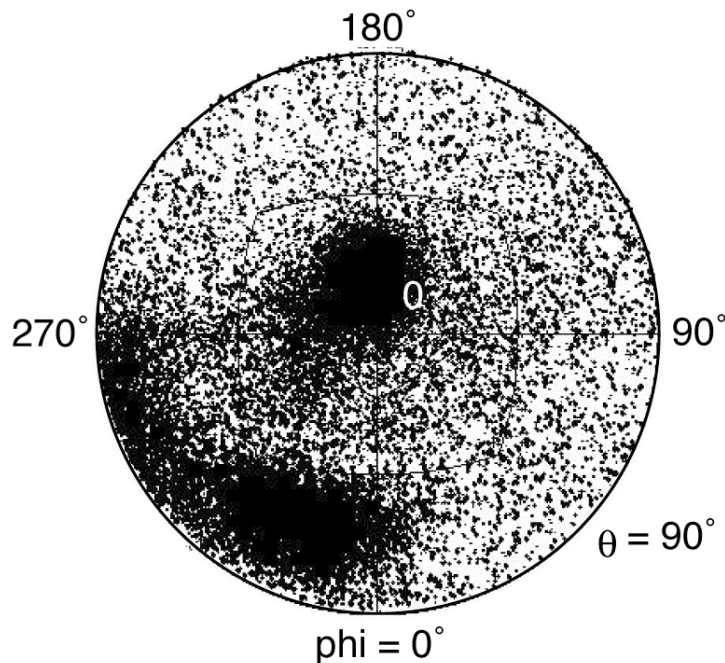
Henderson *et al.* (2012) *Structure* 20: 205-214



# Map validation

## Re-projections and angular distribution

Anaphase  
promoting  
complex



# Map validation

- Compare reference-free averages with projections
  - only checks consistency of 3D map with 2D data
  - also check angle distribution
- Tilt-pair analysis

Henderson *et al.* (2012) *Structure* 20: 205-214

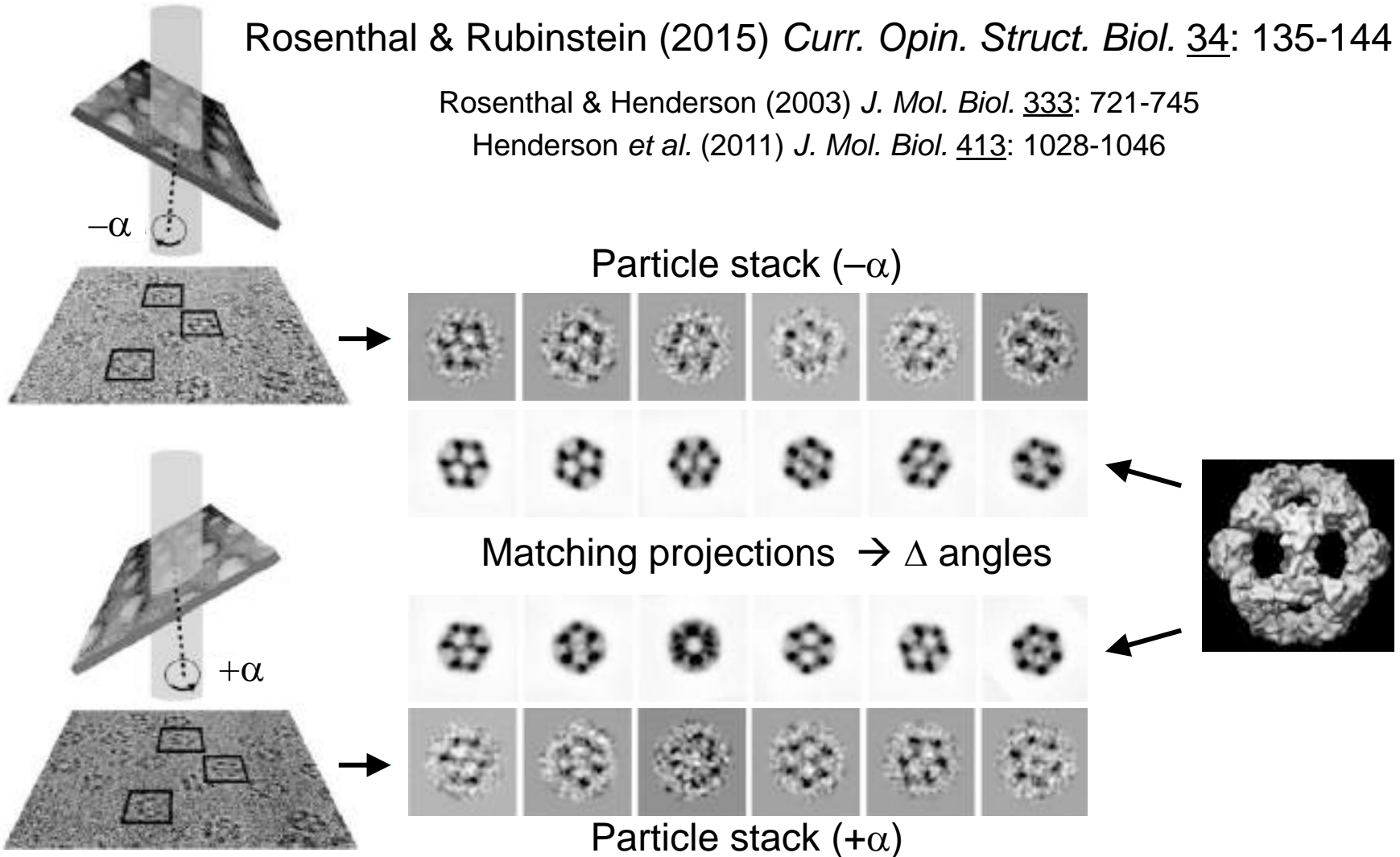
# Map validation

## Tilt-pair analysis

Rosenthal & Rubinstein (2015) *Curr. Opin. Struct. Biol.* 34: 135-144

Rosenthal & Henderson (2003) *J. Mol. Biol.* 333: 721-745

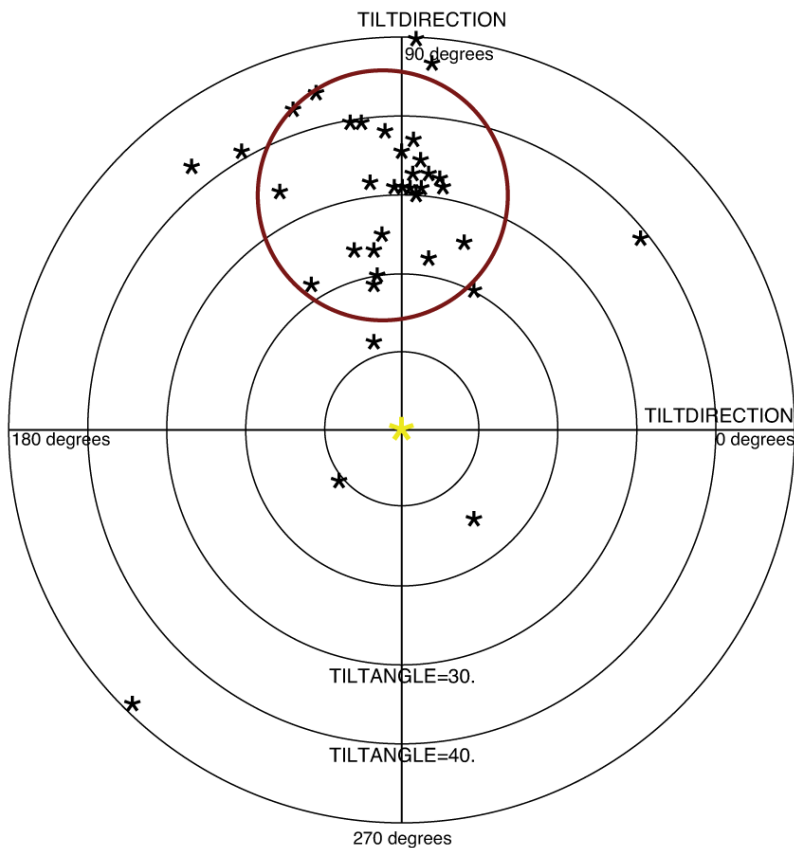
Henderson *et al.* (2011) *J. Mol. Biol.* 413: 1028-1046



