What can we apply MSA / classifications to?

- 2D classification (reference-free alignment: only centered data, not rotationally aligned)
- alignment by classification (alignment against class averages or a typical eigenimage)
- analysis of symmetry (through symmetry in the eigenimages)
- local MSA (focus on an area with high structural variability)
- re-classification of class averages belonging to an object view
- size-classification (e.g. White et al., *J. Mol. Biol.* 336 (2004) 453-460).
- 3D classification of structures (separation of mixed particle populations):

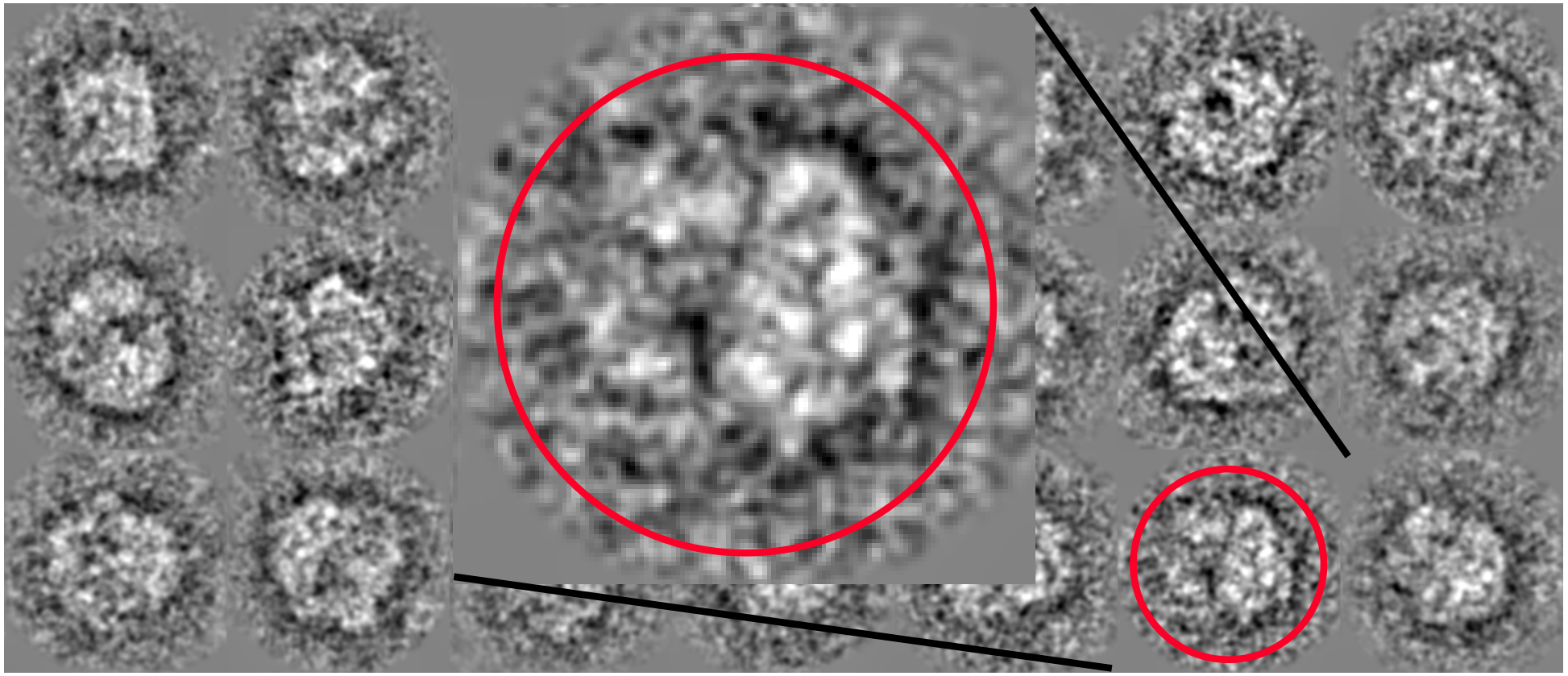
  particles:3D-SC, sub-tomograms
- classification of powerspectra (sorting of defocus classes)

Important to do before MSA:

- normalisation
- filtering
- centered data (aligned if for structure refinement)
- define MSA area: MSA mask

IGBMC

Example of initial class averages:



70S ribosome data set, 500 particles, 50 classes, ~10 members per class

band-pass filter: 300Å - 12Å; *Imagic*

Now find your Euler angles and start the structure…!

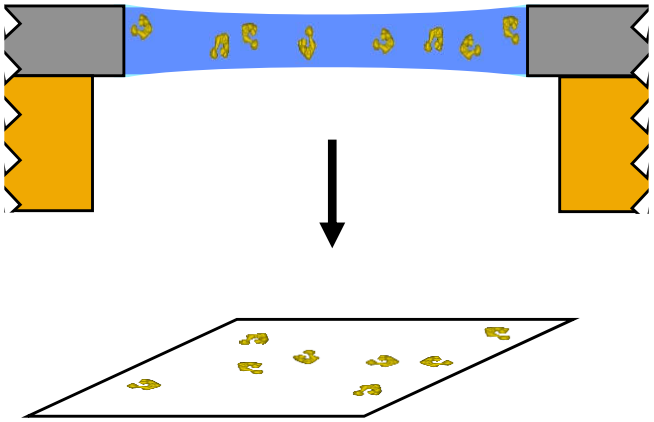
But will it work?

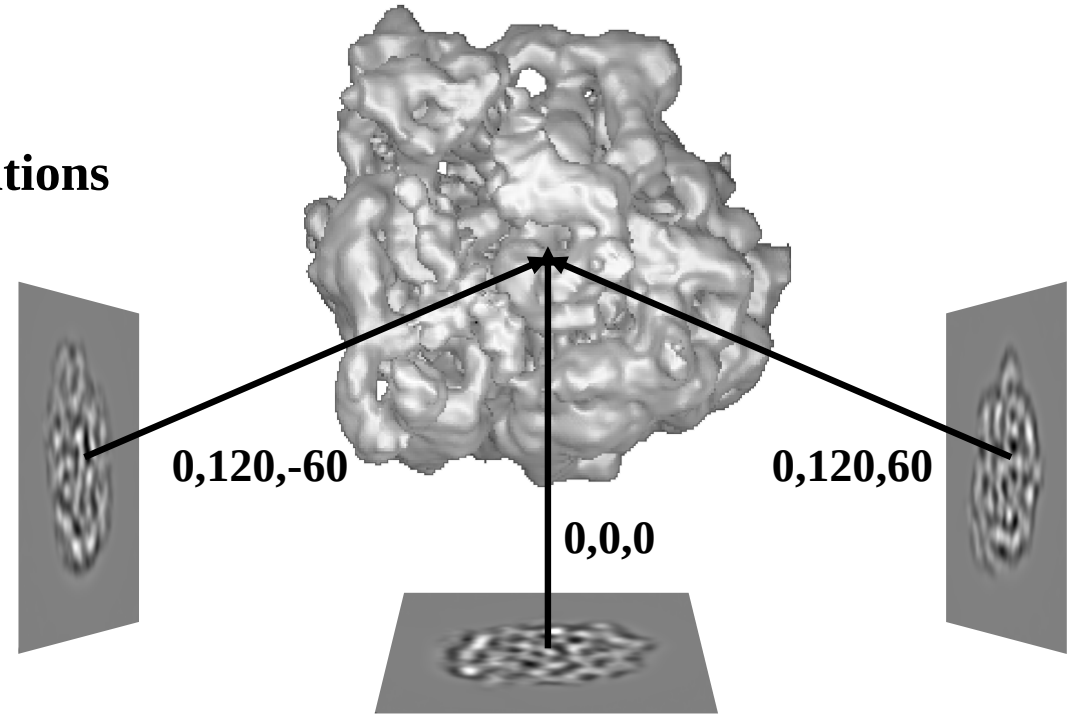
Structure determination by 3D reconstruction

Assumptions?

3D reconstruction of single particles:

unique particle type in random orientations





And what if we have different structures in the sample…?

Conformational changes of cats?

correlated movements?

Different conformations? → Cannot be averaged

IGBMC

What means homogeneity?

- same composition

- same functional state

- same structural state, i.e. same conformational state

Determining structures of multiple conformational states in a single sample

Homogeneity of multi-component system / macromolecular complexes is tricky to get:

- multiple subunits / components (proteins, RNA, DNA,…)

- flexibility of the core structure

Additional, dynamic components:

- factors such as proteins, RNA, DNA binding transiently

- nucleotides (GTP +/- hydrolysable)

- ligands

Determining structures of multiple conformational states in a single sample

Sample heterogeneity:

- **typical problem of multi-component systems, composition & flexibility**
- **structure determination difficult / limits resolution**
- **detection of the problem: disordered 3D map; MSA**
- **advantage: analysis of dynamics**

Structure sorting has important implications for high-resolution structural studies and allows converting the problem of heterogeneity into an advantage by describing structure ensembles to provide insights into the dynamics of multi-component macromolecular assemblies.

see Klaholz, *Open J. Statistics*, 2015.

IGBMC

Determining structures of multiple conformational states in a single sample

How to sort out heterogeneity?

