



# **Improving cryo-EM maps: focused 2D & 3D classifications and focused refinements**

**3rd International Symposium on Cryo-3D Image Analysis**

**23.3.2018, Tahoe**

**Bruno Klaholz**

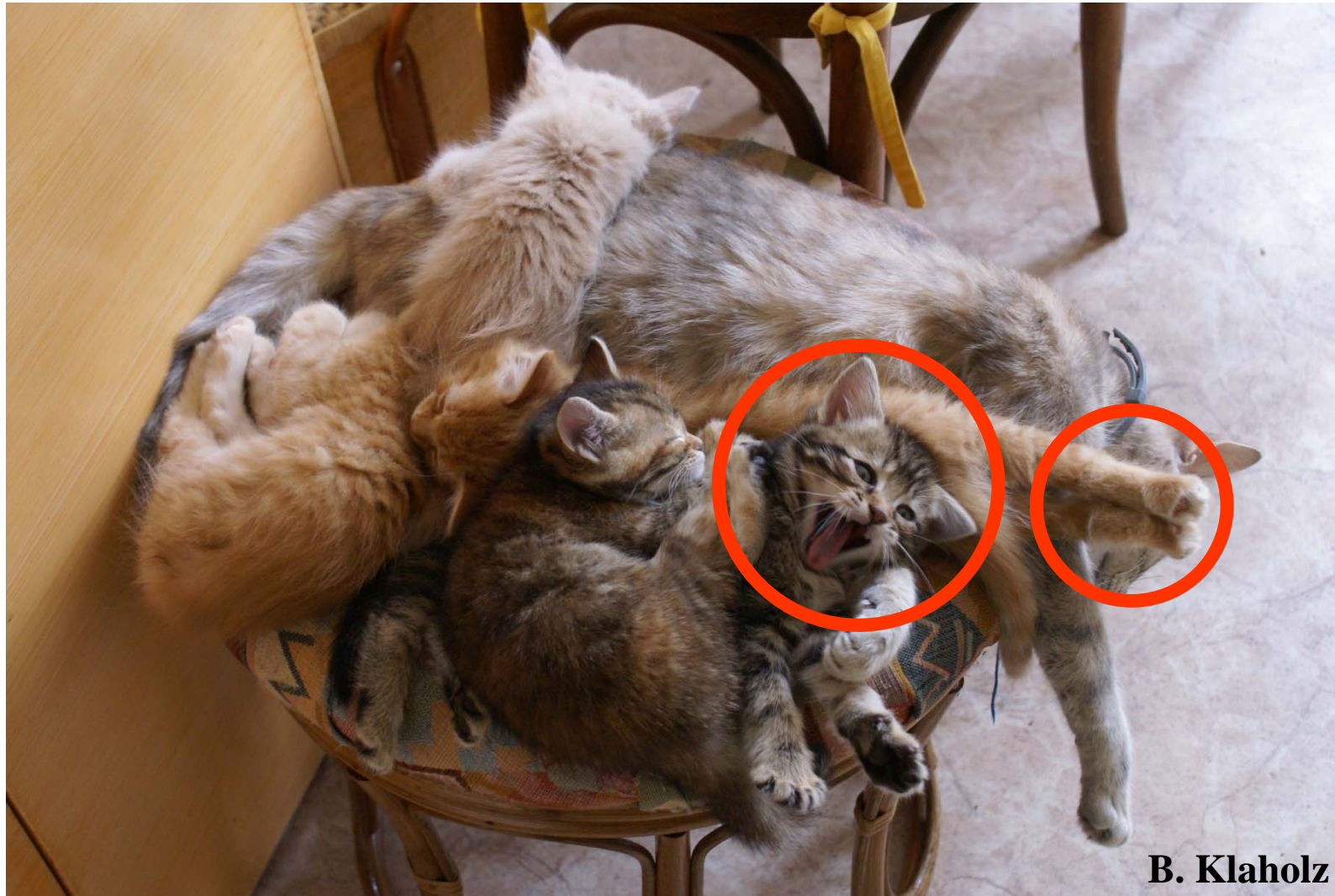
**Centre for Integrative Biology, IGBMC, Illkirch, France**

**<http://www.igbmc.fr/Klaholz>**

*Note: you are welcome to reuse slides elsewhere but please cite your source*



## Conformational changes of cats?



B. Klaholz

# I. Particle sorting

*How to sort out heterogeneity (composition / conformation)?*

*→ particle sorting*

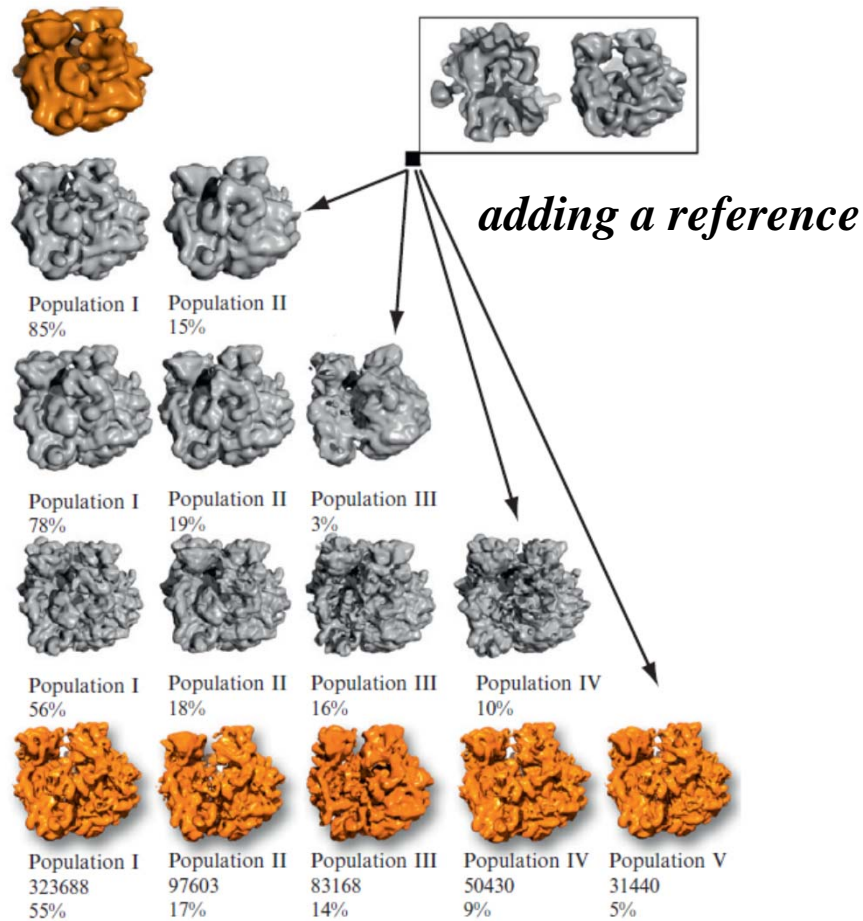
**4 different approaches in the cryo-EM field:**

- 1) reference-based, i.e. cross correlation with forward-projections of known structures**
- 2) multivariate statistical analysis (MSA): 2D classification or 3D classification, variance analysis + resampling, bootstrapping, 3D resampling**
- 3) maximum likelihood based classifications**
- 4) deep learning methods**

# Determining structures of multiple conformational states in a single sample

1) reference-based, i.e. cross correlation with forward-projections of known structures

example:





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

## 2) multivariate statistical analysis (MSA): 2D classification, 3D classification



distinguish: orientational classification and conformational classification

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

## 2) multivariate statistical analysis (MSA): 2D classification, 3D classification



B. Klaholz

front-view, conformation 2

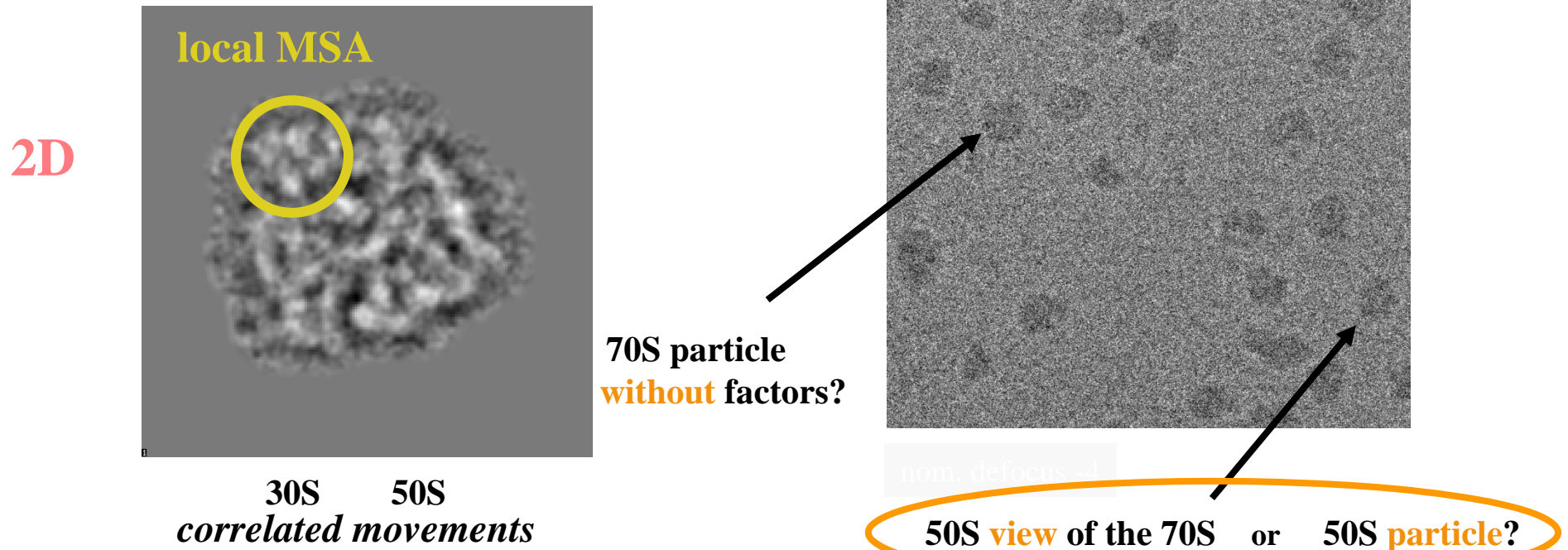
distinguish: orientational classification and conformational classification



# Determining structures of multiple conformational states in a single sample

## 2) multivariate statistical analysis (MSA): 2D classification, 3D classification

### local 2D MSA (focused classification)



Perform 2 classifications:

- (i) global MSA for classification according to particle orientations (i.e. classical class averages)
- (ii) local MSA with a smaller mask for classification according to particle variability.



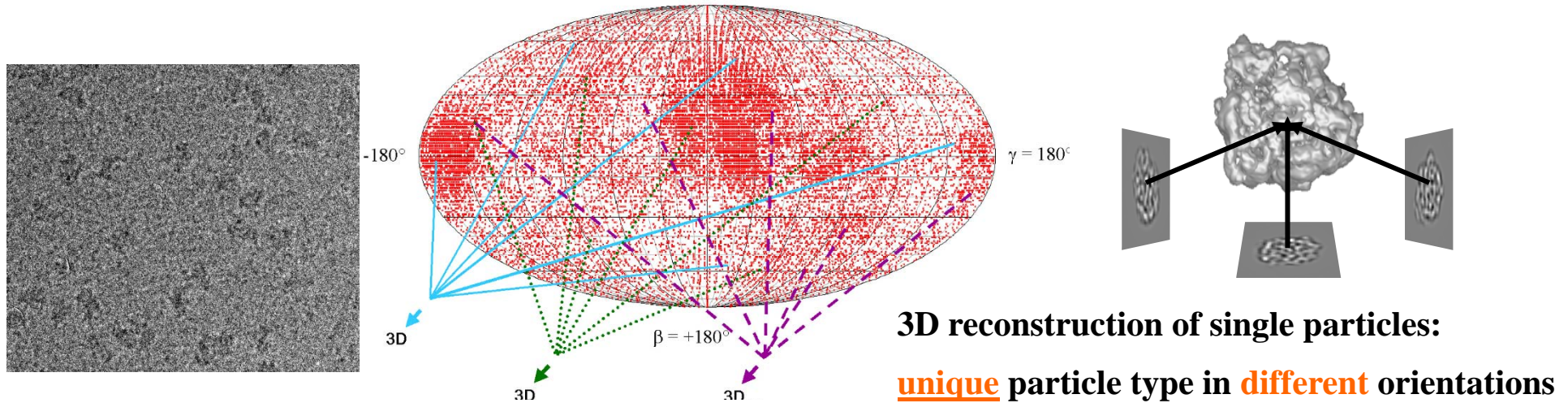
# Sorting out heterogeneity of complexes:

3D

3D statistical analysis and 3D classification:

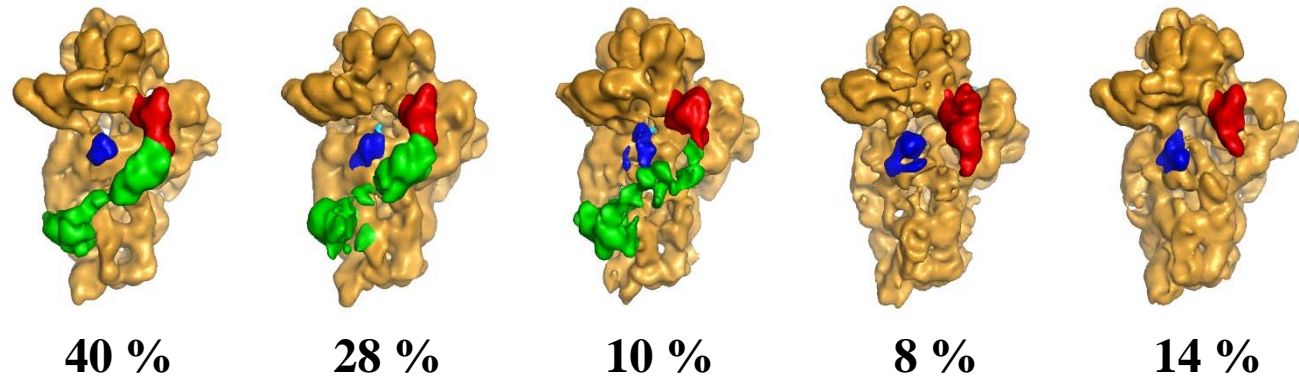
3D sampling and classification (3D-SC)

Klaholz, *Open J. of Stat.*, 2015.



2D → 3D → 4D

200kV FEG data;  
 total 80 000 particles  
 resolution of 3D's: 9Å



→ does both re-sampling and 3D classification;

see also work by P. Penczek (bootstrapping (re-sampling), used primarily to find region of variance)

see also S. Scheres/J-M Carazo (maximum likelihood parameter refinement and classification)

[Simonetti et al., Nature, 2008.](#) used by Fischer et al., *Nature*, 2010; Papai et al., *Nature* 2010.