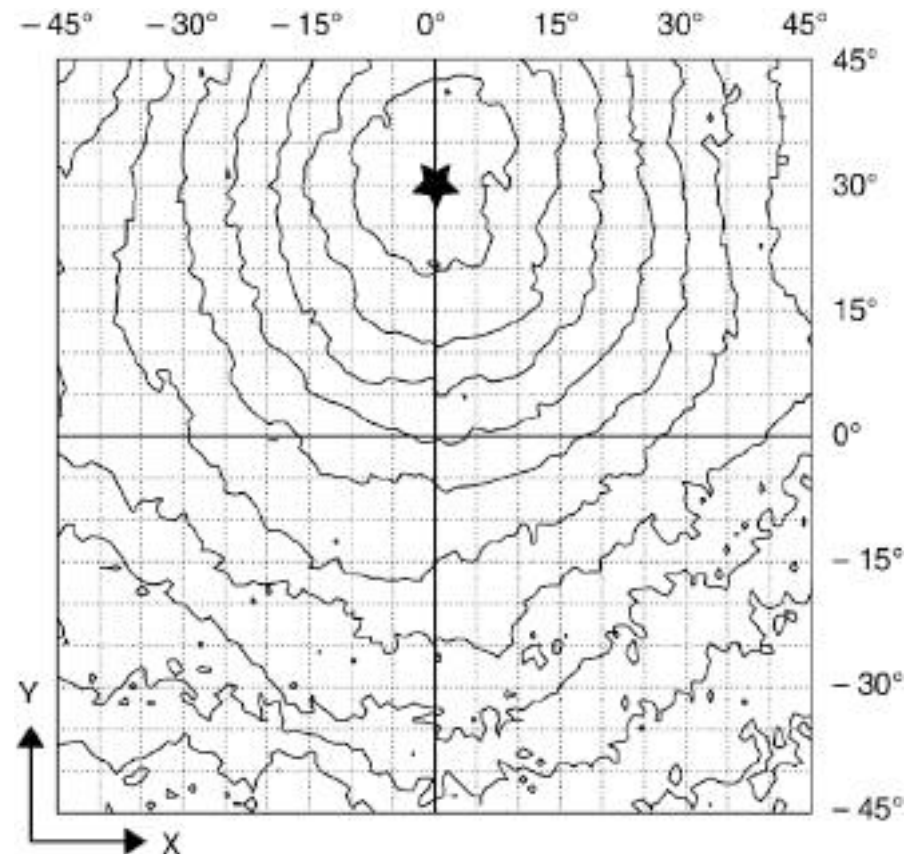
# Map validation

## Tilt-pair analysis

Tilt-pair parameter plot



Tilt-pair phase residual plot





# Map validation

## Tilt-pair analysis

Henderson *et al.* (2011) *J. Mol. Biol.* 413: 1028-1046

**Table 1.** Overview of tilt-pair statistics

Specimen	Symmetry	Particle size (Å)	Molecular mass (MDa)	Number of tilt pairs	Number of particles	Successful alignment (%)	Angular error (°)	
							Mean	Maximum
Rotavirus DLP	I2	700	50	10	95	100/100	0.25	1.0
CAV	I2	255	2.7	1	45	62/82	2.5	3.5
70S ribosomes	C1	270 × 260	2.6	12	220	45/75	4.0	5.0
FAS	D3	260 × 220	2.6	2	44	59/95	4.0	6.0
PDH-E2CD	I1	280	1.6	1	50	62/94	3.0	4.0
<i>Thermus</i> V-ATPase	C1	250 × 140	0.6	1	50	54/80	10.0	16.0
Bovine F-ATPase	C1	250 × 140	0.6	1	29	52/79	20.0	25.0
DNA-PKcs	C1	150 × 120	0.47	14	108	44/81	15.0	17.0
β-Galactosidase	D2	180 × 130 × 95	0.45	2	119	74/91	10.0	14.0

- determines whether overall 3D map is correct at 15-20 Å resolution (but not high-resolution features)
- allows determination of handedness
- can be used to refine parameters used for orientation determination  
→ can thus be used to improve the map
- validates orientation parameters (but not microscope parameters, i.e., defocus, magnification)

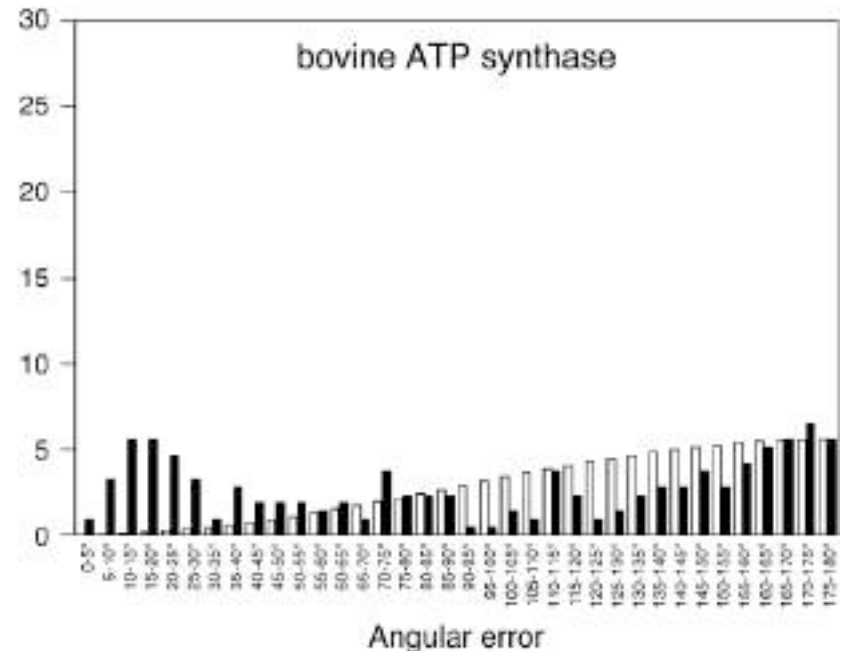
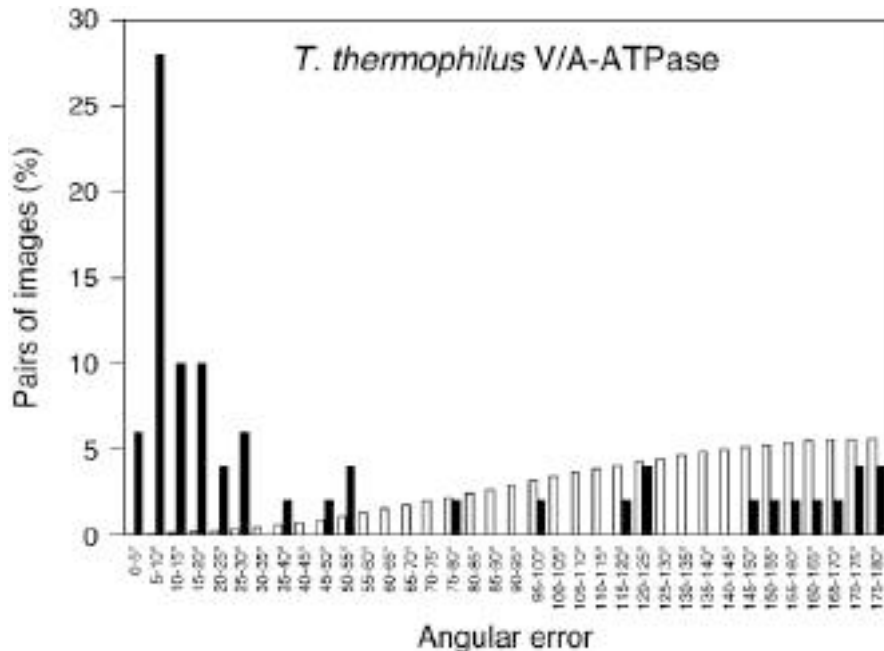
“If less than 60% of particles show a single cluster, the basis for poor orientation parameters should be investigated”

# Map validation

## Tilt-pair analysis

### Tilt-pair alignment test

- angular errors for determination of the tilt transformation of each particle pair
- expected for random orientations



Rosenthal & Rubinstein (2015) *Curr. Opin. Struct. Biol.* 34: 135-144

Baker et al. (2012) *Proc. Natl Acad. Sci. USA* 109: 11675-11680

Russo & Passmore (2014) *J. Struct. Biol.* 187: 112-118

# Map validation

## Tilt-pair web server

Input

[Load previous settings](#)

### Micrograph Parameters

Magnification 4.98 A/px	Defocus 1 58626.0 A	Tilt search range 20.0 degrees
Voltage 300.0 kV	Defocus 2 59084.0 A	Min resolution 100.0 A
Particle size (radius) 30.0 pixels	Angle astigmatism 55.7 degrees	Max resolution 30.0 A

☐ Disable CTF correction

Note: Positive defocus values equate underfocus.

### Input data

3D model  
[Browse...](#) e2map.mrc

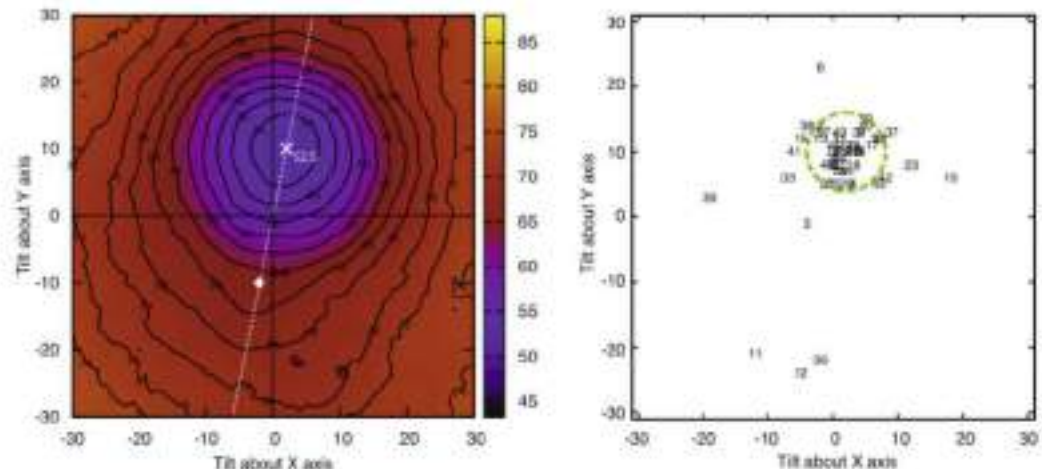
Stack 1 (max 300 particles)  
[Browse...](#) stack1.mrc

Stack 2 (max 300 particles)  
[Browse...](#) stack2.mrc

Parameters for the Stack 1  
[Browse...](#) stack1.par

Parameters format  
[Freeform](#)

Output



### Parameters:

Magnification	4.98 (effective: 9.96)
A/px	
Defocus	58626   59084
Astigmatism	55.7
Voltage	300 kV
Resolution Range	100.0 - 30.0 A
Tilt Range	20
Particle radius	20 (effective: 10) px
Optimized box size (after binning)	48
Effective binning	2

### Summary of the results for all submitted particles:

Minimal Phase Residual: **52.53°**  
 Minimum at the position: **2.0°, 10.0°**  
 Tilt axis (angle with respect to the X axis): **78.7°**  
 Tilt angle: **16.2°**  
 Rand Phase Difference: **12.48°**  
 Average distance to the global minimum: **5.24°**

Particles in the cluster (0.50 - **4.13°**) near the minimum average phase residual:  
 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27  
 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50

Particles outside the cluster:  
 2 6 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50

# Map validation

<http://www.ebi.ac.uk/pdbe/emdb/validation/tiltpair/>

Tilt-pair validation server

www.ebi.ac.uk/pdbe/emdb/validation/tiltpair/

This website uses cookies. By continuing to browse this site, you are agreeing to the use of our site cookies. To find out more, see our [Terms of Use](#).