→ particle sorting, advanced image processing

3 different approaches:

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

2) multivariate statistical analysis (MSA): 2D classification or 3D classification

3) maximum likelihood based class assignment

Determining structures of multiple conformational states in a single sample

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

"supervised classification":
uses projections calculated from a known 3D template (i.e. external reference)
for comparison with particle images,
assign group according to best cross-correlation coefficients, i.e. projection matching

if class membership and orientation parameters are estimated simultaneously:
***"unsupervised* classification" (K-means clustering, etc.)**

Practically:
use projections from several known 3D structures as references
(careful with the usage of crystal structures! use a strongly band-passed version),
run multi-reference alignment,
after a few cycles of refinement: add a reference (e.g. apo-form of a complex), iterate

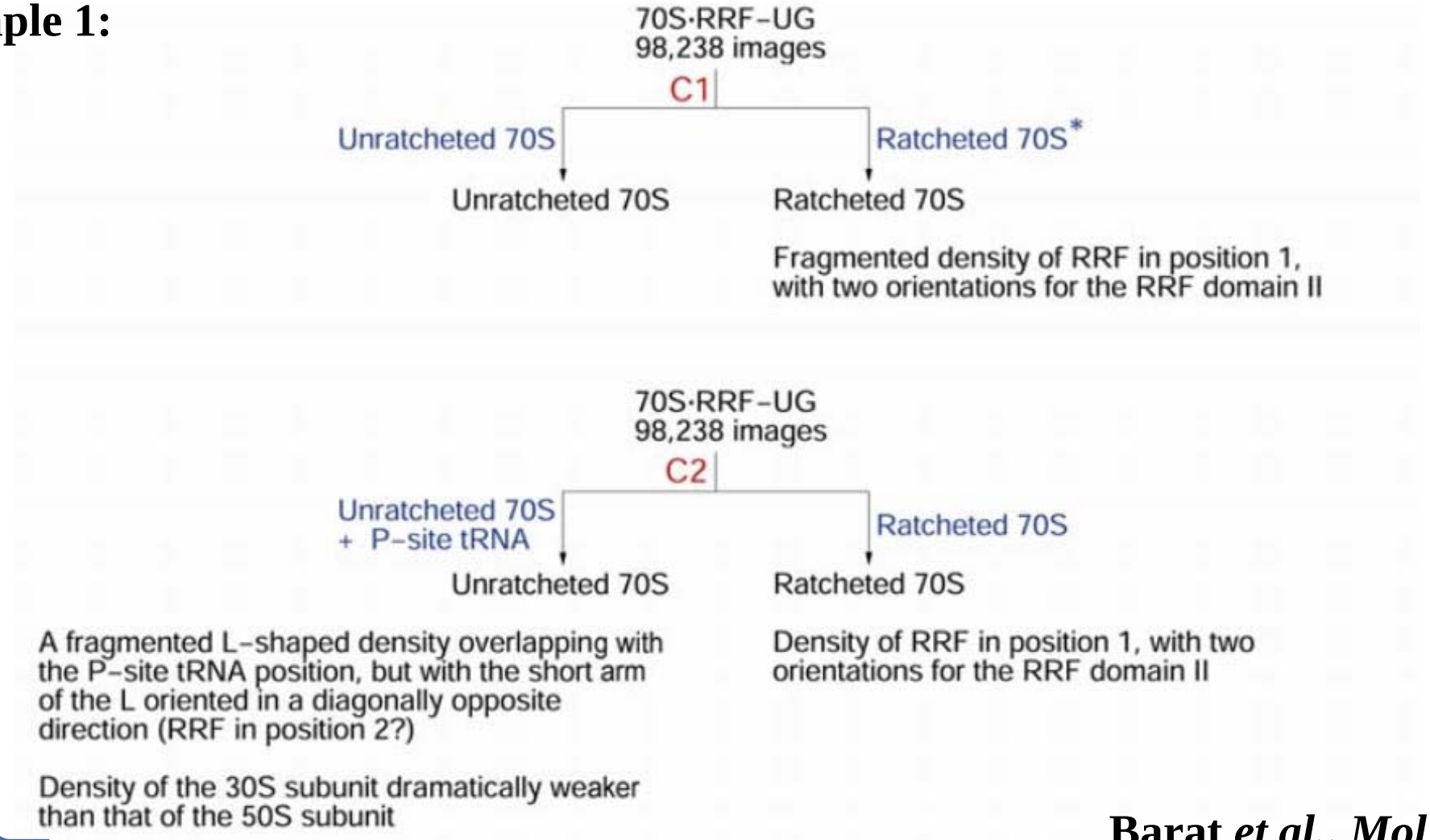
e.g. Gao *et al.*, 2004; Connell *et al.*, 2008; etc.

IGBMC

Determining structures of multiple conformational states in a single sample

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

example 1:



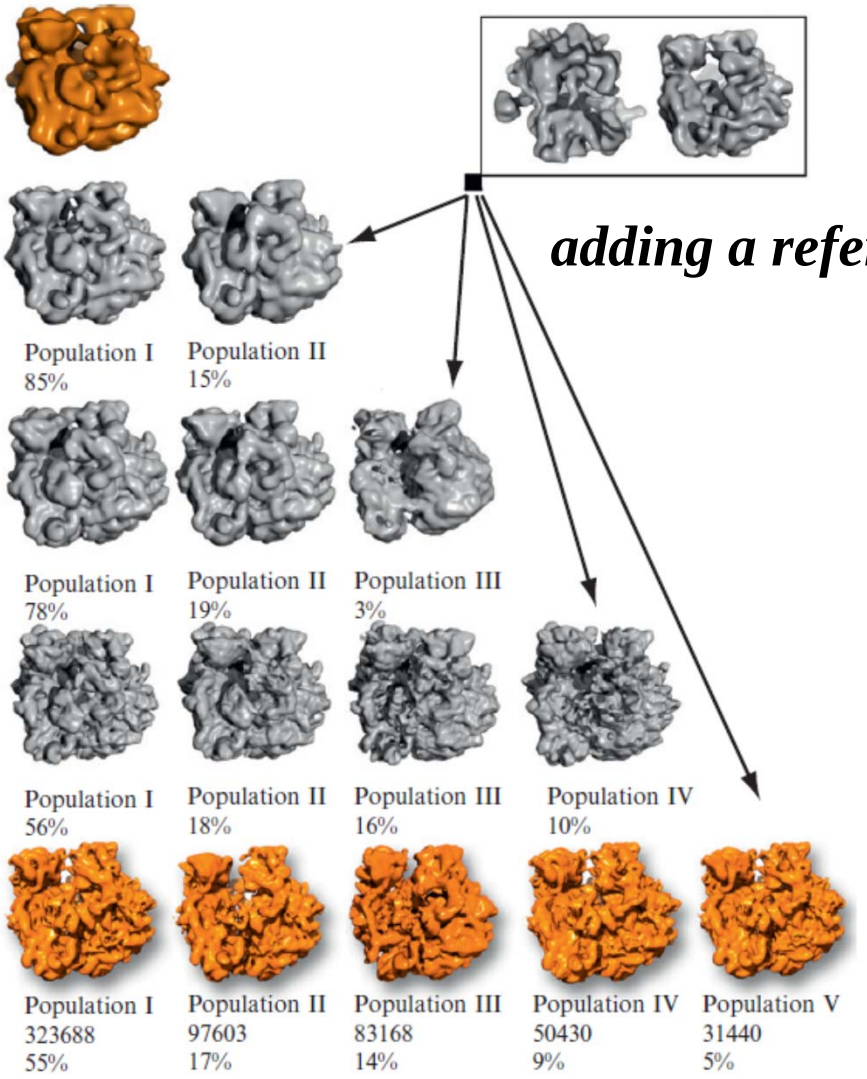
Barat *et al.*, *Mol. Cell* 2007.