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

### **3) maximum likelihood based classifications**

→ assign particles to different 3D classes based on maximum likelihood  
(probability distribution; uses randomly selected references + ML-weighting)

**Practically:**

random subsets are optimized and a low-resolution average structure is used as reference,  
i.e. resampling is used in combination with likelihood optimization

e.g. Scheres *et al.*, *JMB* 2005; *Meth. Enzymol.* 2010;

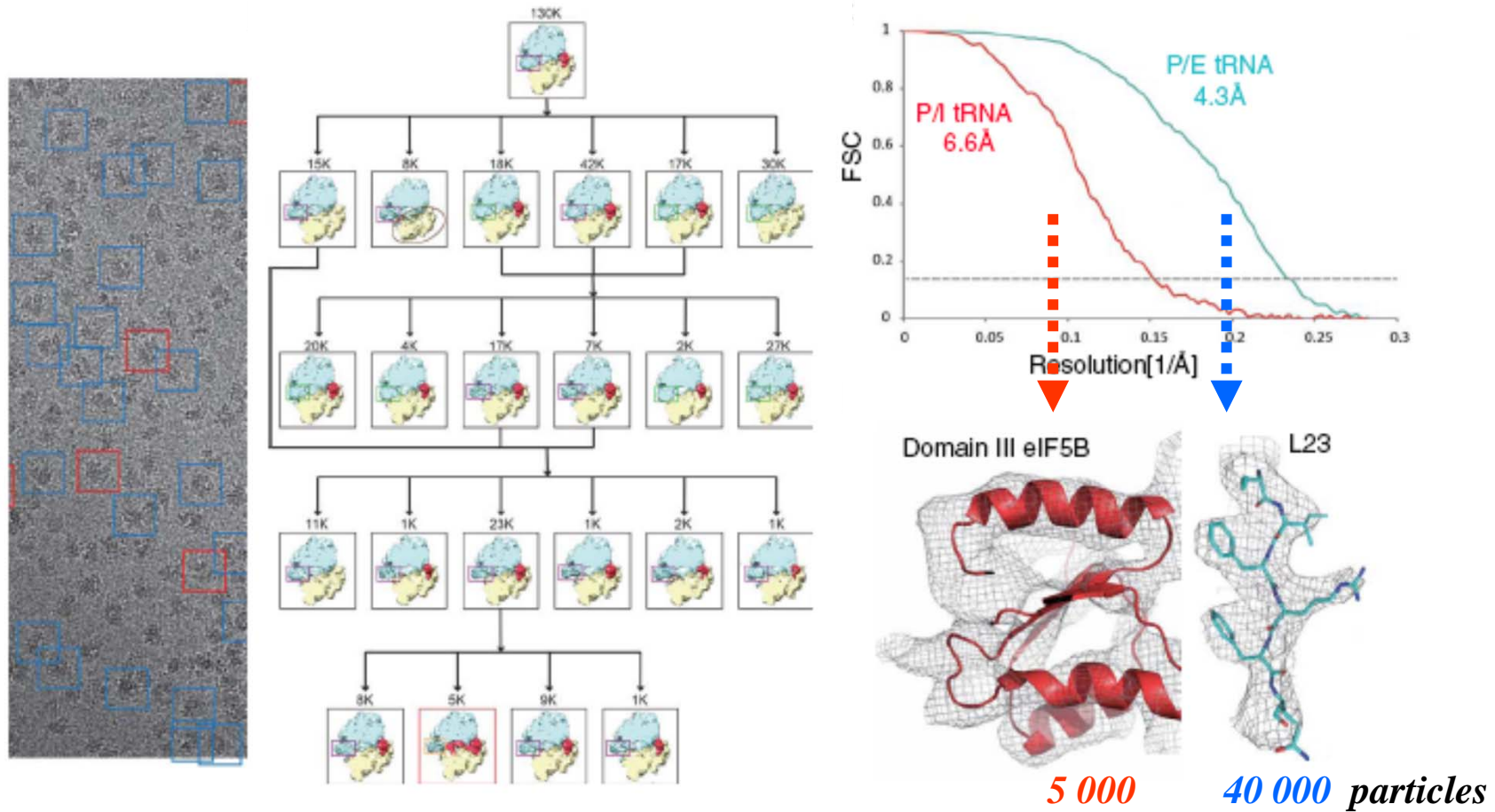
Lyumkis *et al.*, *JSB* 2013

Introduction of the ML concept in cryo-EM: F. Sigworth, *JSB* 1998;

in X-ray crystallography: G. Bricogne, *Acta Cryst A*, 1991



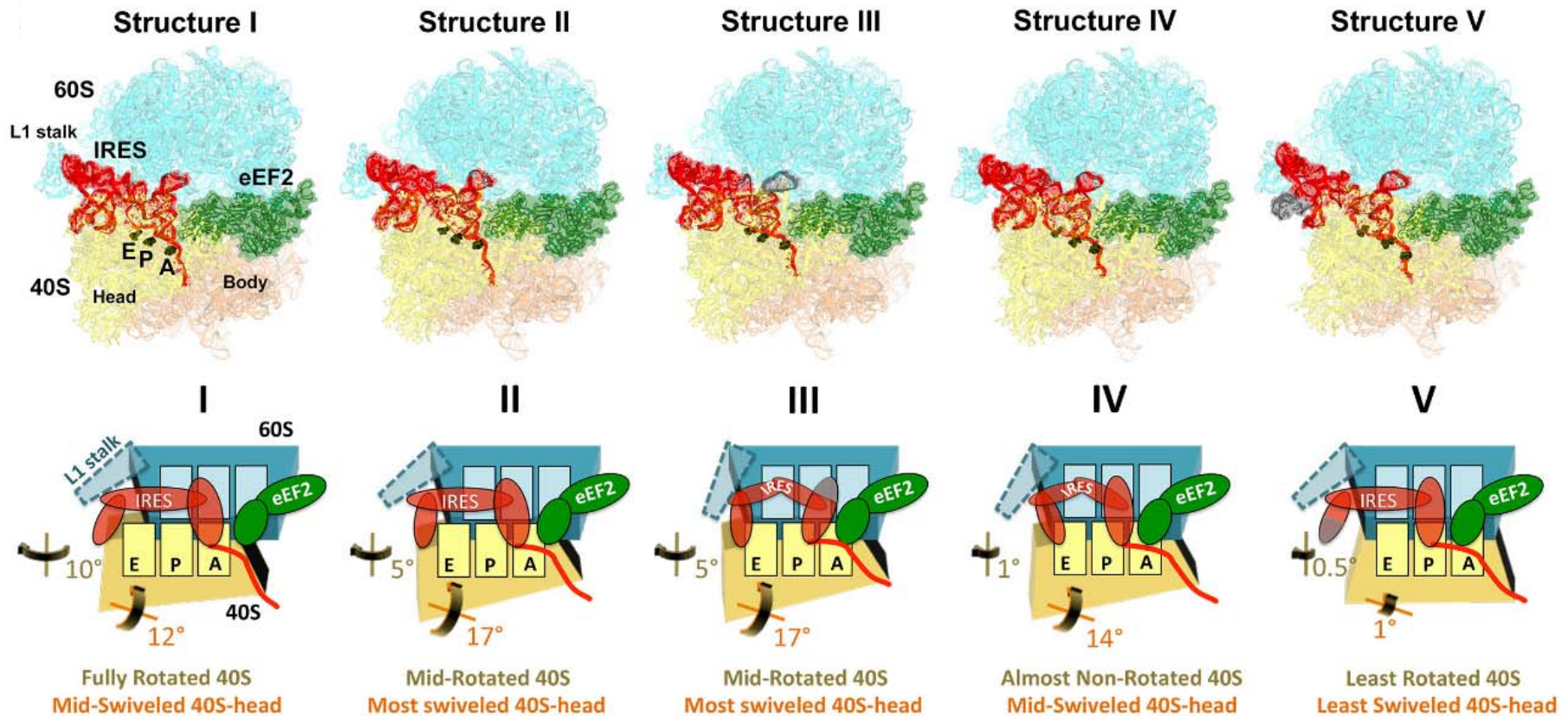
## Examples of ML-based 3D classification



Strong heterogeneity of a reconstituted eukaryotic translation initiation (eIF5B) complex:  
 sorting → 5143 particles, representing **3%** of the population in the sample, **6.6 Å** reconstruction.