EMBL-EBI Protein Data Bank in Europe Bringing Structure to Biology

EM resources

Services Research Training About us

Share Feedback

### EM Resources

- Home
- Statistics
- Validation
- EMDataBank
- EMPIAR
- Test data

### EMDB

- Latest maps
- Latest headers
- Latest updates
- Search
- Browse
- FTP archive
- Deposit EM map/model
- EMDB data model

### Tilt pair validation server

Welcome to the PDBe tilt pair validation server!

Tilt-pair validation analysis (Rosenthal and Henderson, 2003) can be used to assess the accuracy of initial angle assignment in single-particle processing. To perform this analysis you need to collect two corresponding sets of particle images - one untitled and the other tilted, then upload the stacks of images along with a 3D reconstruction based on the untitled images. This server is based on the [Tilt-pair server](#) developed at MRC National Institute for Medical Research (Wasilewski and Rosenthal, 2014), and we thank Sebastian Wasilewski and Peter Rosenthal for their help in developing and testing the current server.

You may upload map files in MRC or CCP4 format, and parameter files (containing Euler angles for individual particles) in Spider or Frealign format. We have some test data sets that you can use to try out the service [here](#). We are still developing the server and appreciate your [feedback](#)!

Map (3D volume):  no file selected

Untitled stack:  no file selected

Orientation parameters for stack 1:  no file selected

Tilted stack:  no file selected

Pixel size (Å):

Mask radius (pixels):

Tilt search range (degrees):

Resolution range (low to high; Å):

Email address:

Job name:

Perform CTF correction? ☐

EMDataBank

PDBe is a member of

EMBL-EBI

News  
Brochures  
Contact us  
Intranet

Services

By topic  
By name (A-Z)  
Help & Support

Research

Overview  
Publications  
Research groups  
Postdocs & PhDs

Training

Overview  
Train at EBI  
Train outside EBI  
Train online

Industry

Overview  
Members Area  
Workshops  
SME Forum

About us

Overview  
Leadership  
Funding  
Background



# Map validation

- Compare reference-free averages with projections
  - only checks consistency of 3D map with 2D data
  - also check angle distribution
- Tilt-pair analysis
  - excellent, also establishes handedness
- “Gold standard” FSC
  - not necessarily needed (but certainly not bad)
- Randomize phases

Henderson *et al.* (2012) *Structure* 20: 205-214

# Map validation

## Randomize phases

Rosenthal & Rubinstein (2015) *Curr. Opin. Struct. Biol.* 34: 135-144

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

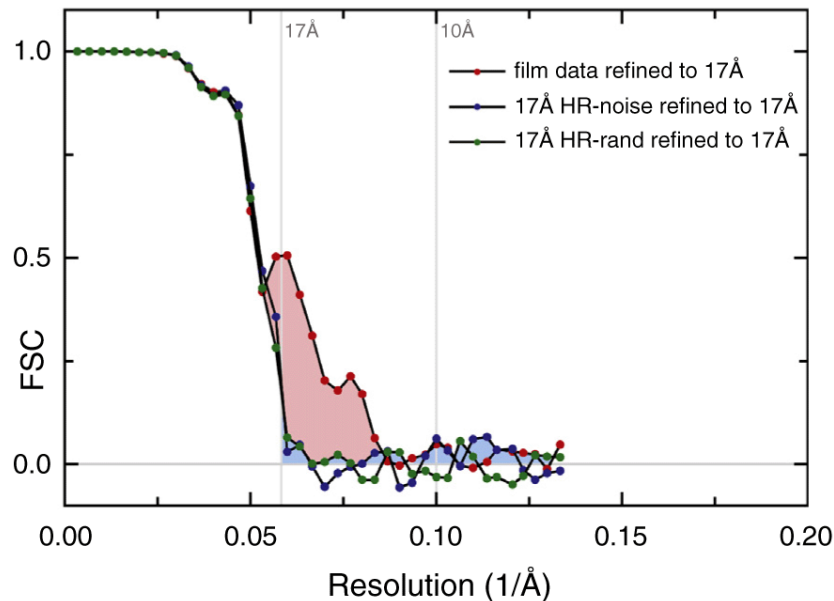
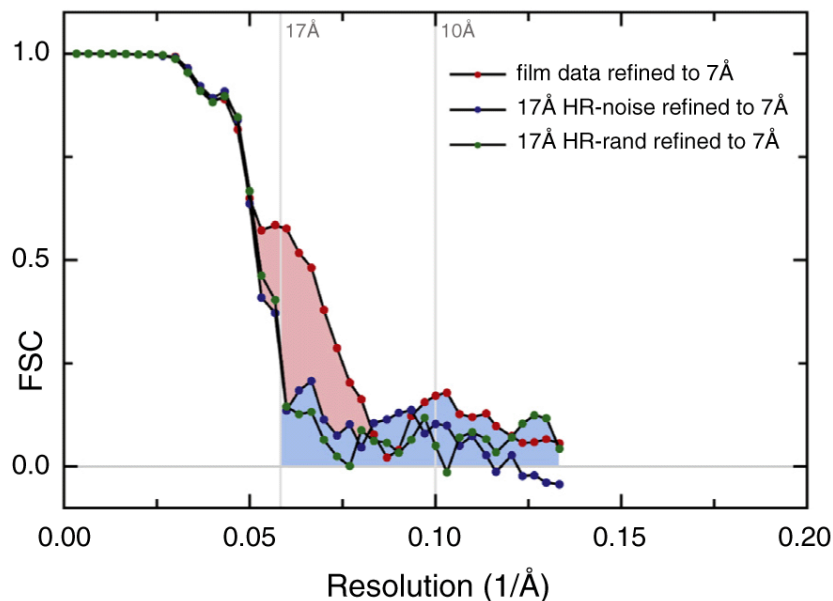
- Do single-particle reconstruction / refinement
- Determine resolution (FSC)
- Take raw data, randomize phases beyond which  $FSC_T$  falls below a threshold (75 or 80%)
- Redo the same analysis and recalculate FSC curve
- Any signal in region of randomized phases indicates issues with noise alignment in that region
- Can be implemented in any package

# Map validation

## Randomize phases

Rosenthal & Rubinstein (2015) *Curr. Opin. Struct. Biol.* 34: 135-144

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



- FSC signal due to over-fitting (noise)
- FSC signal due to true structural information

# Map validation

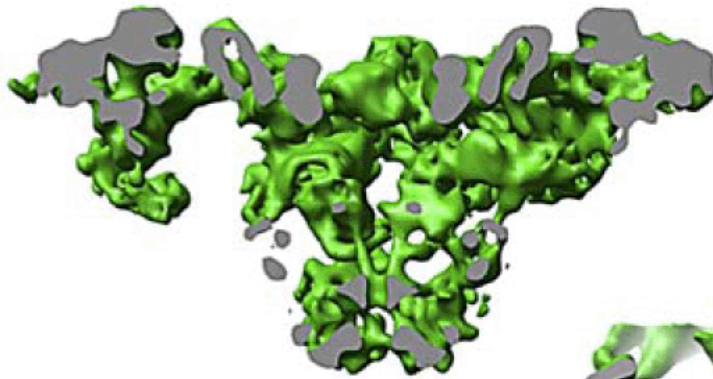
- Compare reference-free averages with projections
  - only checks consistency of 3D map with 2D data
  - also check angle distribution
- Tilt-pair analysis
  - excellent, also establishes handedness
- “Gold standard” FSC
  - not necessarily (but certainly not bad)
- Randomize phases
  - excellent, but not commonly used
- Appearance of expected secondary structure elements