IGBMC

Determining structures of multiple conformational states in a single sample

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

example 2



Loerke *et al.*, *Meth. Enzymol.* 2010

IGBMC

Determining structures of multiple conformational states in a single sample

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

Determining structures of multiple conformational states in a single sample

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

distinguish: <u>orientational</u> classification and <u>conformational</u> classification

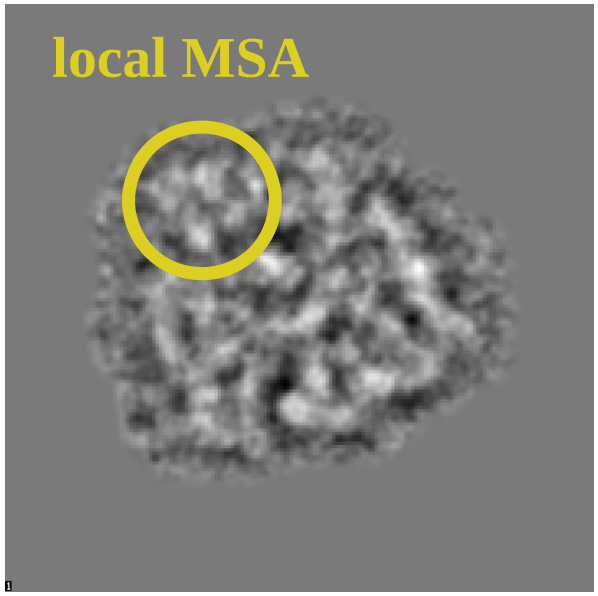
Determining structures of multiple conformational states in a single sample

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

distinguish: <u>orientational</u> classification and <u>conformational</u> classification

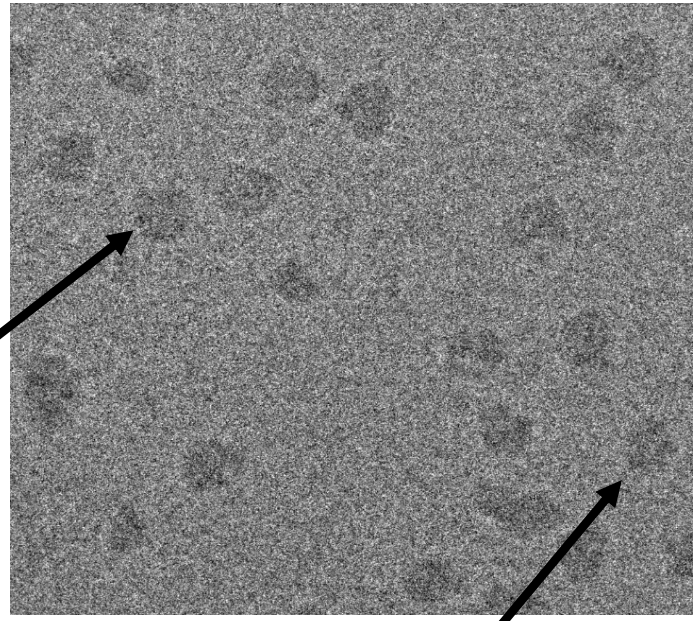
Determining structures of multiple conformational states in a single sample

local 2D MSA (focused classification)



30S 50S

70S particle
without factors?



50S view of the 70S or 50S particle?

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.

Klaholz *et al.*, *Nature* 2004; see Suppl. Mat.

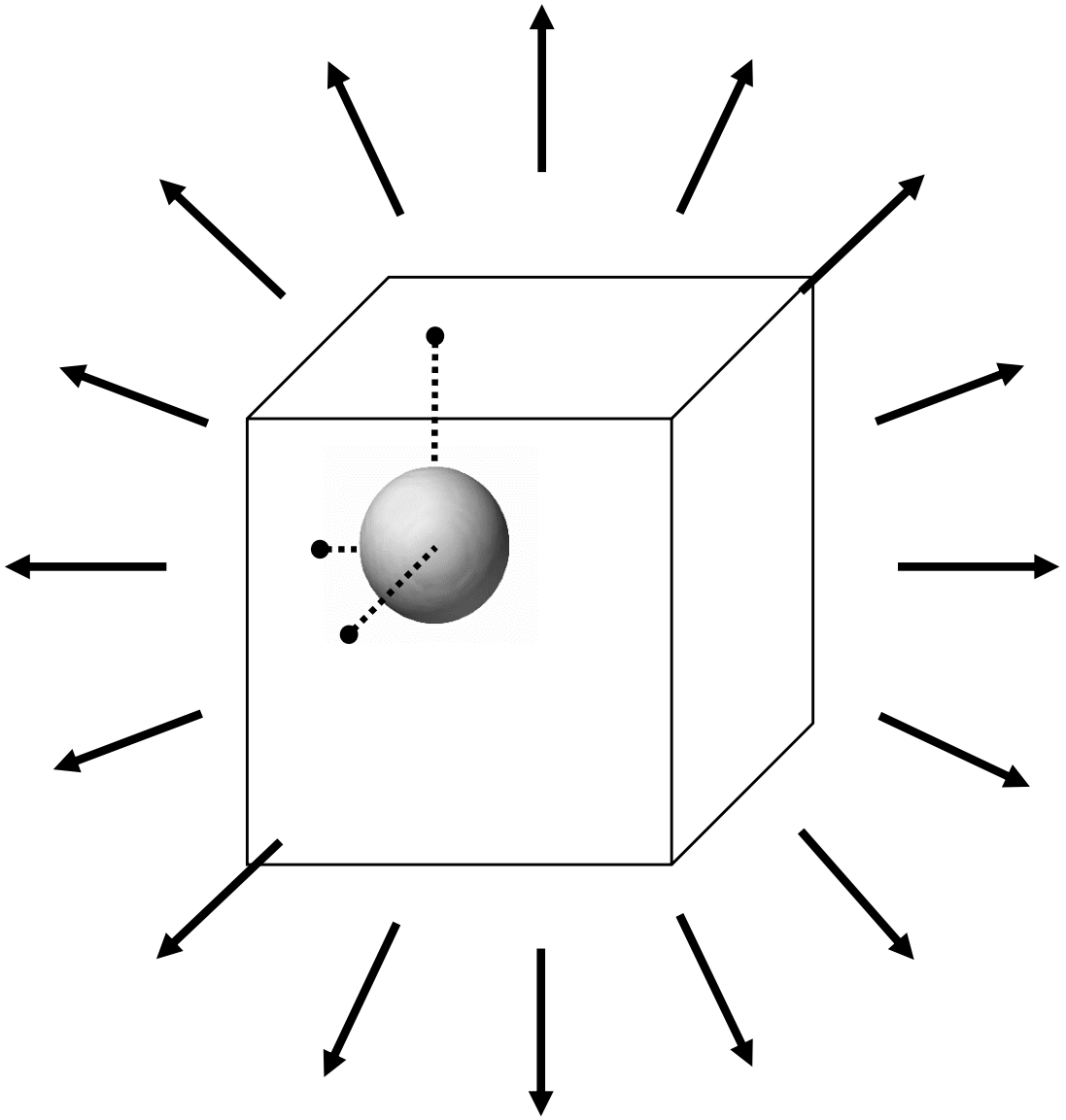
IGBMC

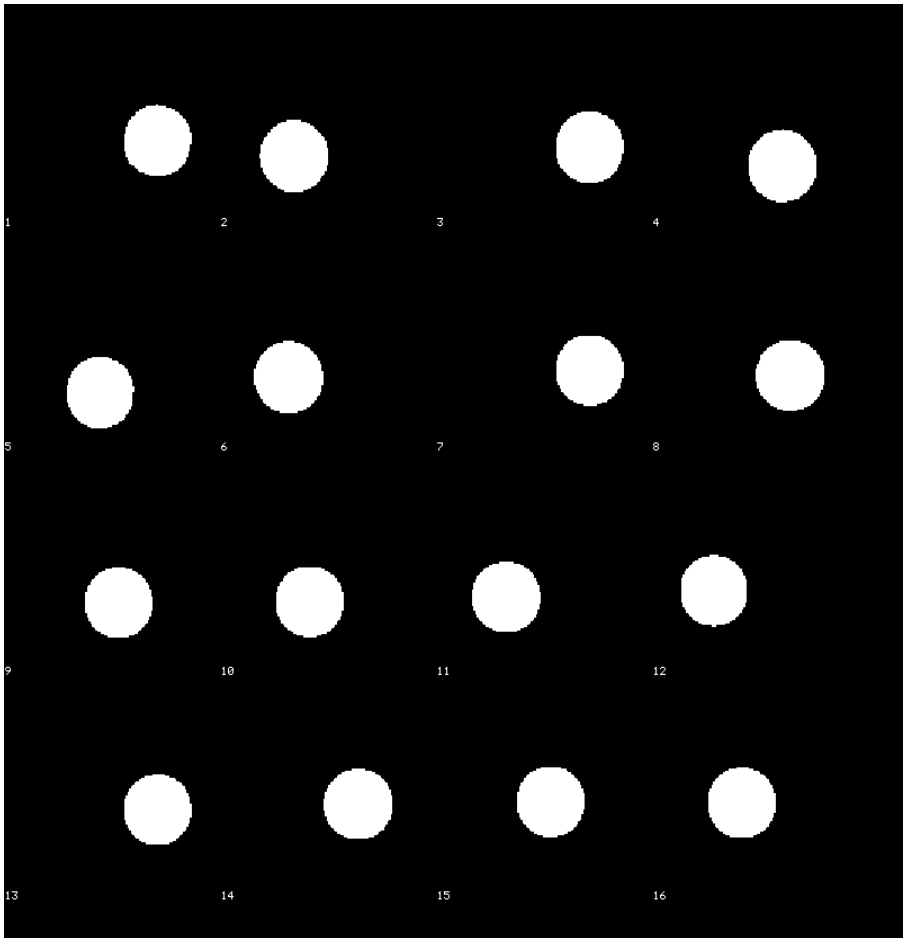
Determining structures of multiple conformational states in a single sample