Fernández *et al.*, *Science* 2013; V. Ramakrishnan & S. Scheres.

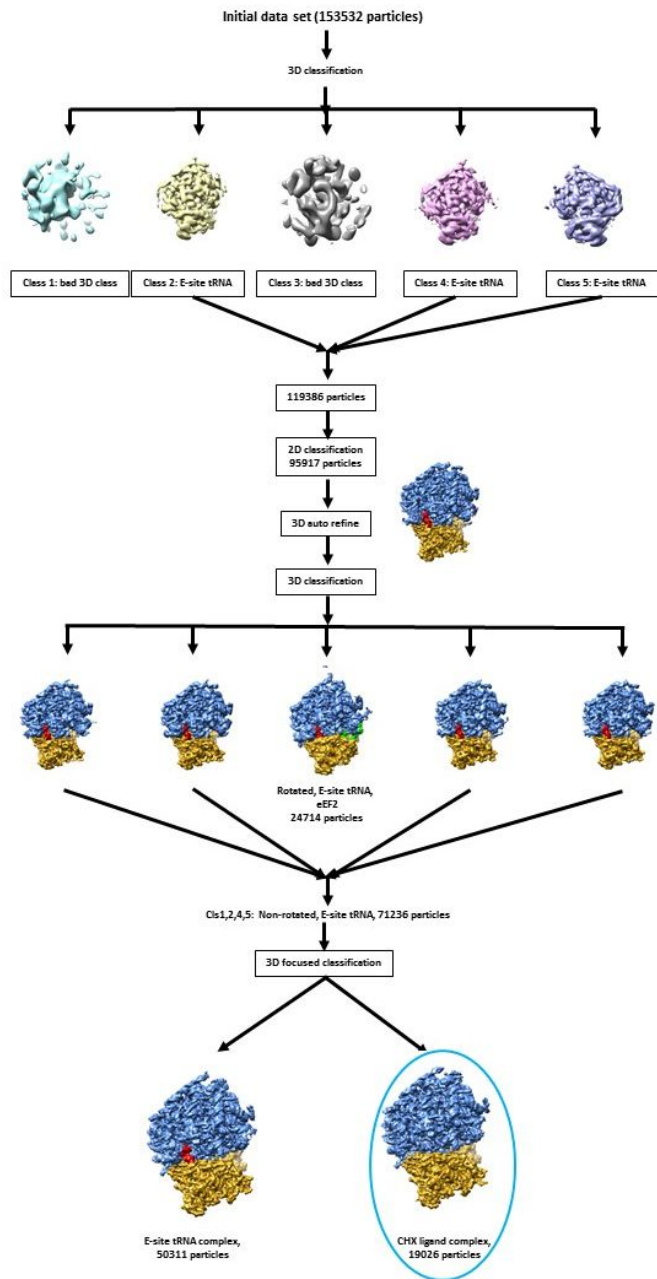
e.g. ML-based focused classification  
of 80S / TSV IRES complex with eEF2/GDP/sordarin



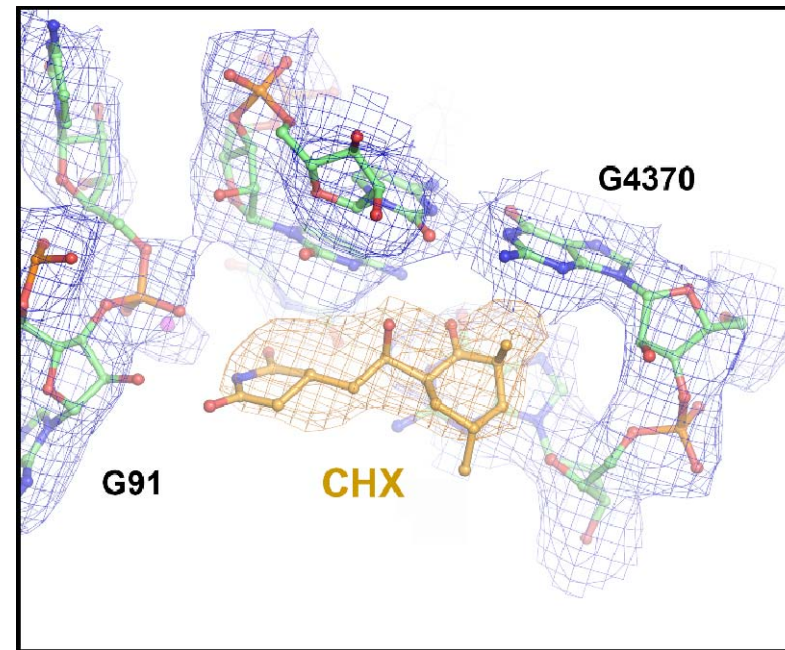
Abeyrathne *et al.*, *eLife* 2016

See also: von Loeffelholz *et al.*, *Curr. Opin. Struct. Biol.* 2017.





e.g. ML-based focused classification  
sorting scheme for human 80S/antibiotic complex

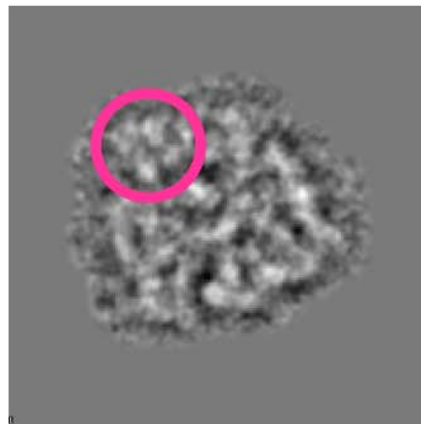


Myasnikov *et al.*, *Nat. Comm.* 2016.

## **II. Focused refinement**



## Local MSA / focused 2D/3D classification & focused refinement:



Concept of local 2D MSA /  
focused classification

see also:

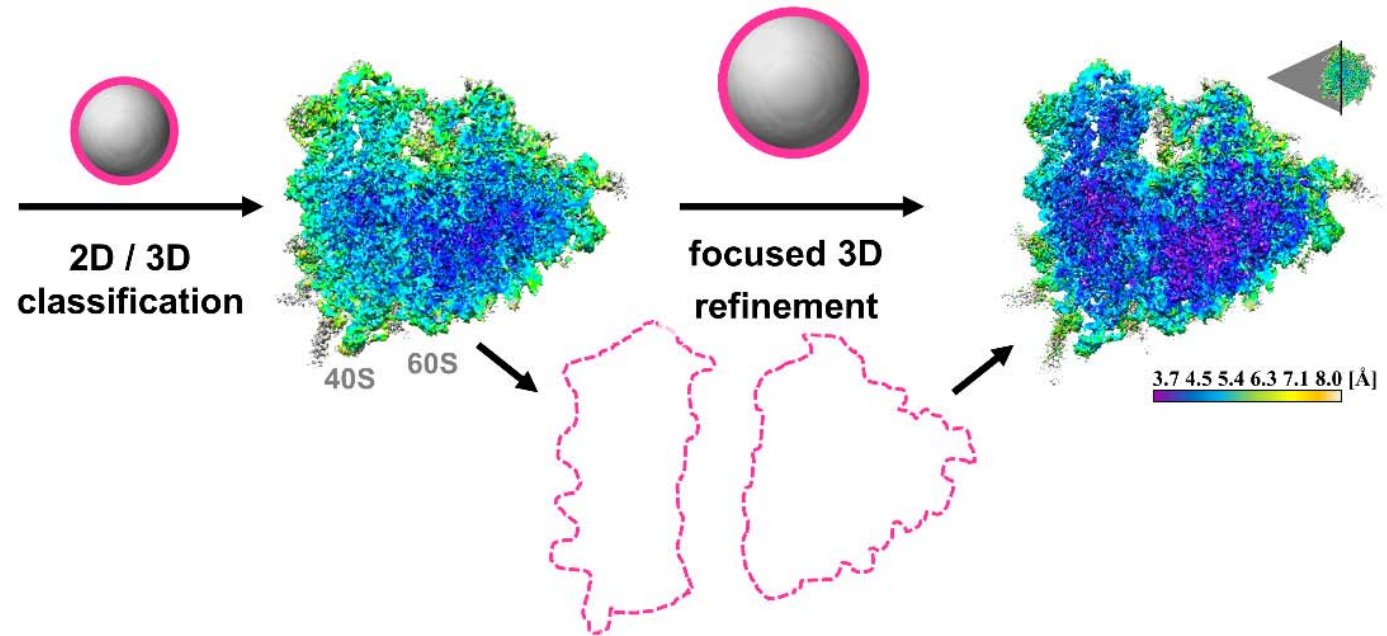
Klaholz *et al.*, *Nature* 2004;

White *et al.*, *JSB* 2004;

Penczek *et al.*, *JSB* 2006;

Wong *et al.*, *Elife* 2014;

...

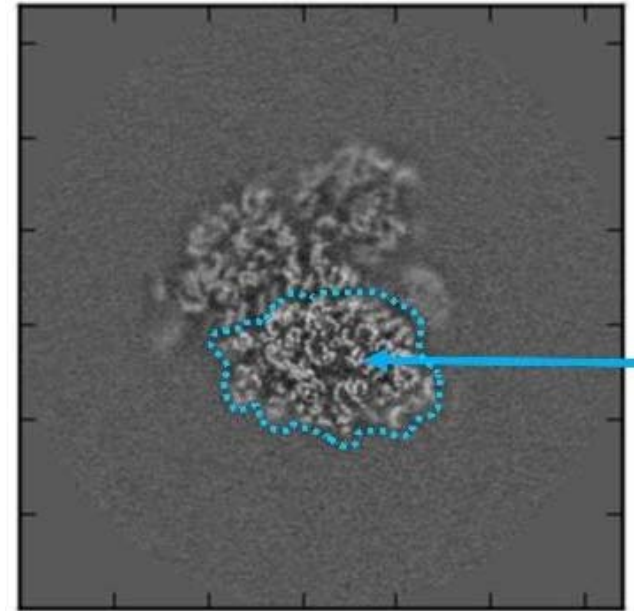
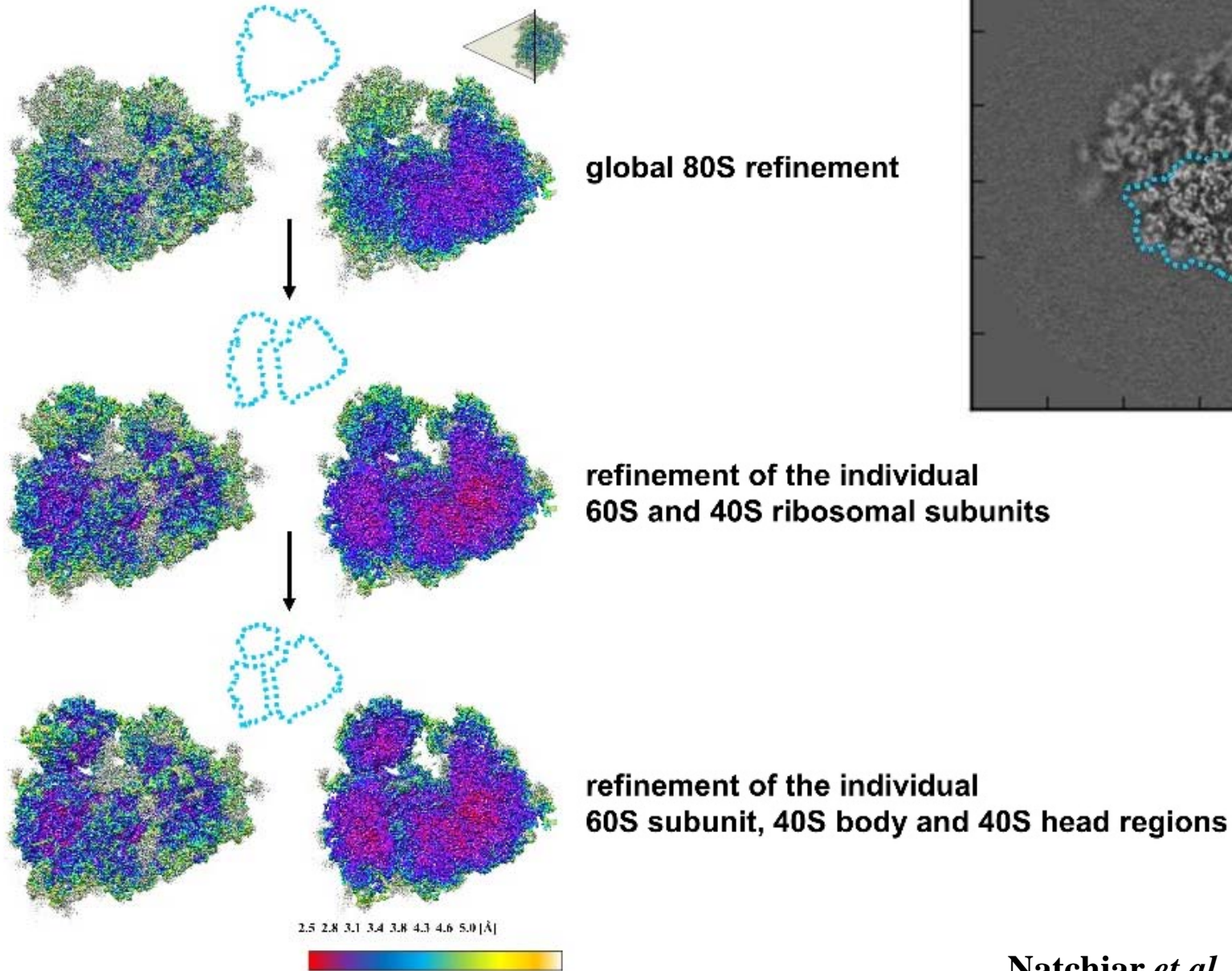


Concept of focused cryo-EM structure refinement through  
- 3D resampling & 3D classification (3D-SC) / bootstrapping  
- maximum likelihood 3D classification  
using spherical mask or dilated, binarized map of region of interest

*Helps: use a slightly larger region than the region of interest, e.g. 30-50 Å in diameter*

von Loeffelholz *et al.*,  
*Curr. Opin. Struct. Biol.* 2017.