Henderson *et al.* (2012) *Structure* 20: 205-214

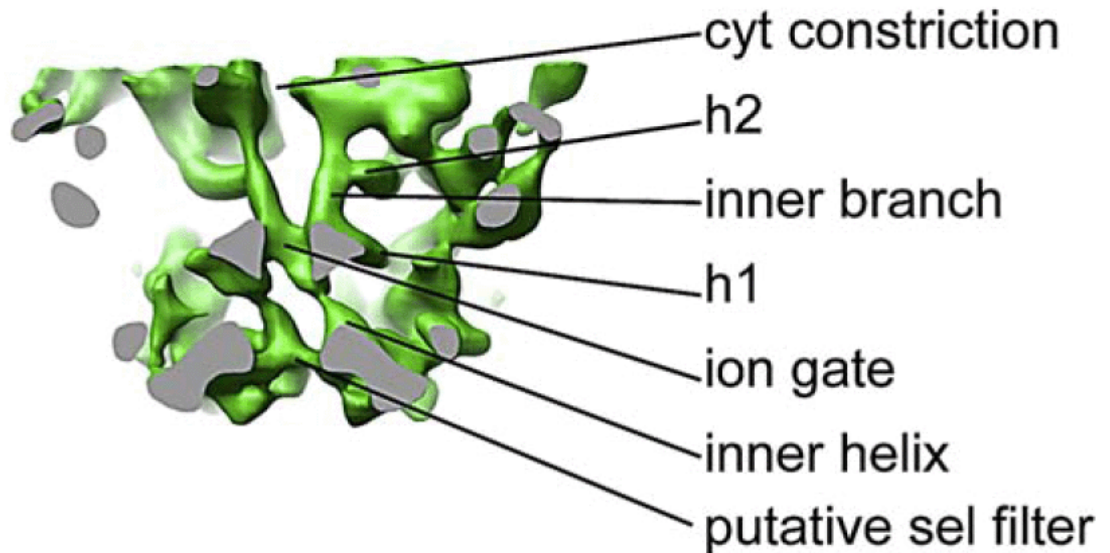


# Map validation

## Expected secondary structure



Ryanodine receptor 1  
at 10.2 Å resolution



Samso *et al.* (2009) *PLoS Biol.* 7: e1000085

# Map validation

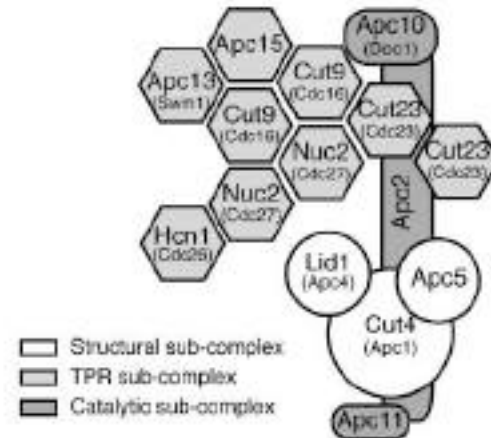
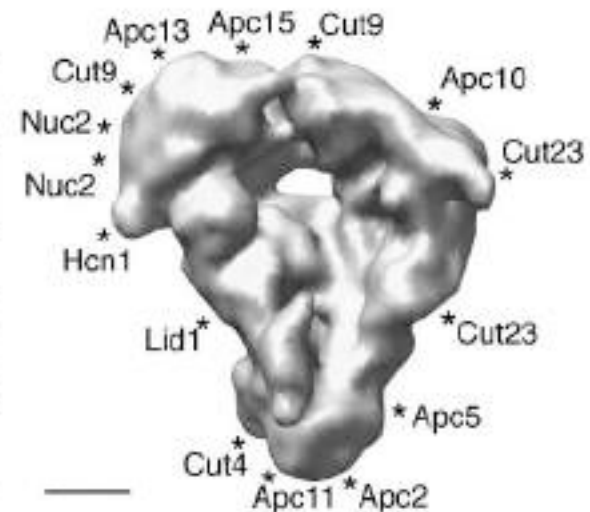
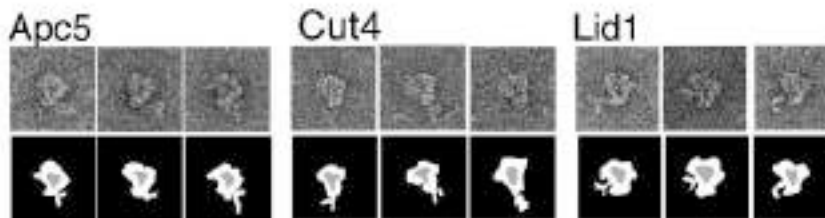
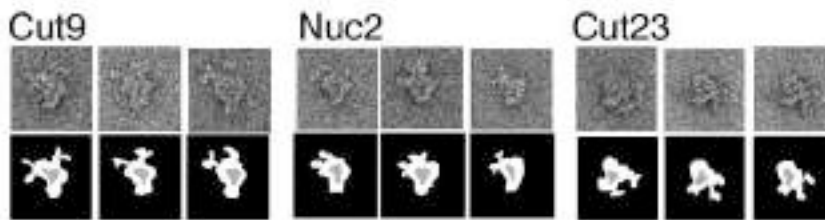
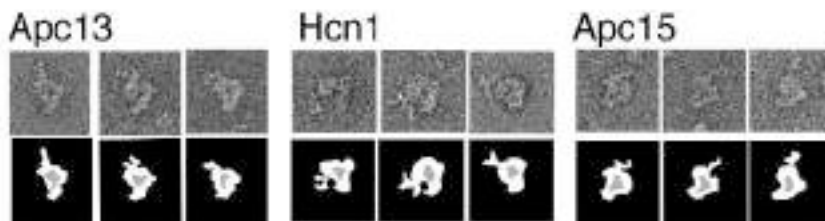
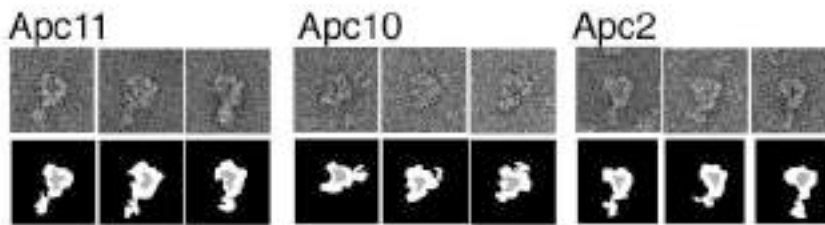
- Compare reference-free averages with projections
  - only checks consistency of 3D map with 2D data
  - also check angle distribution
- Tilt-pair analysis
  - excellent, also establishes handedness
- “Gold standard” FSC
  - not necessarily (but certainly not bad)
- Randomize phases
  - excellent, but not commonly used
- Appearance of expected secondary structure elements
- Evaluate with published information

Henderson *et al.* (2012) *Structure* 20: 205-214

# Map validation

## Evaluation with published information

### Anaphase promoting complex



# Map validation

- Compare reference-free averages with projections
  - only checks consistency of 3D map with 2D data
  - also check angle distribution
- Tilt-pair analysis
  - excellent, also establishes handedness
- “Gold standard” FSC
  - not necessarily (but certainly not bad)
- Randomize phases
  - excellent, but not commonly used
- Appearance of expected secondary structure elements
- Evaluate with published information
  - yeast two-hybrid analysis
  - pull-down experiments
  - cross-link mass spectrometry
- Dock known atomic structures into map

Henderson *et al.* (2012) *Structure* 20: 205-214

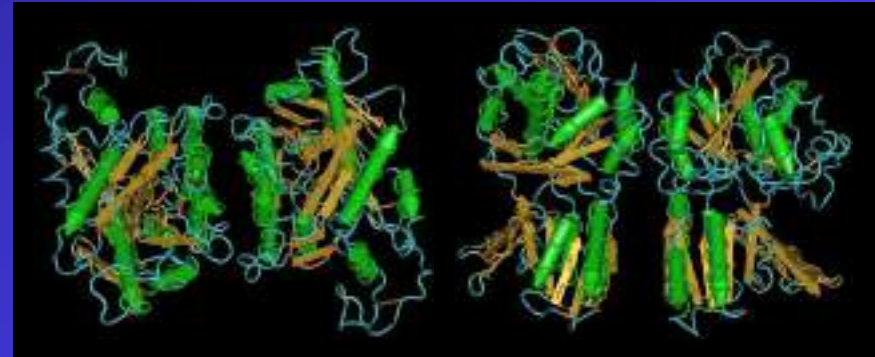


# Map validation

## Docking of atomic models

### mGluR1

Kunishima et al. 2000  
(K. Morikawa)



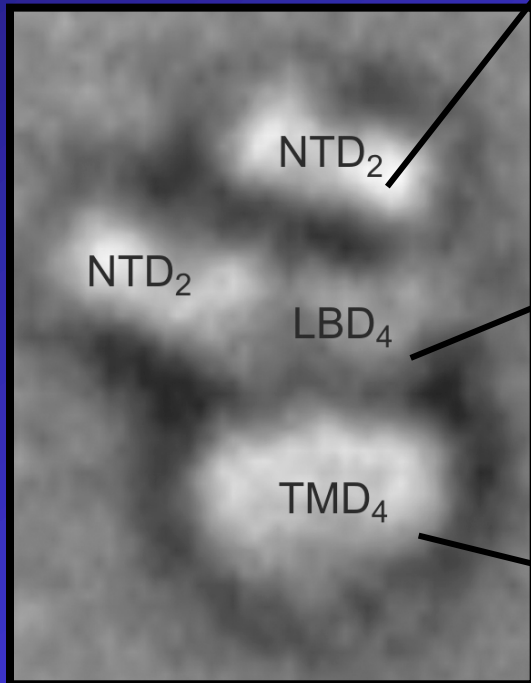
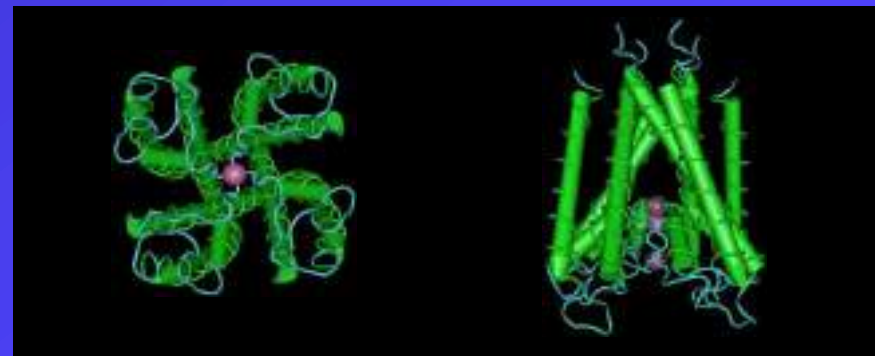
### GluR2

Armstrong et al. 2000  
(E. Gouaux)