Local / focused classification

IGBMC

Determining structures of multiple conformational states in a single sample





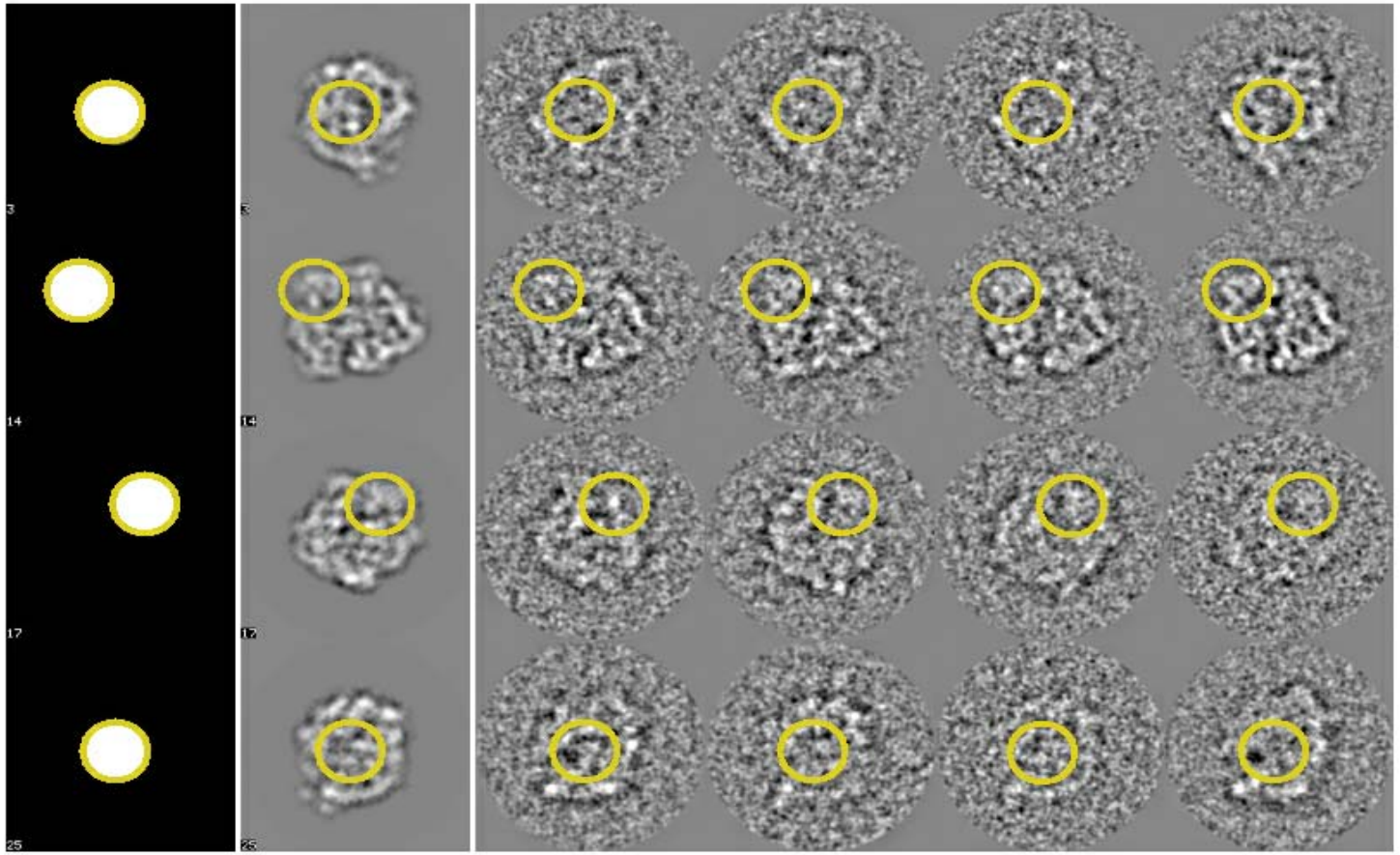
Klaholz, *Open J. Stat.* (2015).

IGBMC

Determining structures of multiple conformational states in a single sample

local 2D MSA

series of MSA's on particle views (class averages or extracted particles):



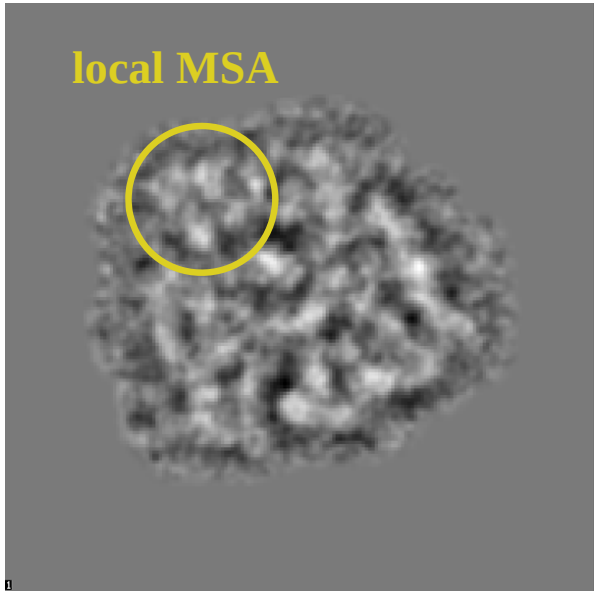
Klaholz *et al.*, *Nature* 2004; see Suppl. Mat.

→ allows re-classification after orientational classification

IGBMC

Determining structures of multiple conformational states in a single sample

local 2D MSA



30S 50S

"seeding" with different structures, from the sample (no external refs)

Image processing procedure

initial (merged) structure:
localise 3D area of disorder

↓

msa of classums-particles on restricted area

↙ ↘ cross-validation of group assignment

3D reconstruction 3D reconstruction

↘ ↙

reproject and append reprojections (reference 1 + 2)

↓ MRA

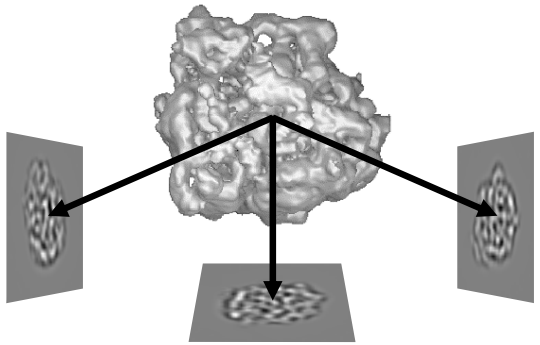
sort particles according to reference number