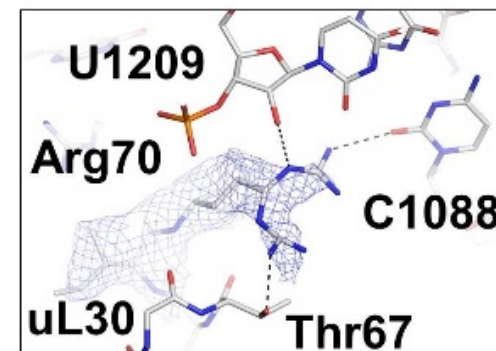
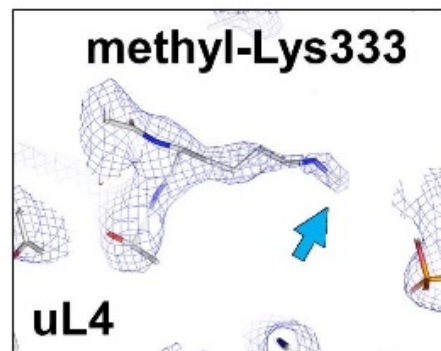
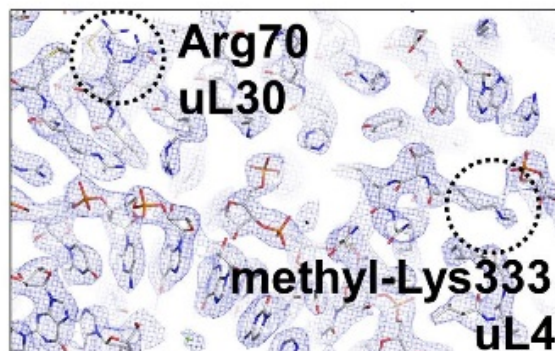
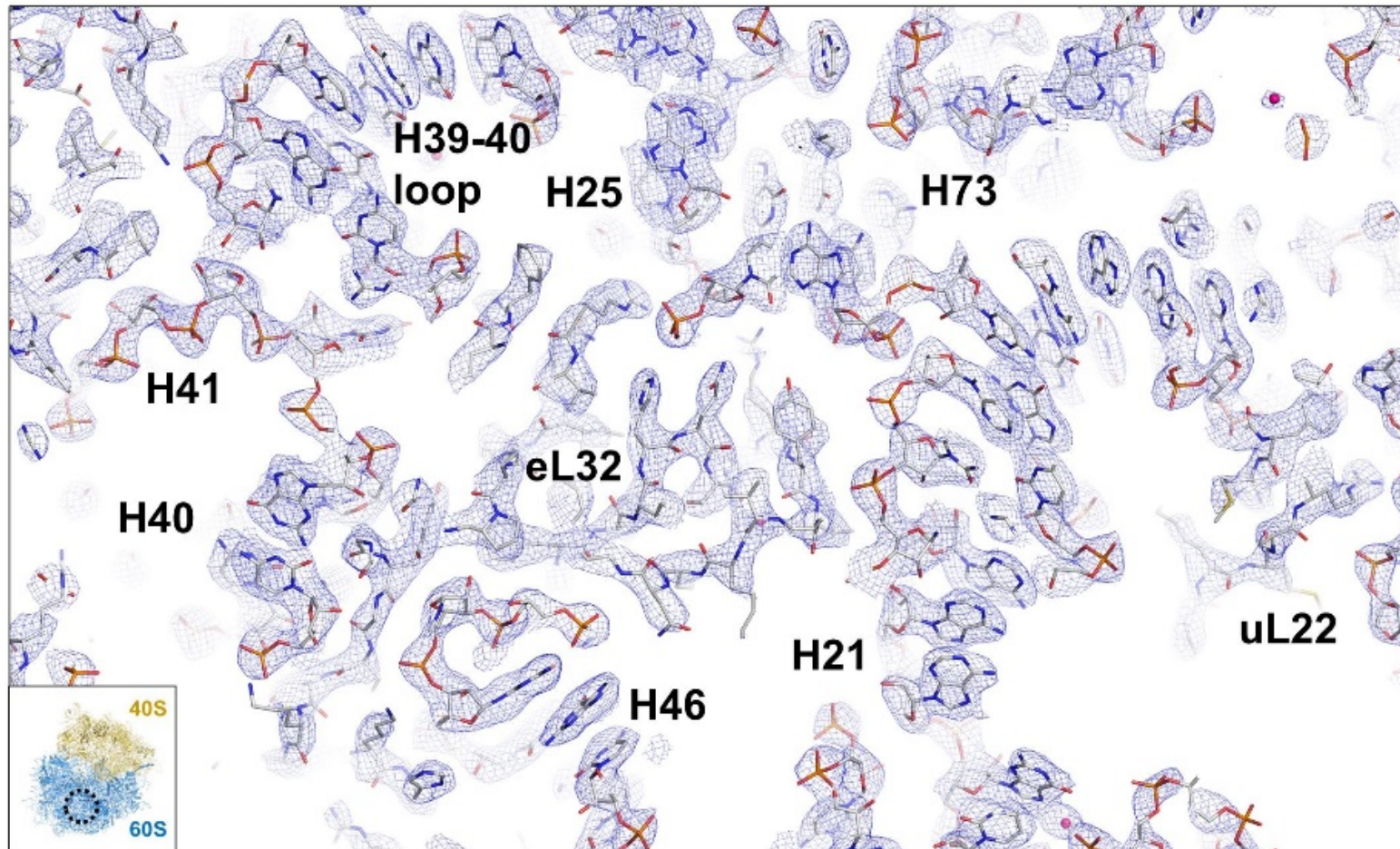
**Focused refinement:**



Natchiar *et al.*, *Nature* 2017.

See also: von Loeffelholz *et al.*, *Curr. Opin. Struct. Biol.* 2017.