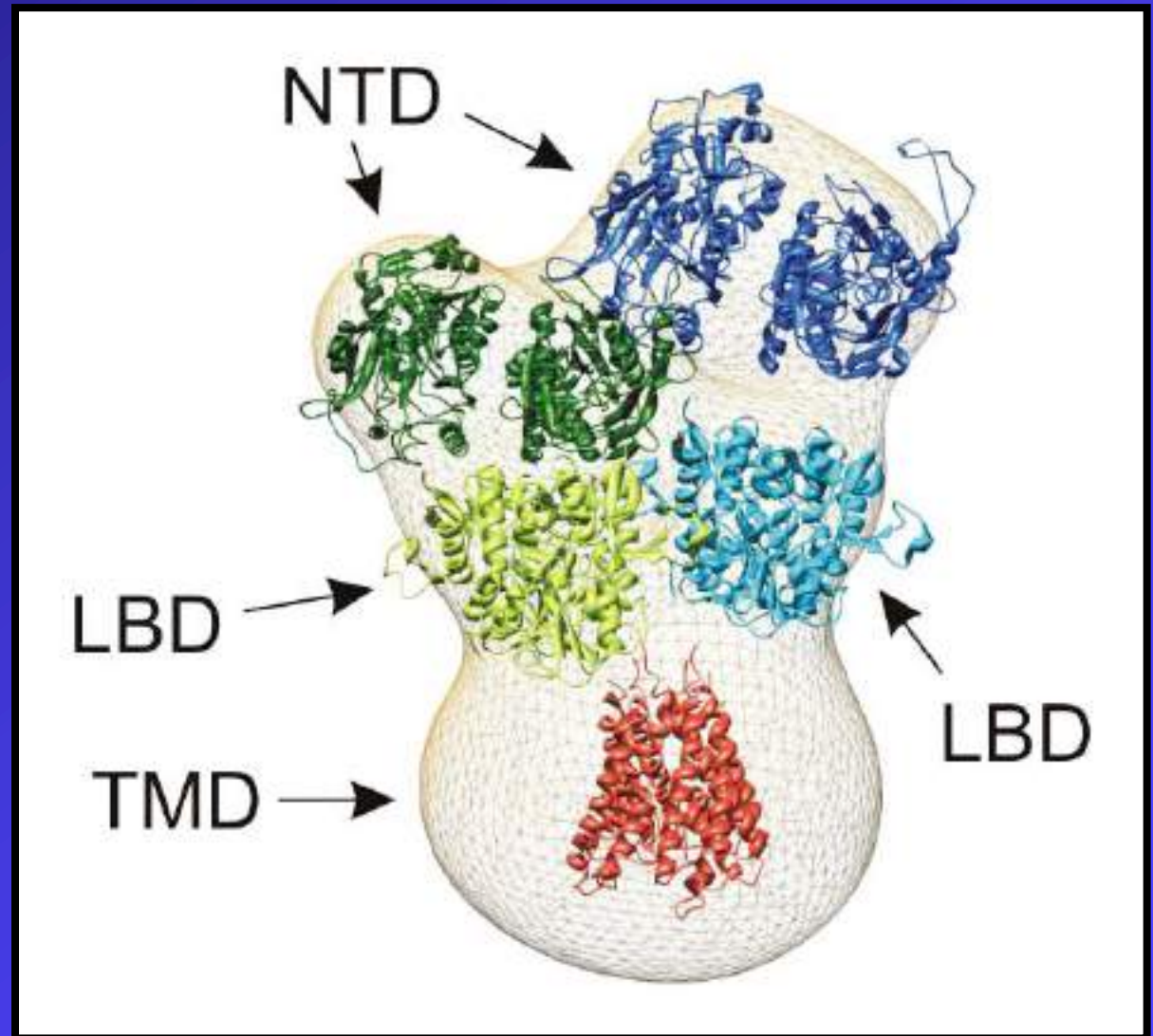
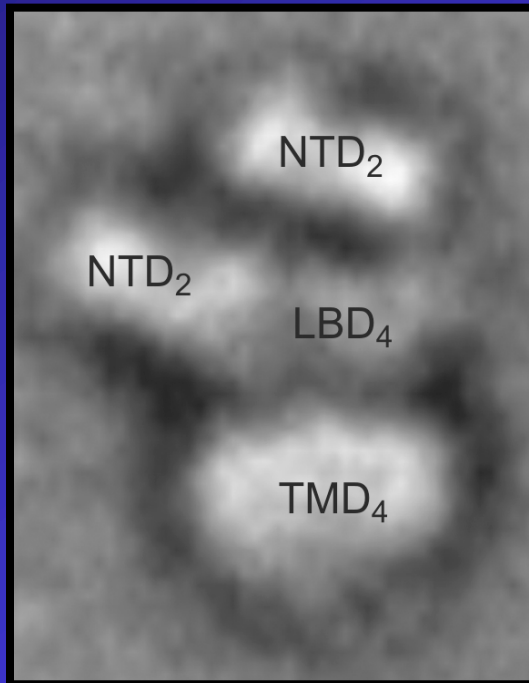
### KcsA

Doyle et al. 1998  
(R. MacKinnon)