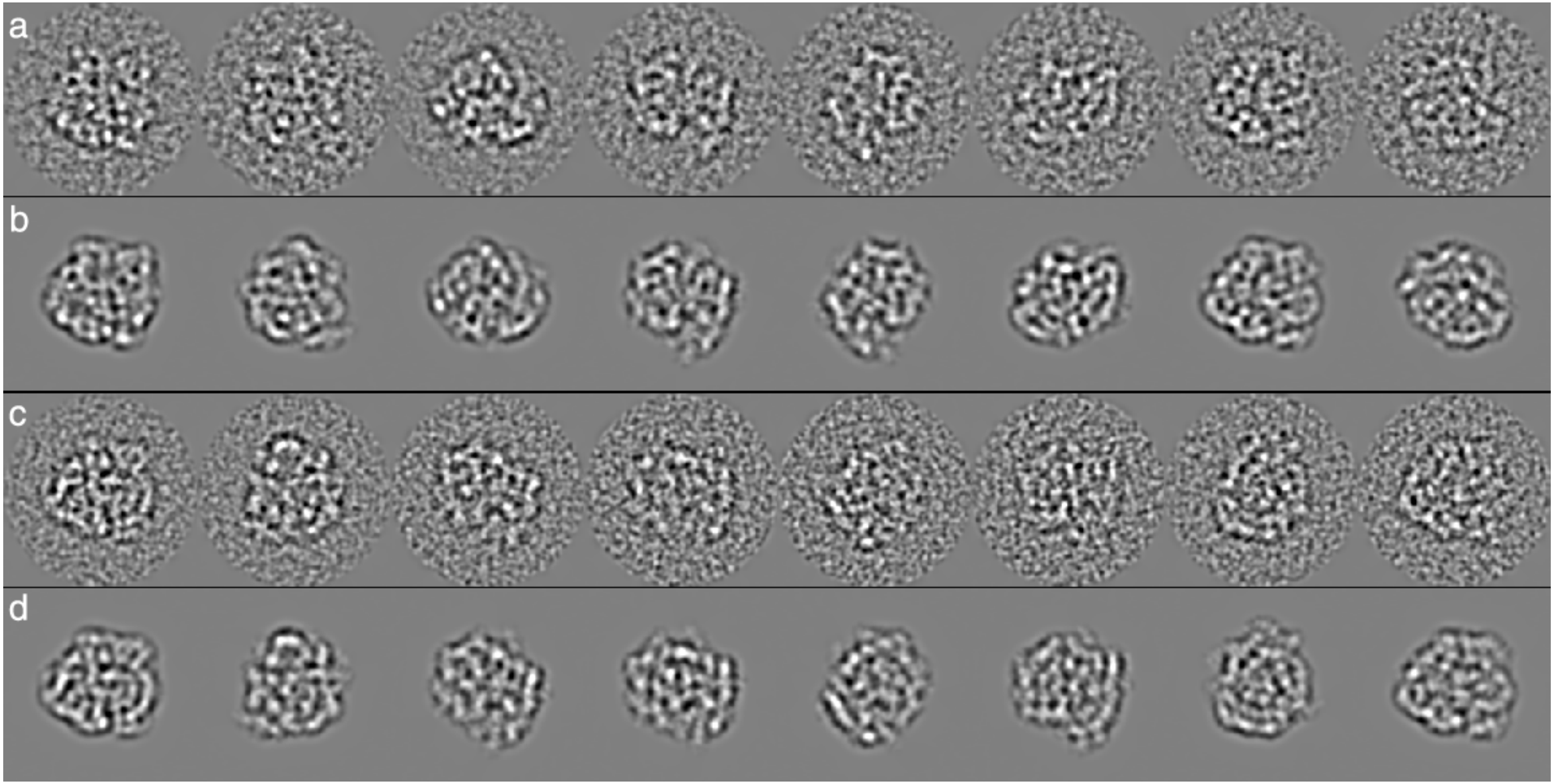
↙ ↘

MSA, AR, 3D MSA, AR, 3D

IGBMC

group assignment by iterative cross-validation with corresponding re-projections



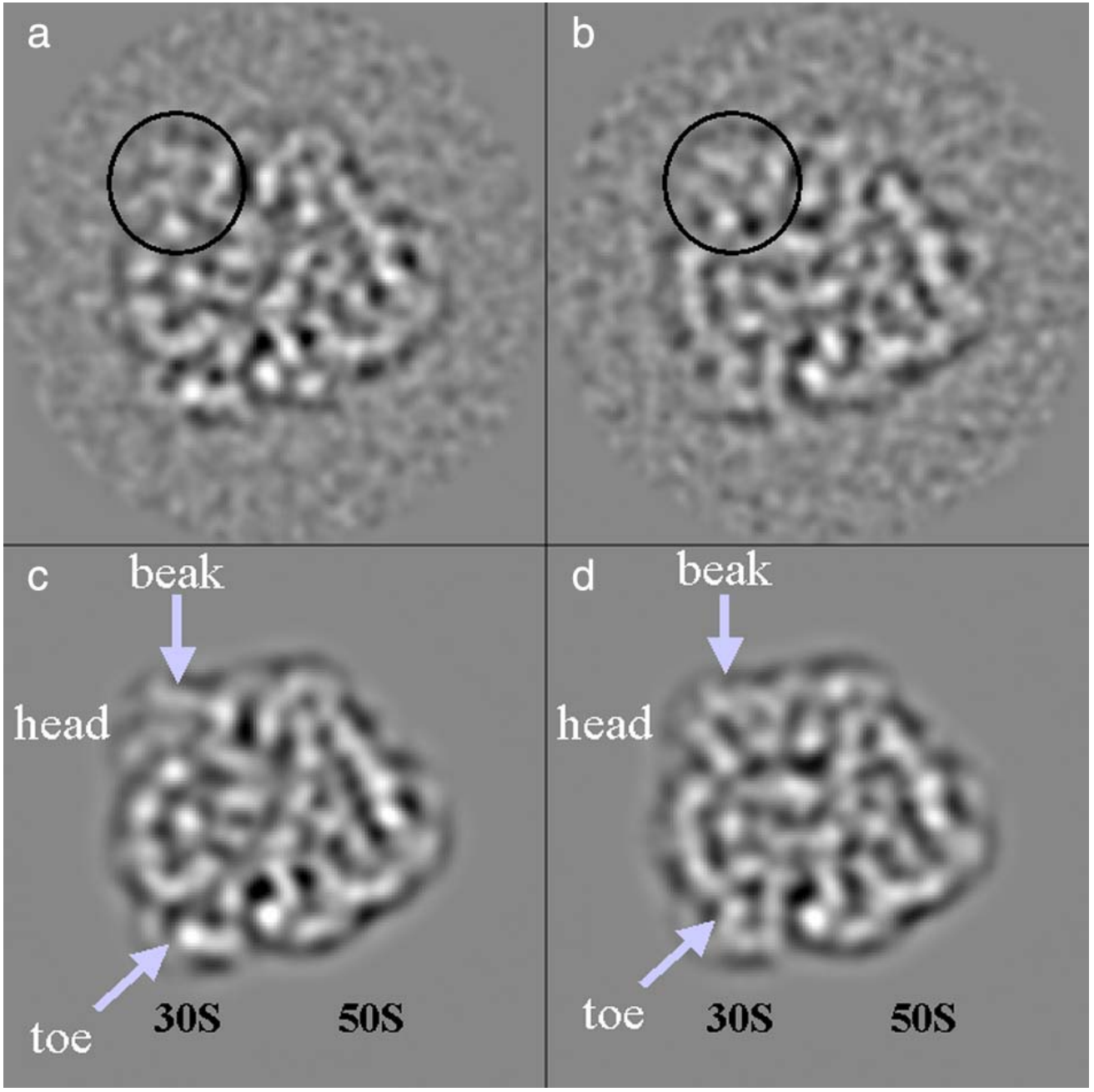


set 1

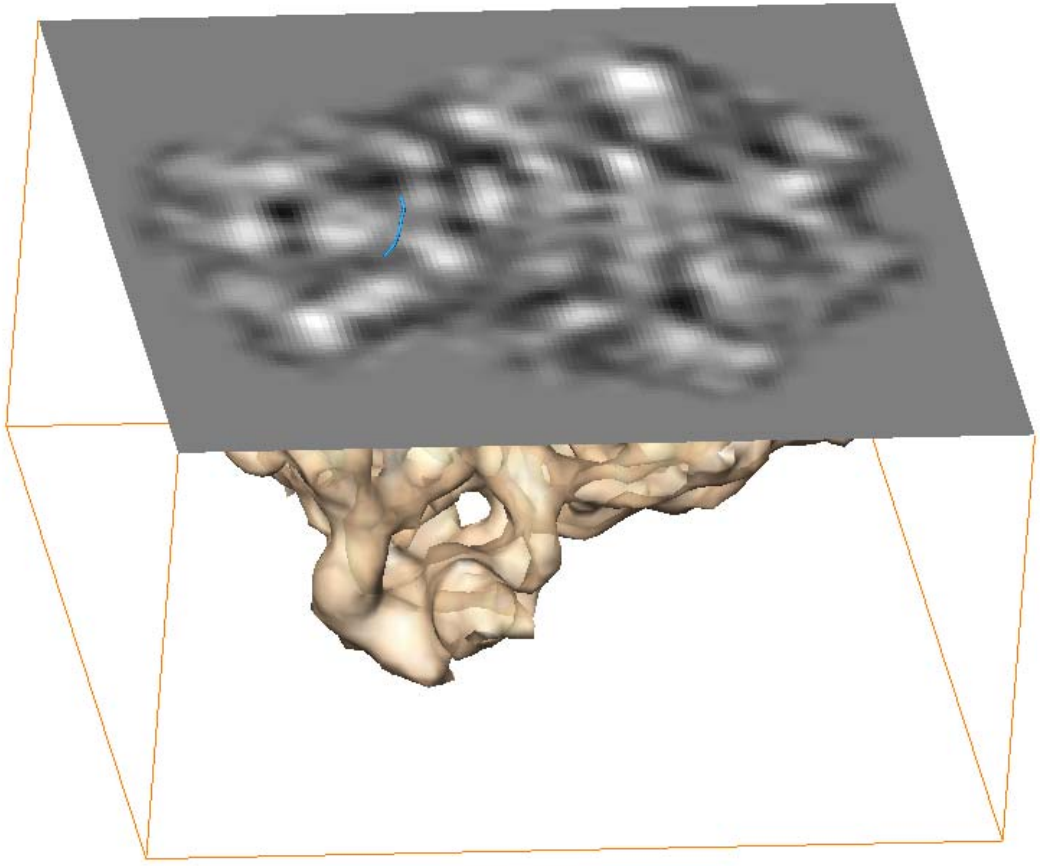
set 2

IGBMC

Procedure applied to distinct areas of the ribosome: conformational changes are correlated:

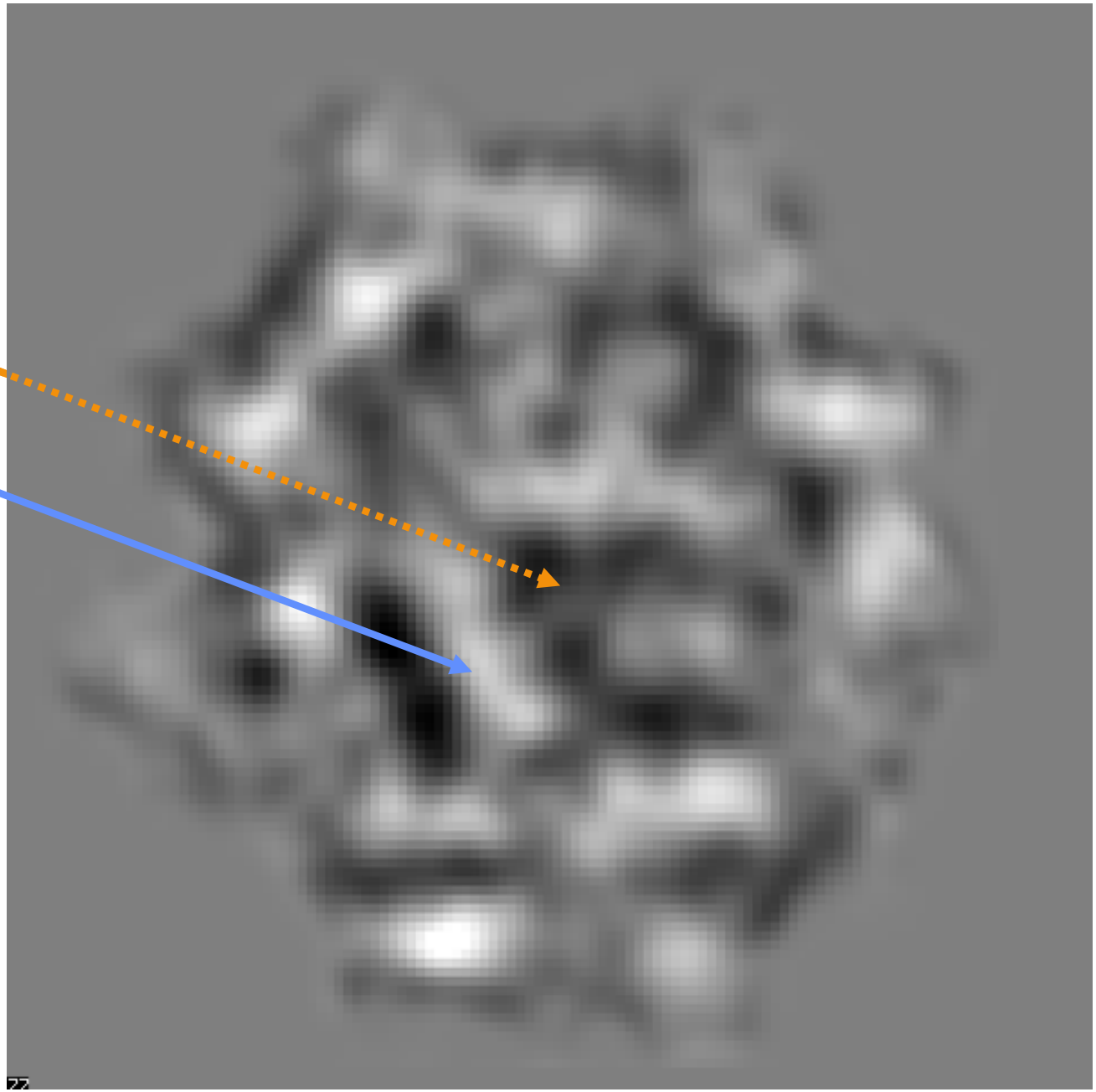


Sections through the 3D map:



P-site tRNA

E-site tRNA



Concept of structure sorting is true at 10-20 Å resolution and needs to be validated at that resolution,

but
also at 2 Å resolution
(e.g. multiple side-chain conformations seen in crystal structures)

tRNA translocation

IGBMC

Intrinsic limitations of 2D-based particle sorting:

(i) usually requires user-knowledge of the structure because some typical molecular views are needed to visually detect structural heterogeneity; solution: use variance map

(ii) it harbours the problem of assigning a particle image to a precise group (i.e. one structural state or another) across different viewing angles (addressed in part by automatic iteration of the cross-validation with re-projections);

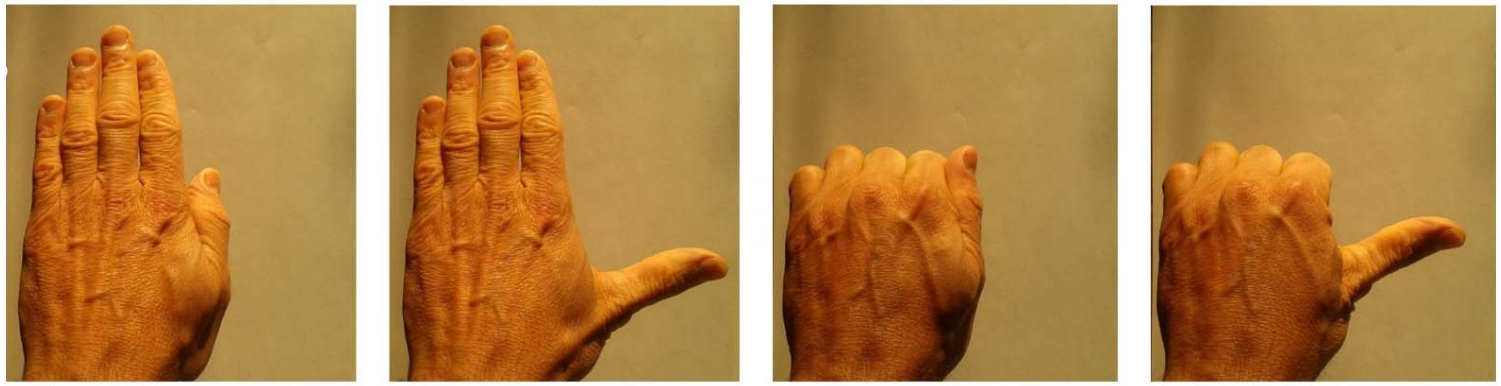
(iii) the procedure is difficult to extend to more than two different states.

MSA-based 3D classification

Determining structures of multiple conformational states in a single sample

Concept of 3D re-sampling and classification (3D-SC)

different conformations of the 3D objects



(here: conformational variability within a given orientation)

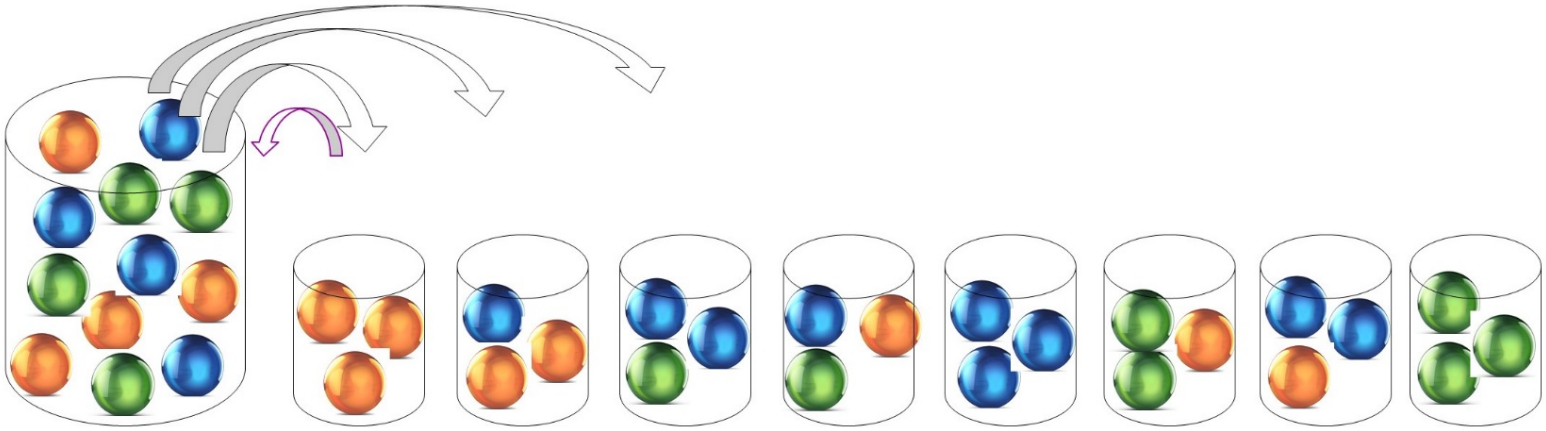
Klaholz, *Open J. Statistics*, 2015.

Determining structures of multiple conformational states in a single sample

Concept of 3D re-sampling and classification (3D-SC)

statistical resampling:

jack-knifing /
bootstrapping
("resampling")



- jack-knifing: selection of small subsets
- bootstrapping: random selection of small subsets, part of which can be re-selected (resampling with replacement;
repeated random resampling is a Monte Carlo approach)

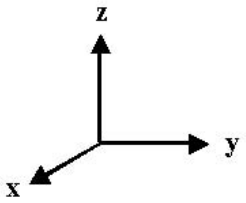
see: Quenouille, 1949; Efron, 1979; Simon, 1969 / 1997; Good, 2005.

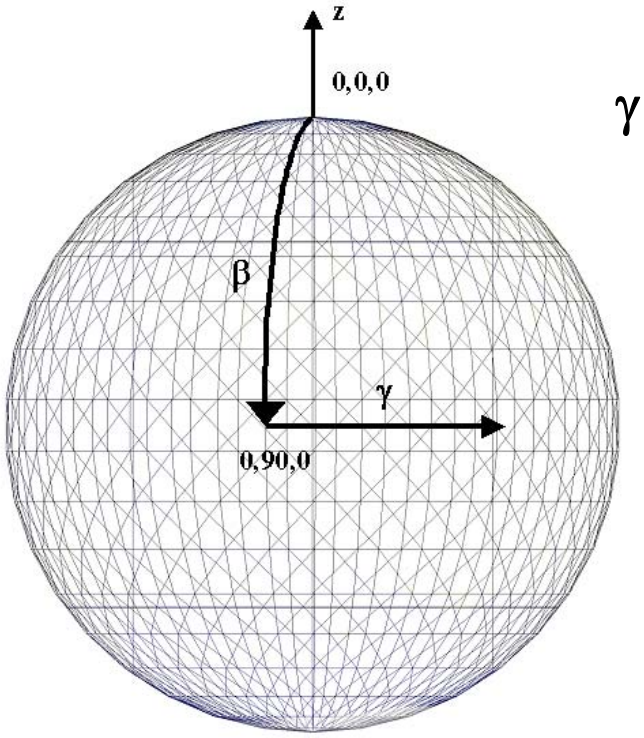
Klaholz, *Open J. Statistics*, 2015.