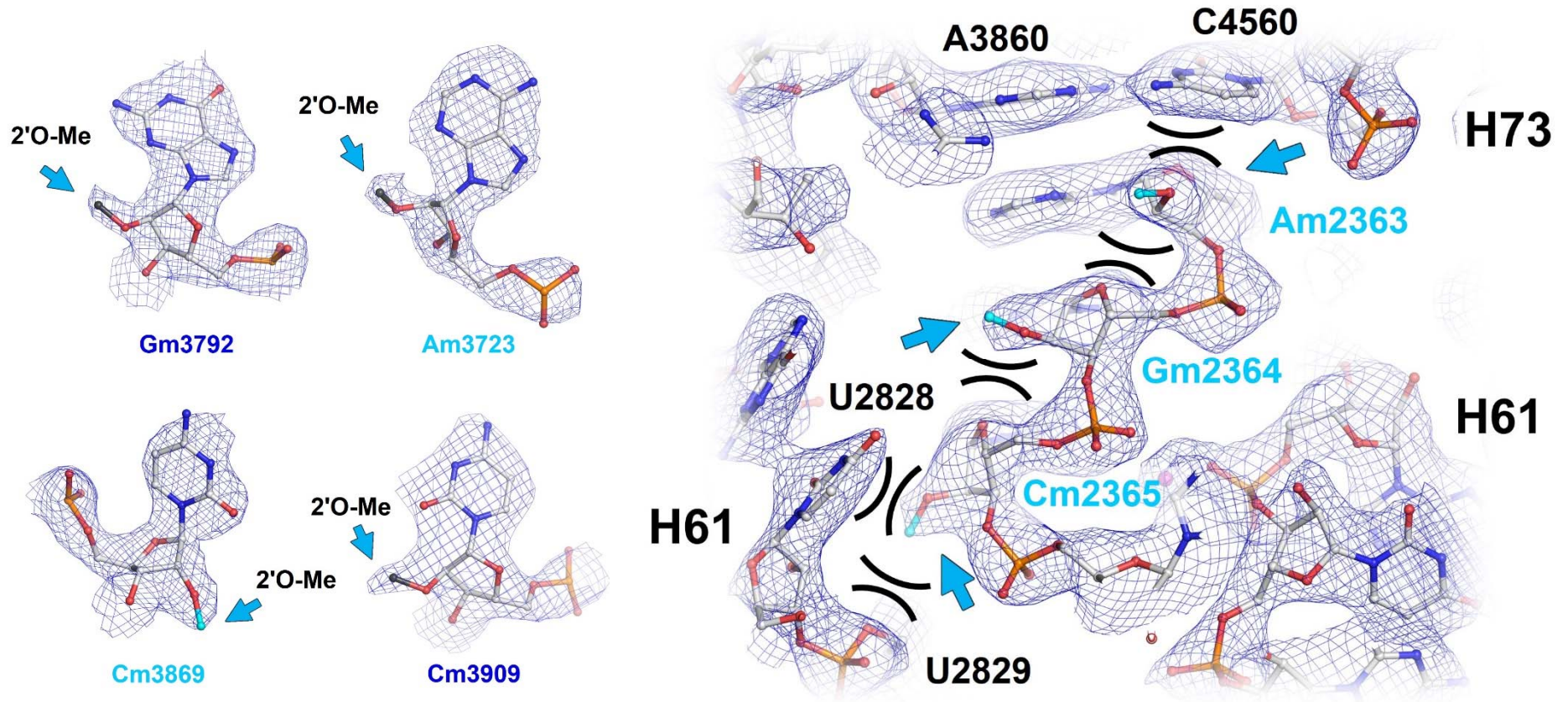




Natchiar *et al.*, *Nature* 2017.



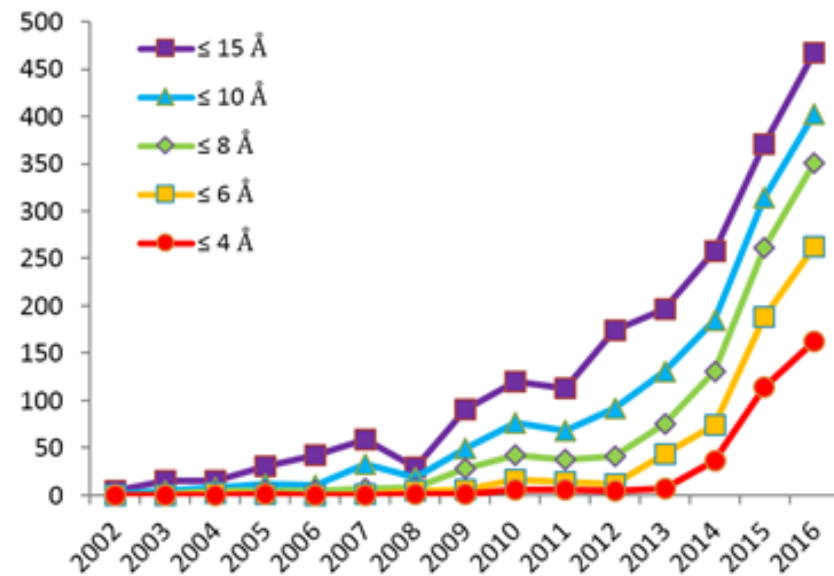
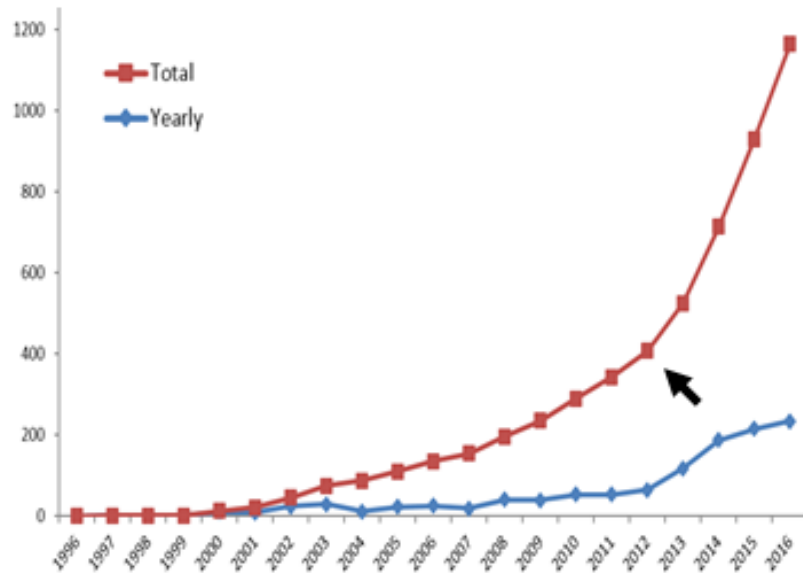
# Allows visualization of chemical modifications in rRNA:



Natchiar *et al.*, *Nature* 2017.

### **III. Improving data collection quality**

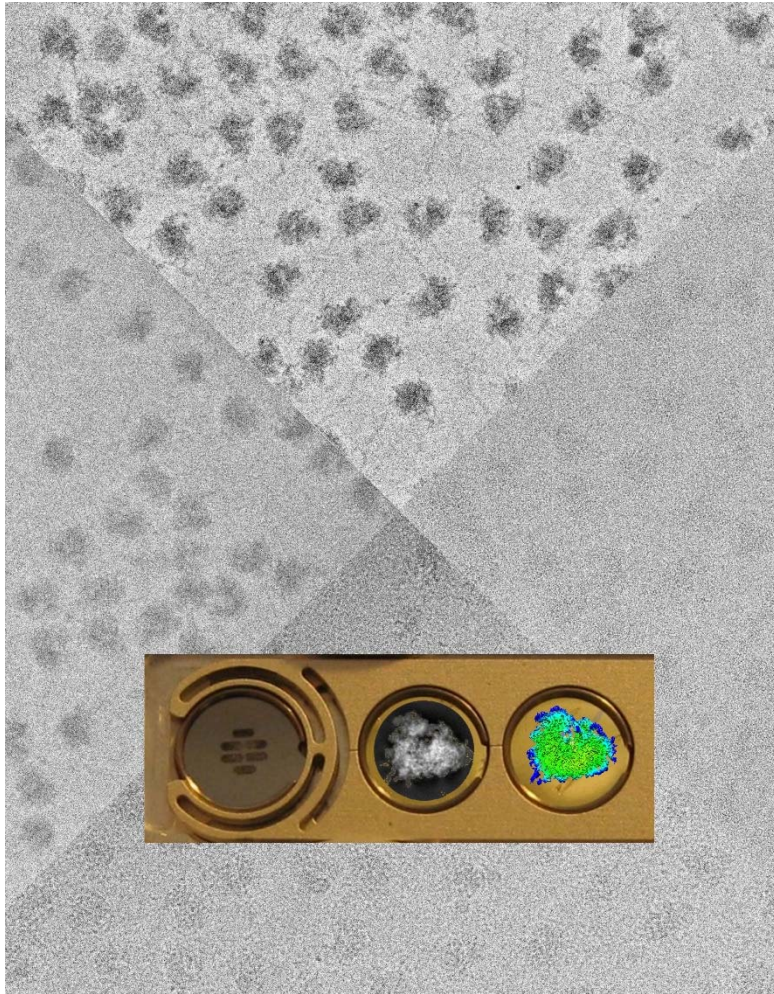
# Detectors, dose weighting, movie alignment



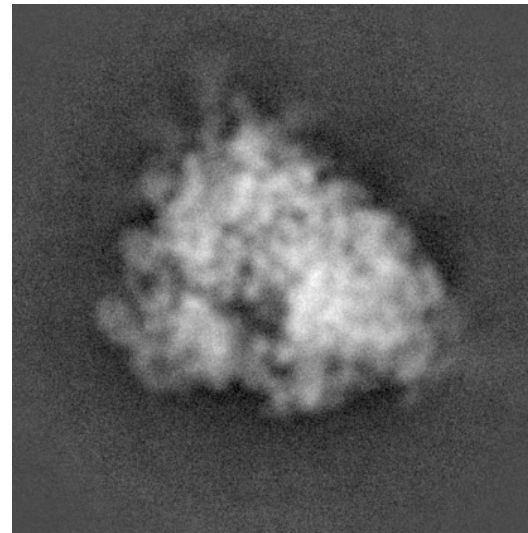
Orlov *et al.*, *Biology of the Cell*, 2017.