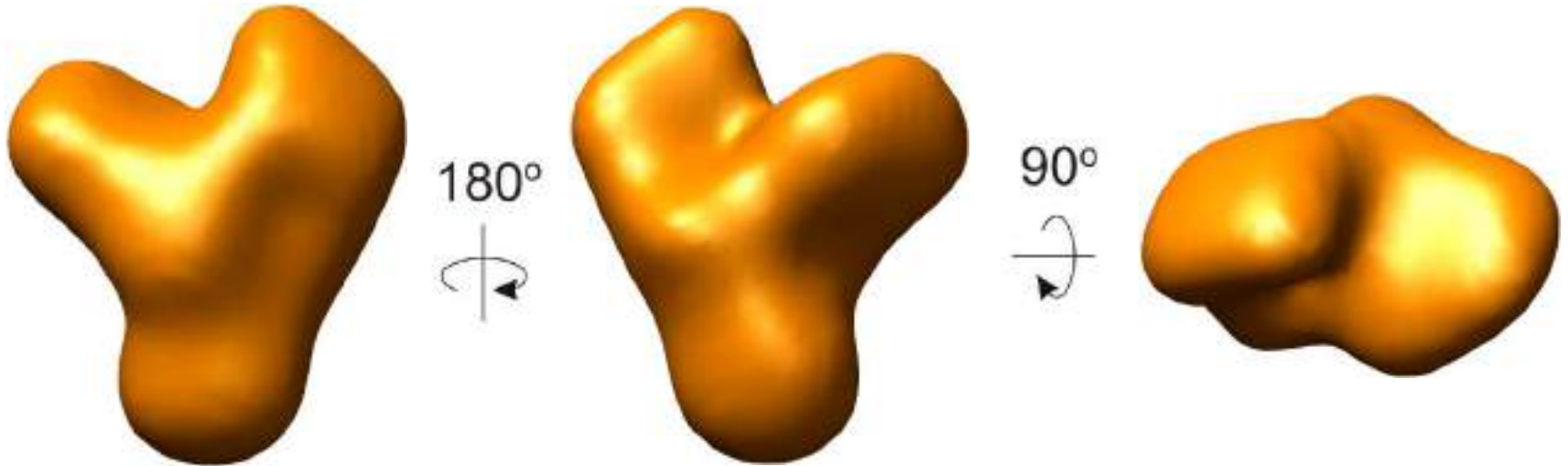
# Map validation

## Docking of atomic models



# Map validation

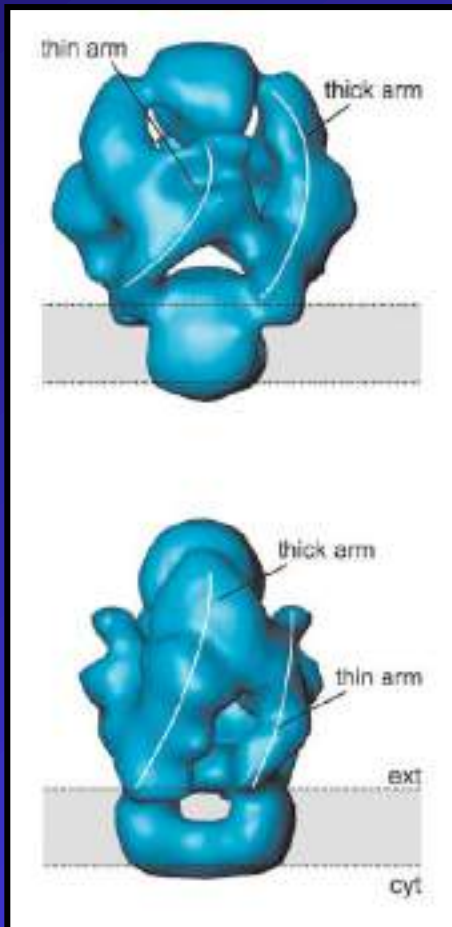
## Docking of atomic models



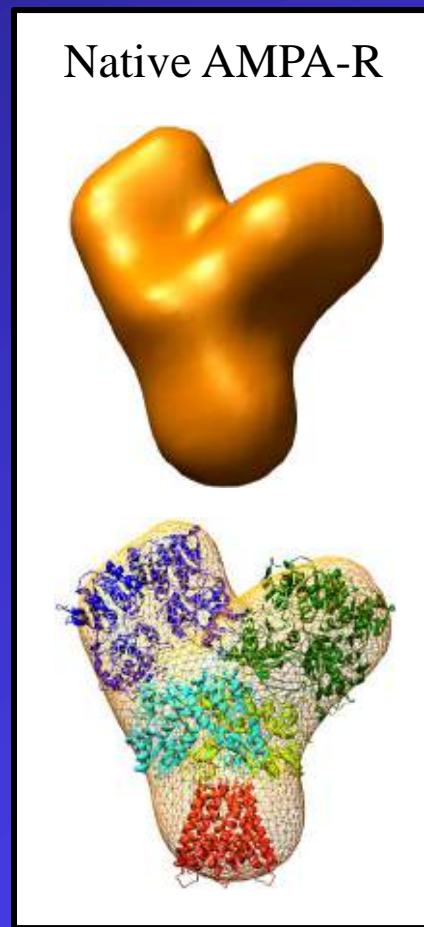


# Map validation

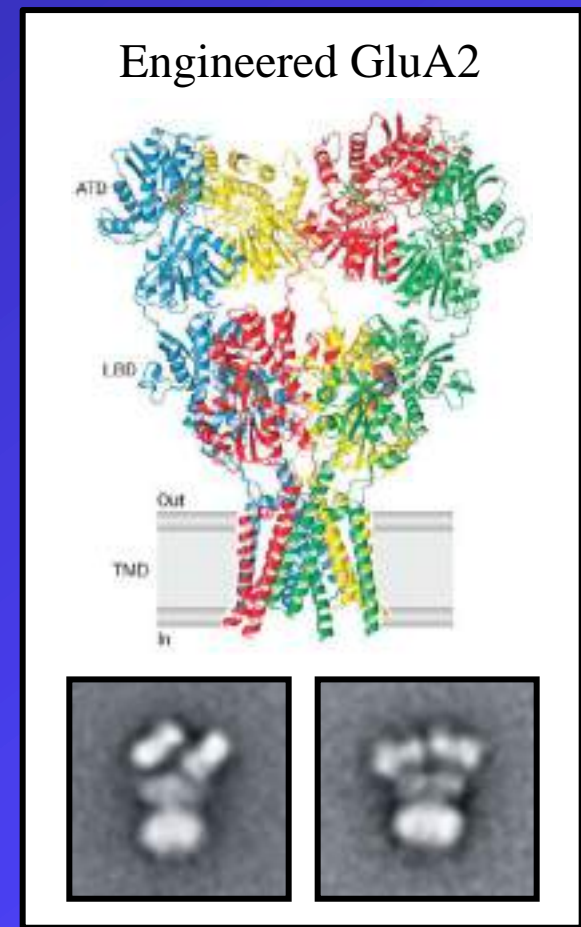
## Docking of atomic models



Tichelaar *et al.* (2004)  
*JMB* 344: 435-442



Nakagawa *et al.* (2006)  
*Biol. Chem.* 387: 179-187

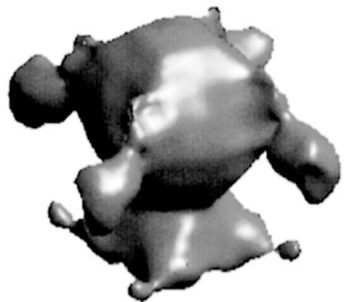


Sobolevsky *et al.* (2009)  
*Nature* 462: 745-758

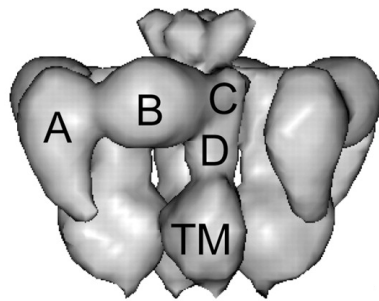
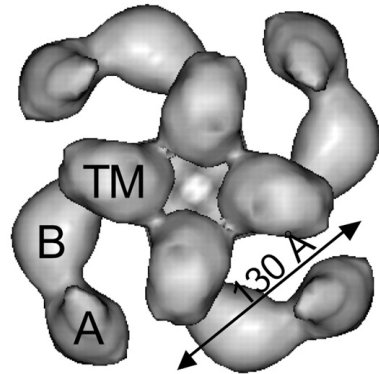


# Map validation - IP3 receptor

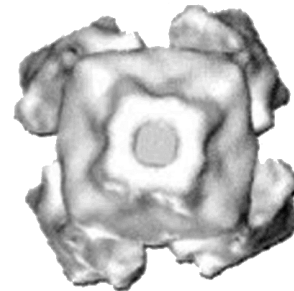
## Different maps of the IP3 receptor



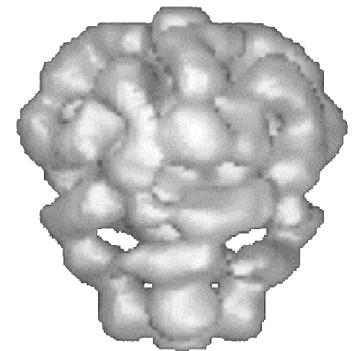
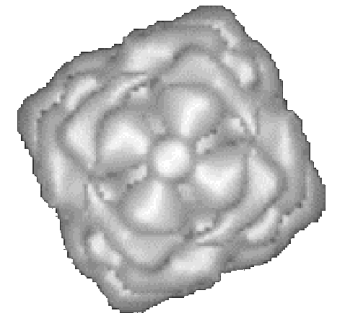
Jiang *et al.*,  
2002



Serysheva *et al.*,  
2003



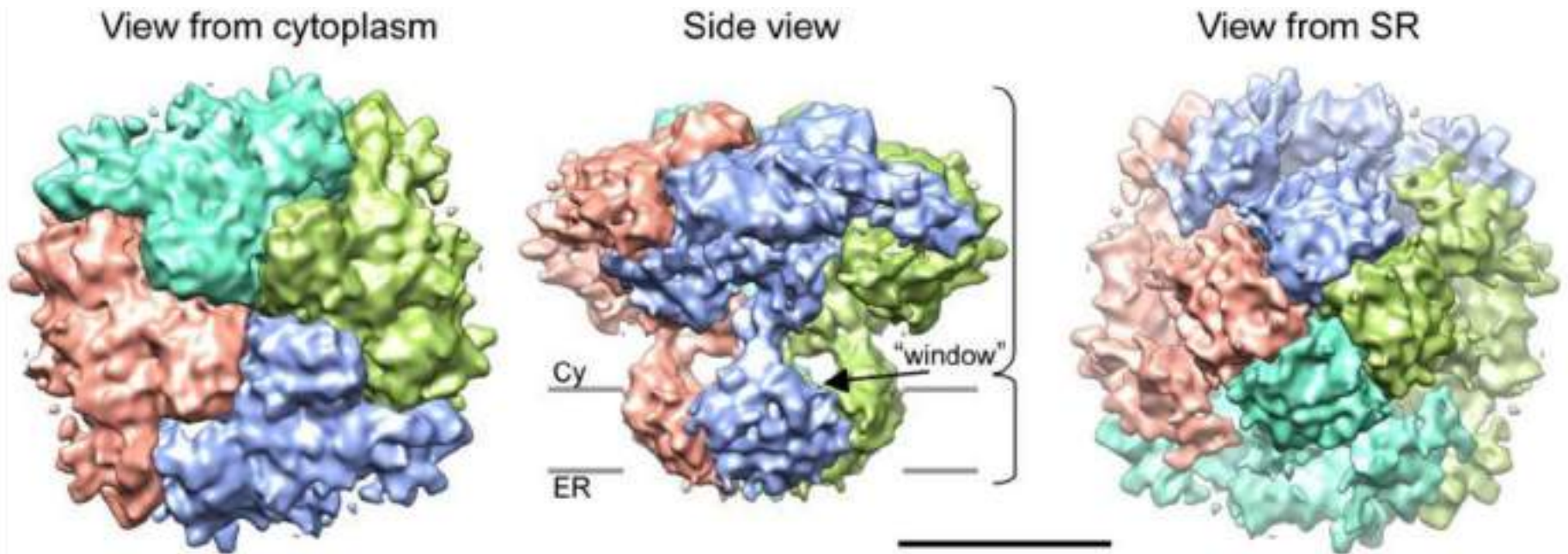
Jiang *et al.*,  
2003



Sato *et al.*,  
2004

# Map validation - IP3 receptor

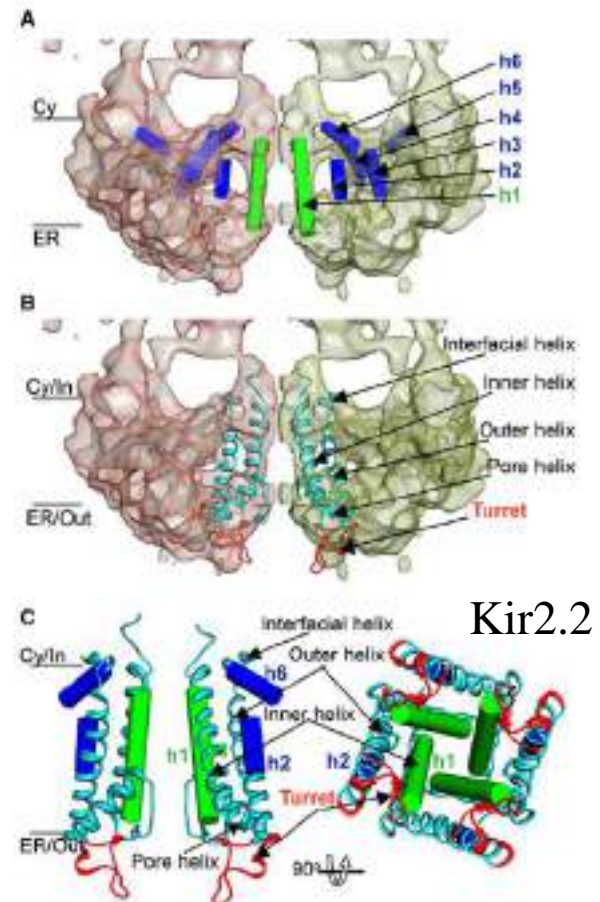
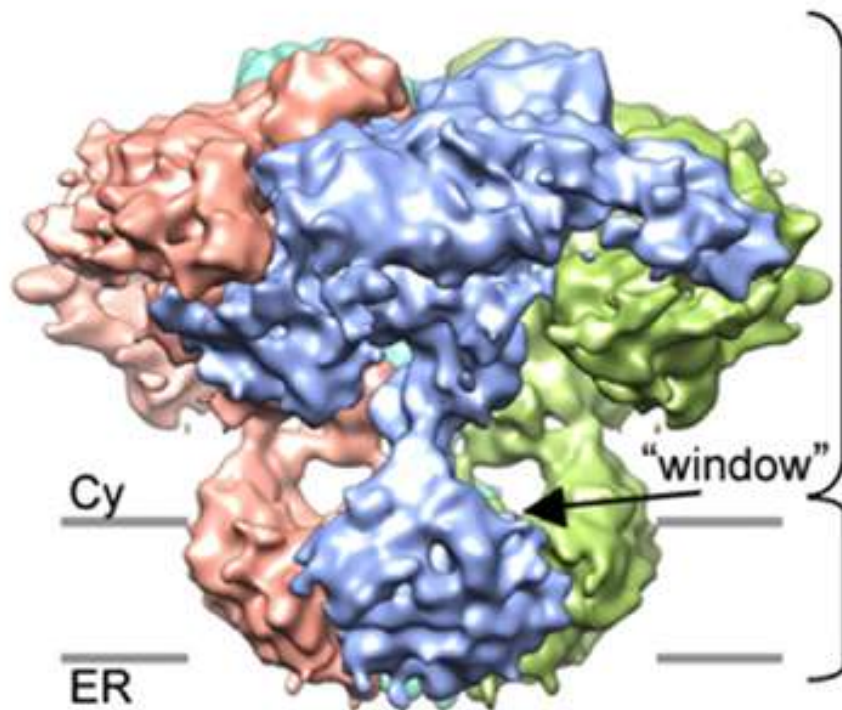
New density map in 2011 at 11 Å resolution