IGBMC

Determining structures of multiple conformational states in a single sample

3D MSA

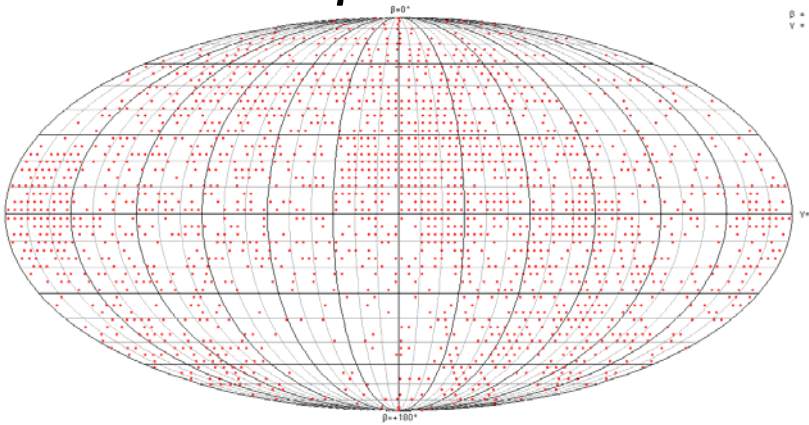




Particle angles plotted on sphere:

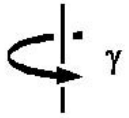
β = 0°

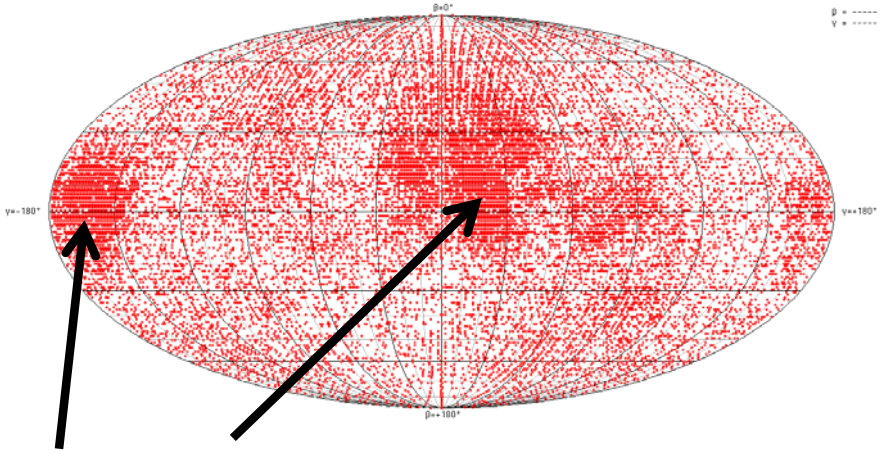
γ = -180°



γ = 180°

β = +180°



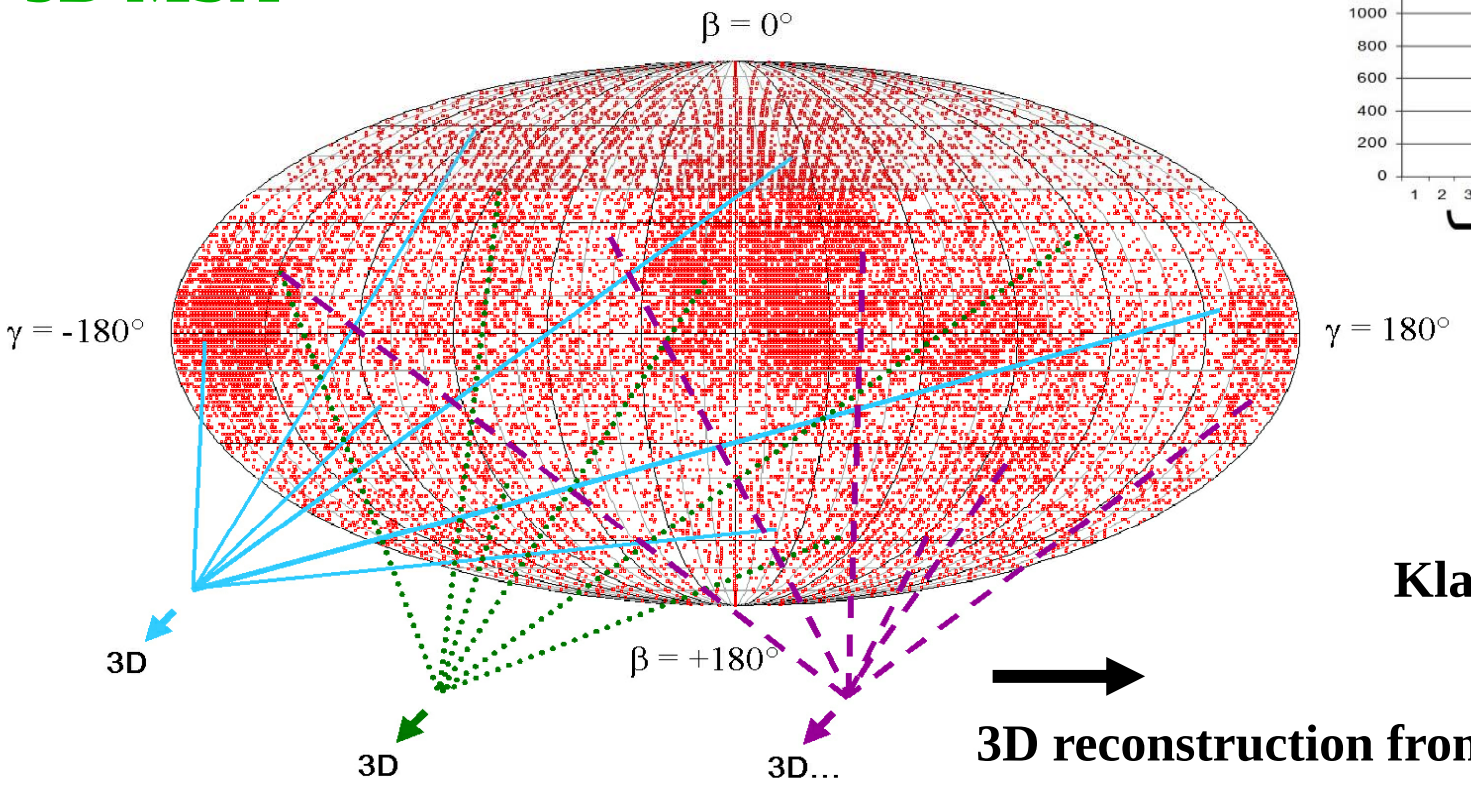


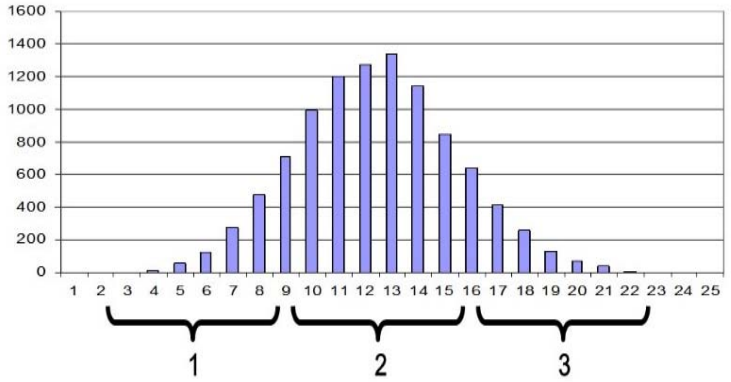
(preferential views)

IGBMC

Determining structures of multiple conformational states in a single sample

3D MSA





Klaholz, *Open J. Statistics*, 2015.