# Volta phase plate data collection facilitates image processing and cryo-EM structure determination



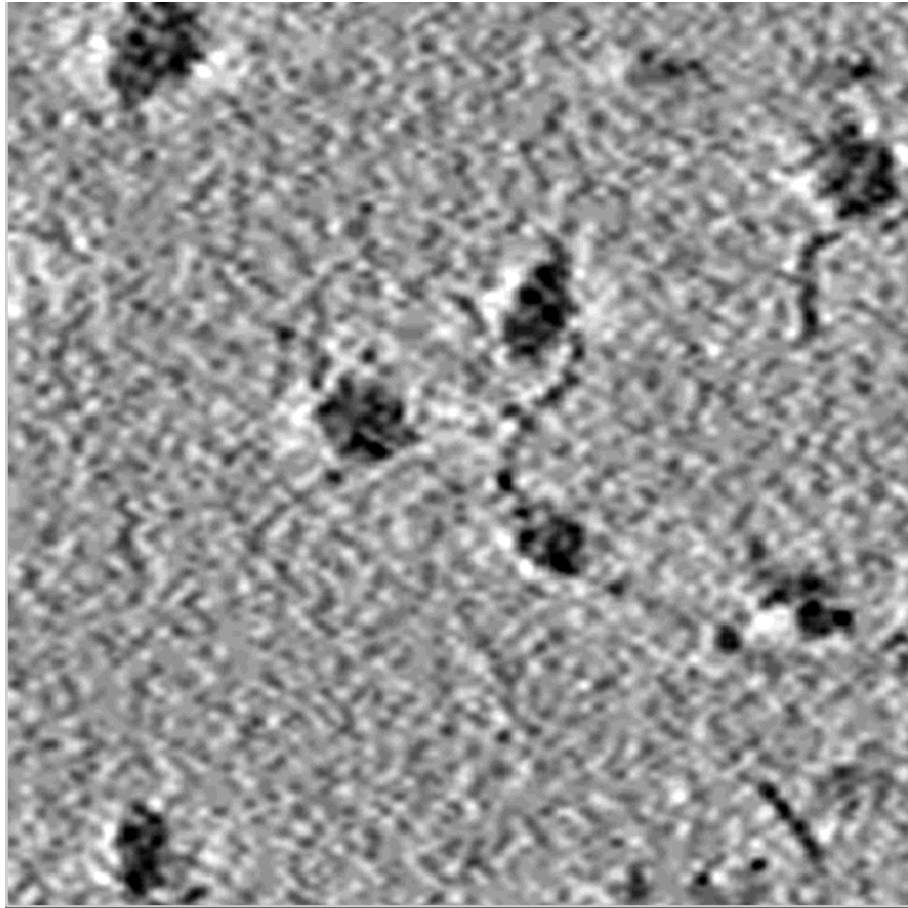
**dfVPP data behave more robustly during image processing:  
particle selection, accuracy in alignments,  
2D & 3D classifications, map interpretation**



**MSA-based classification (3 particles)**

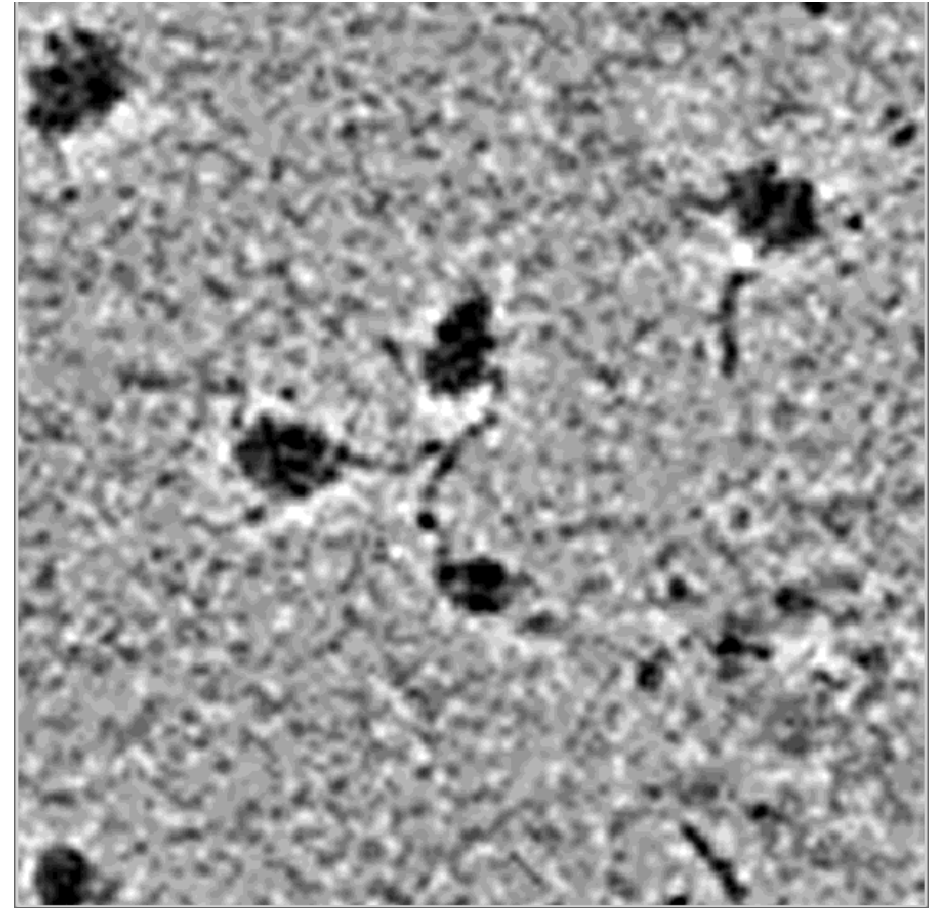
**von Loeffelholz *et al.*, JSB 2018.**

## Single- / dual-tilt cryo electron tomography



*Stripes*  
(*missing wedge*)

**single tilt**



*Less artefacts*  
(*missing pyramid*)

**dual tilt**

*Titan Krios, GIF/K2, VPP, SerialEM, Tom toolbox, IMOD* **T. Frosio & J. Ortiz / Klaholz lab.**

## *Summary*

### **Improving 3D reconstructions and cryo-EM maps:**

- 1) particle sorting (to address heterogeneity)**
- 2) focused refinement (→ composite maps; for a given conformation)**
- 3) improve data quality (CMOS detectors, VPP, dual-axis tomography)**
- 4) to help interpretation: map sharpening**  
**(B-factor, bp-filtering, LocScale, phenix.autosharpen etc.)**

### *Keep in mind:*

- cryo-EM maps are electrostatic potential maps (→ Glu's, Asp's etc.)**
- initial 3D reconstruction (not "model")**
- atomic model building & refinement, proper geometry & B-factors**