Ludtke *et al.* (2011) *Structure* 19: 1192-1199

# Map validation - IP3 receptor

## Expected secondary structure elements

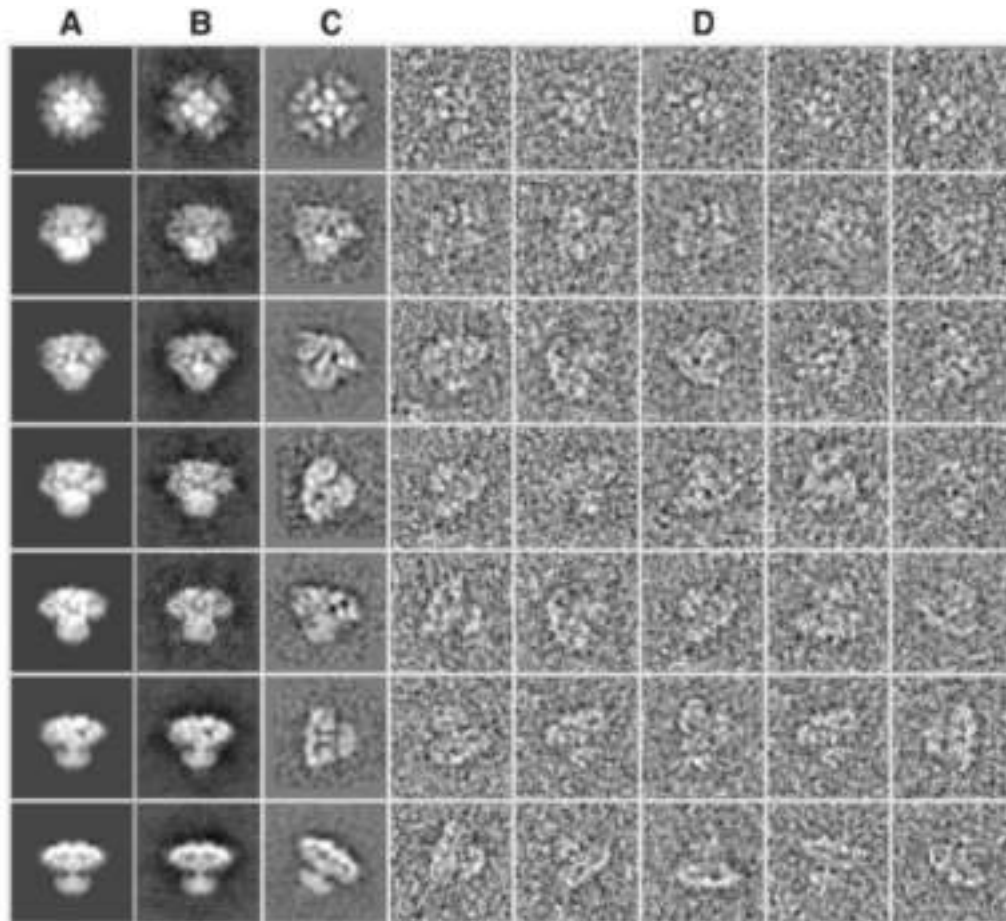


Ludtke *et al.* (2011) *Structure* 19: 1192-1199



# Map validation - IP3 receptor

## Comparison of reference-free averages with projections



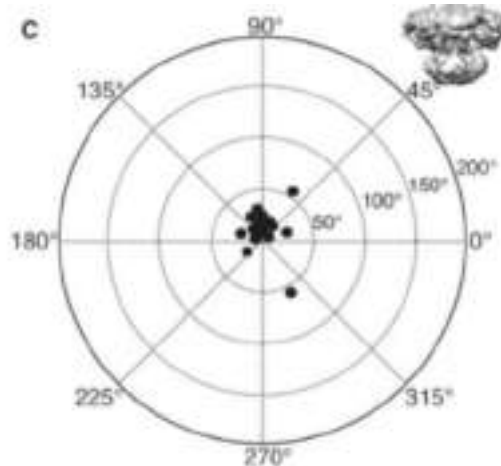
- A: Map projection
- B: Reference-based class average
- C: Reference-free class average
- D: Selected particles

Murray *et al.* (2013) *Structure* 21: 900-909

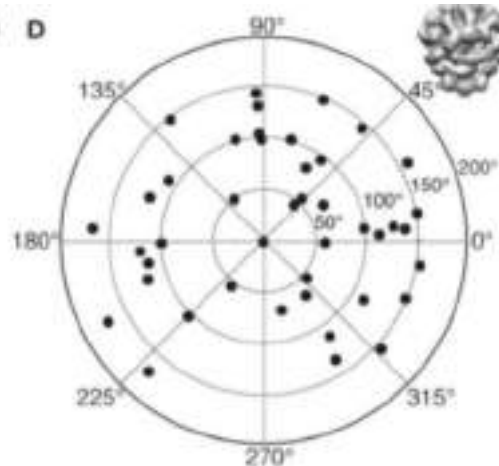
# Map validation - IP3 receptor

## Tilt pair test

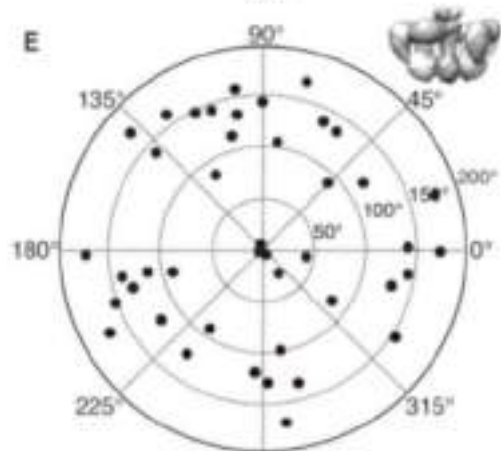
Ludtke *et al.*  
2011



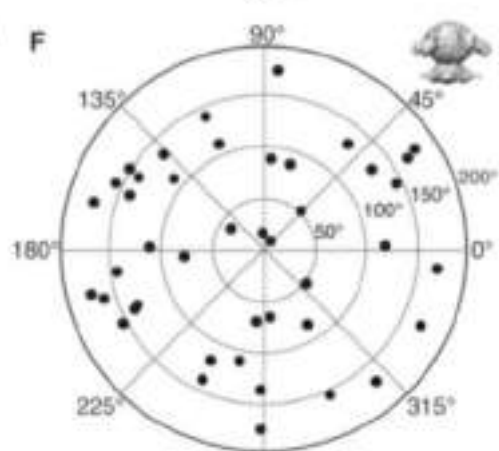
Sato *et al.*  
2004



Serysheva *et al.*  
2003



Jiang *et al.*  
2002

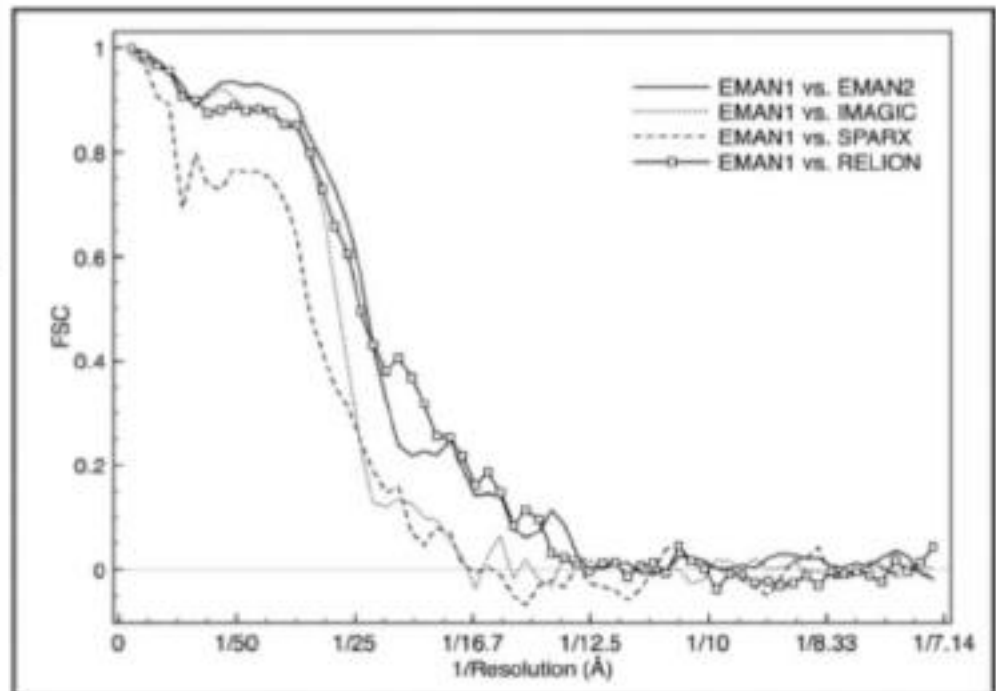
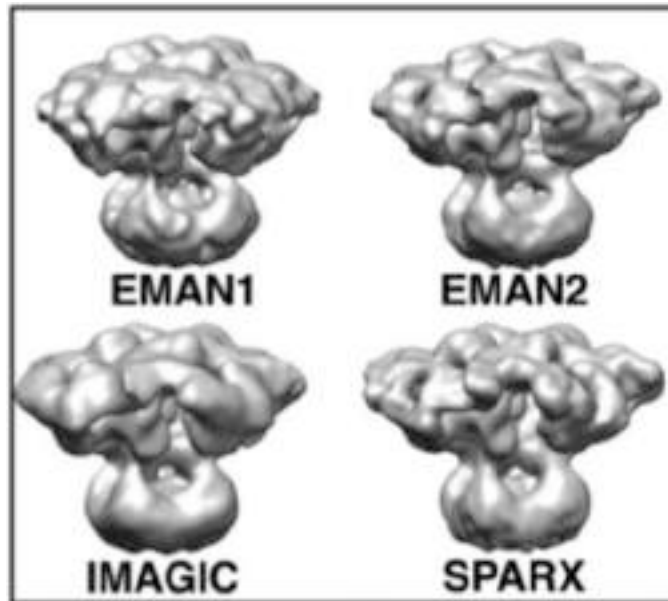


Murray *et al.* (2013) *Structure* 21: 900-909



# Map validation - IP3 receptor

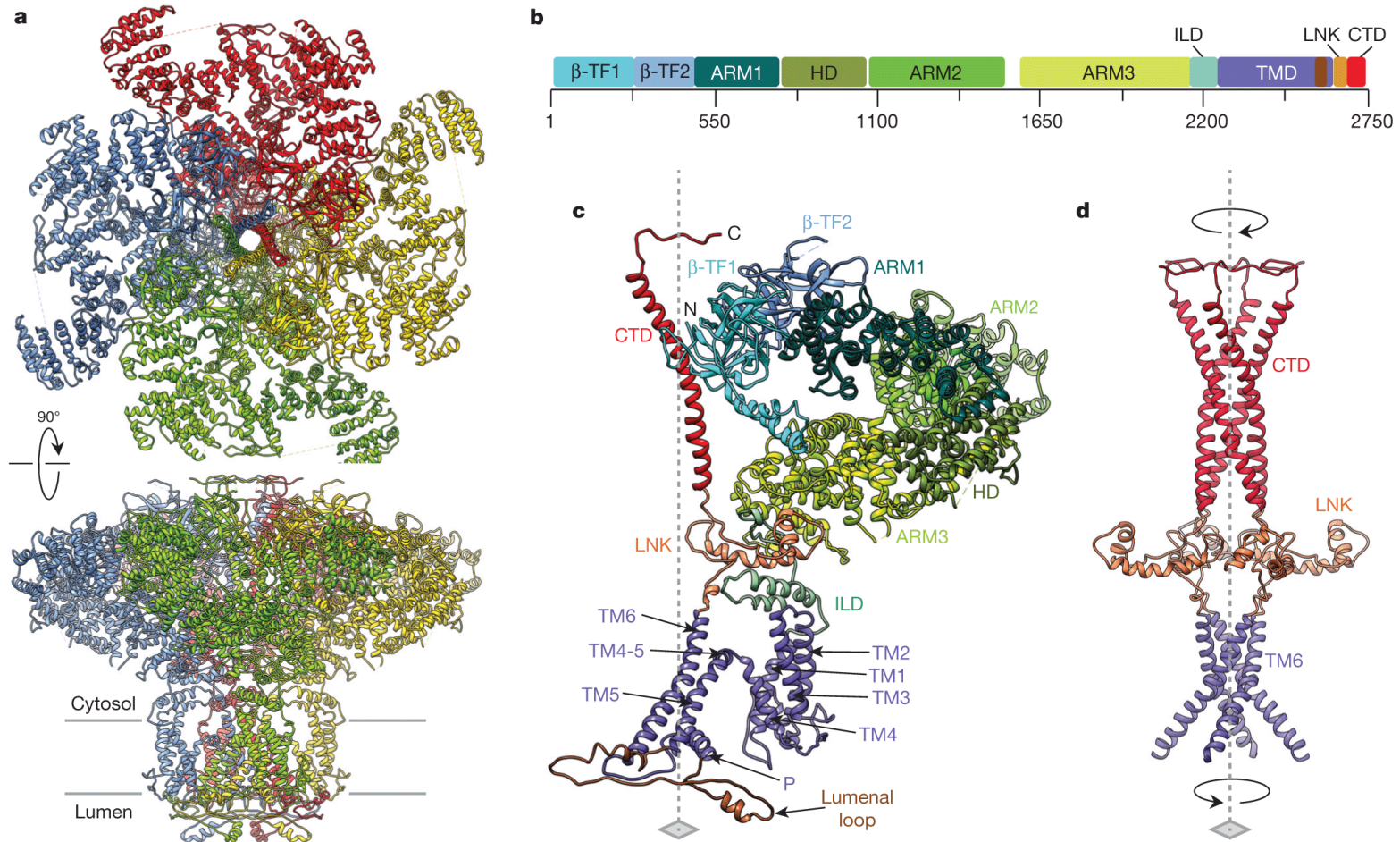
## Comparison of maps from different programs



Murray *et al.* (2013)  
*Structure* 21: 900-909

# Map validation - IP3 receptor

## 4.7 Å resolution structure (2015)

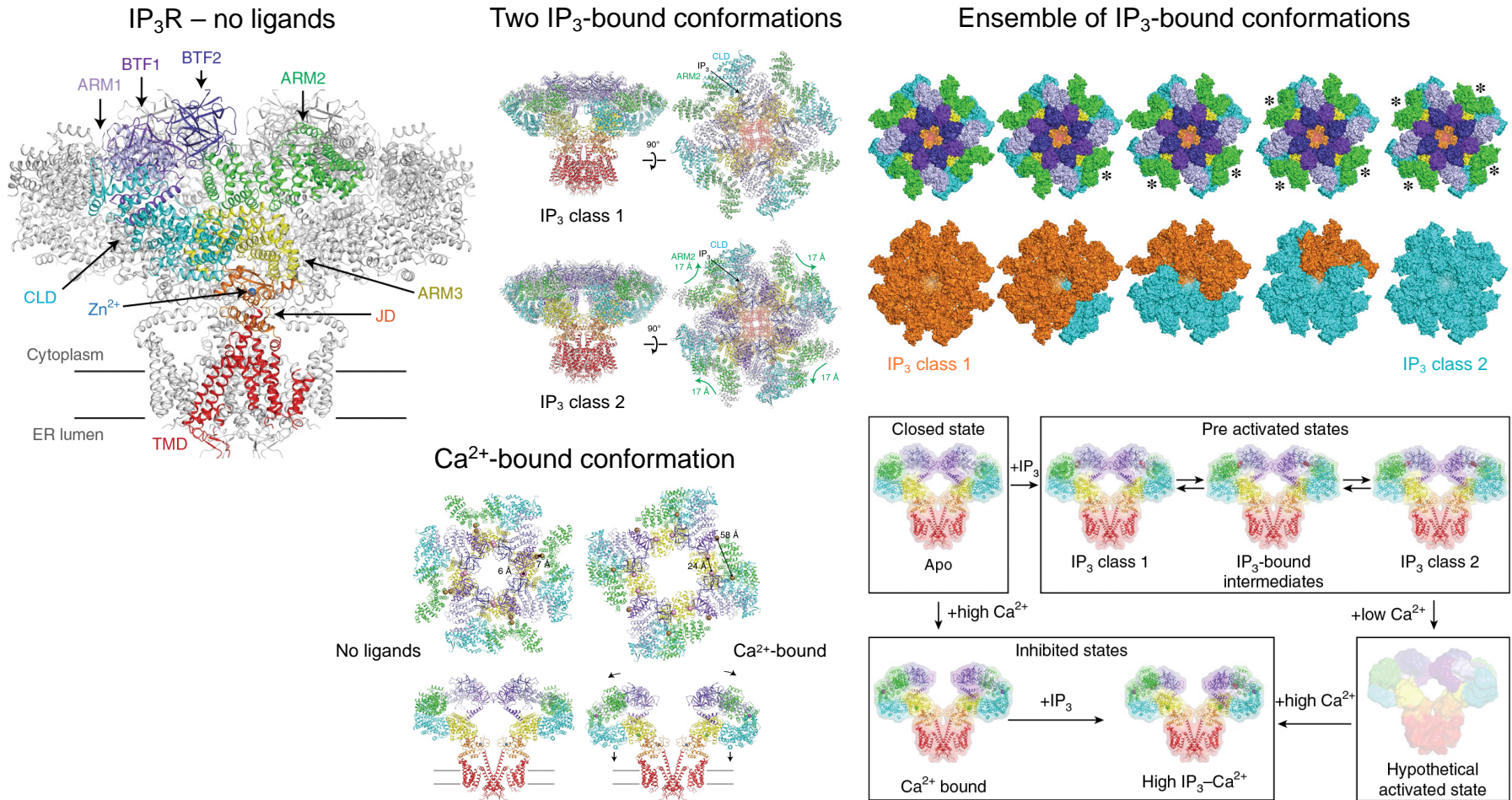


Fan *et al.* (2015) *Nature* 527: 336-341



# Map validation - IP3 receptor

## 3.5 Å resolution structure (2018)



Paknejad & Hite (2018) *Nat. Struct. Mol. Biol.* 25: 660-668