3D reconstruction from many subsets (resampling)
3D statistical analysis and 3D classification:
3D sampling and classification (3D-SC)

→ does both re-sampling and 3D classification, 3D variance map;
see also work by P. Penczek (bootstrapping (re-sampling), used primarily to find region of variance, i.e. estimation of 3D variance)

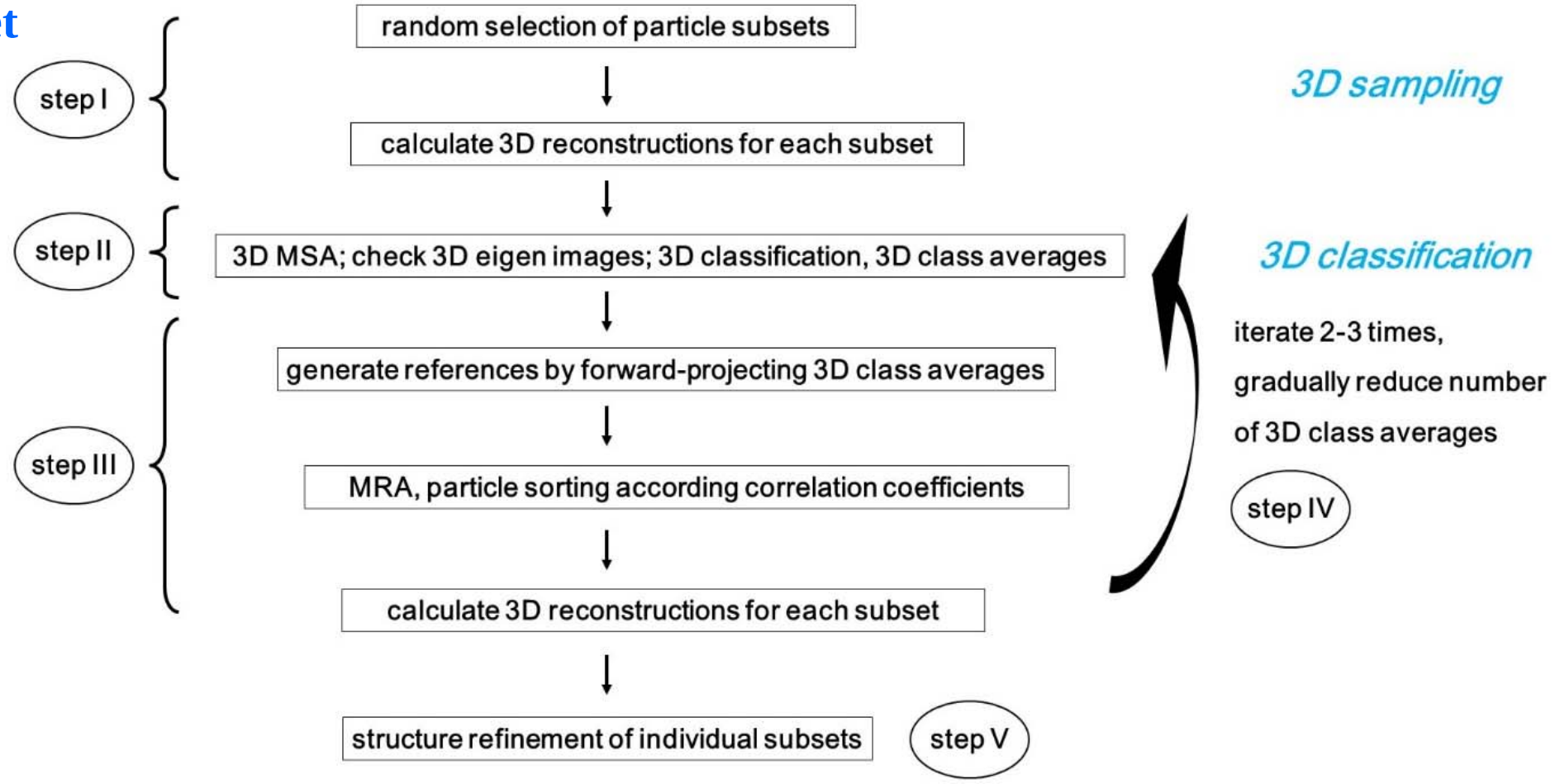
IGBMC

Determining structures of multiple conformational states in a single sample

Concept of 3D sampling and classification (3D-SC)

10-50 particles per set

e.g. 10 000 structures



initially use coarsened (binned) data to speed up the process

Klaholz, *Open J. Statistics*, 2015.

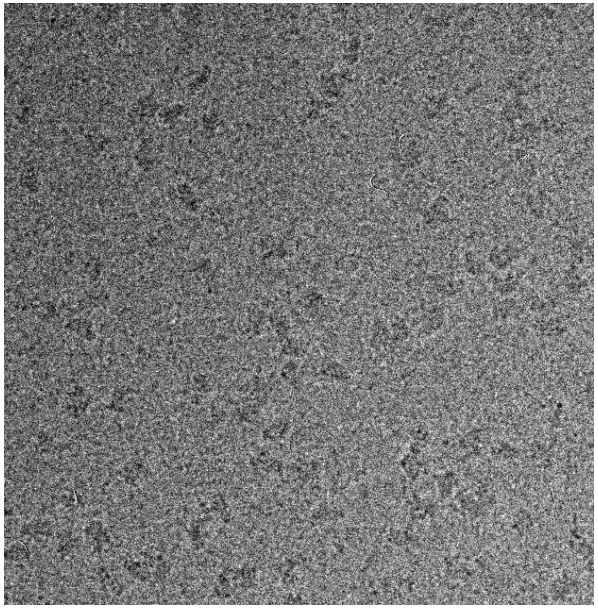
IGBMC

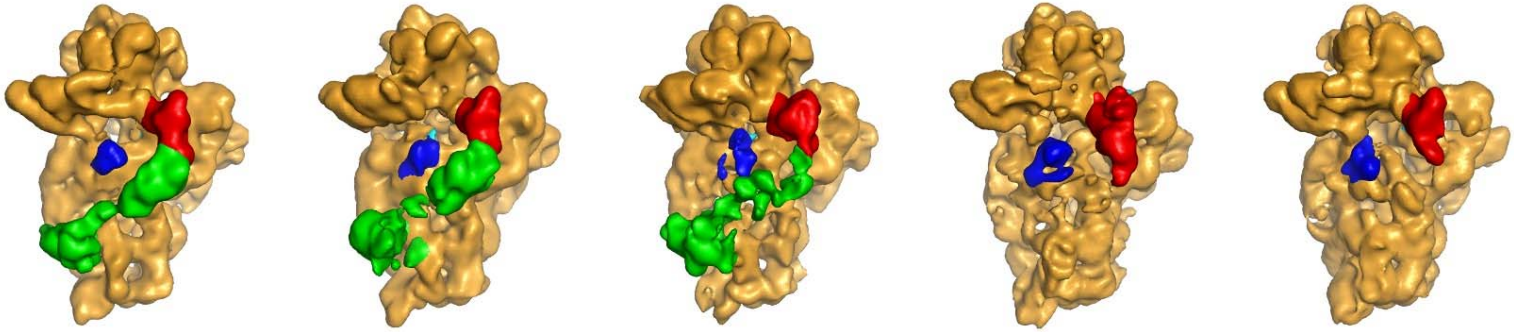
Determining structures of multiple conformational states in a single sample

3D MSA

3D reconstruction from many subsets (resampling),
3D statistical analysis and 3D classification:
3D sampling and classification (3D-SC)

Multiple states in the 30S initiation complex





40 % **28 %** **10 %** **8 %** **14 %**

no mRNA

no IF2

Imagic

200kV FEG data;

total 80 000 particles

resolution of 3D's: 9Å

2D → 3D → 4D

Simonetti *et al.*, *Nature*, 2008.

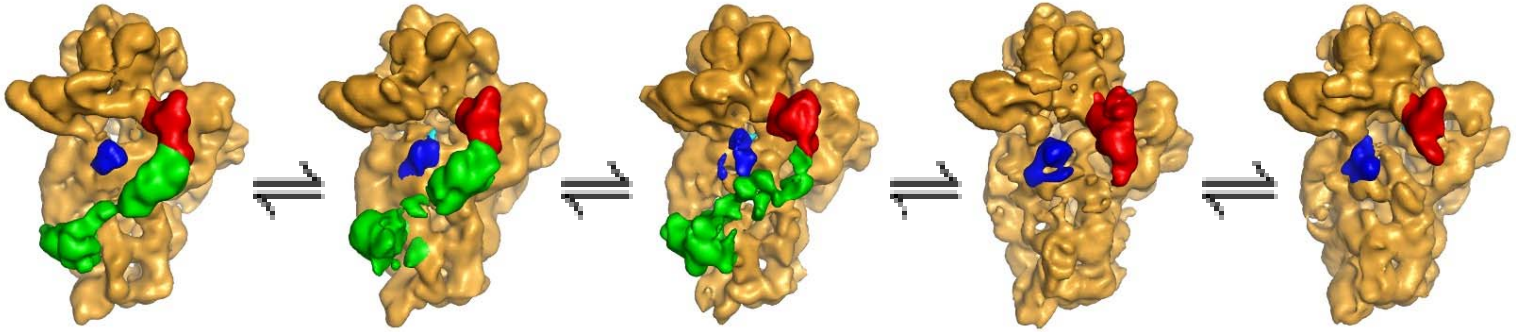
IGBMC

Determining structures of multiple conformational states in a single sample

3D MSA

Addressing the structural state of reaction intermediates that are in equilibrium with each other!

particle populations



$\alpha A + \beta B \rightleftharpoons \sigma S + \tau T$ **40 %** **28 %** **10 %** **8 %** **14 %**

forward reaction rate = $k_+ A^\alpha B^\beta$

backward reaction rate = $k_- S^\omega T^\iota$

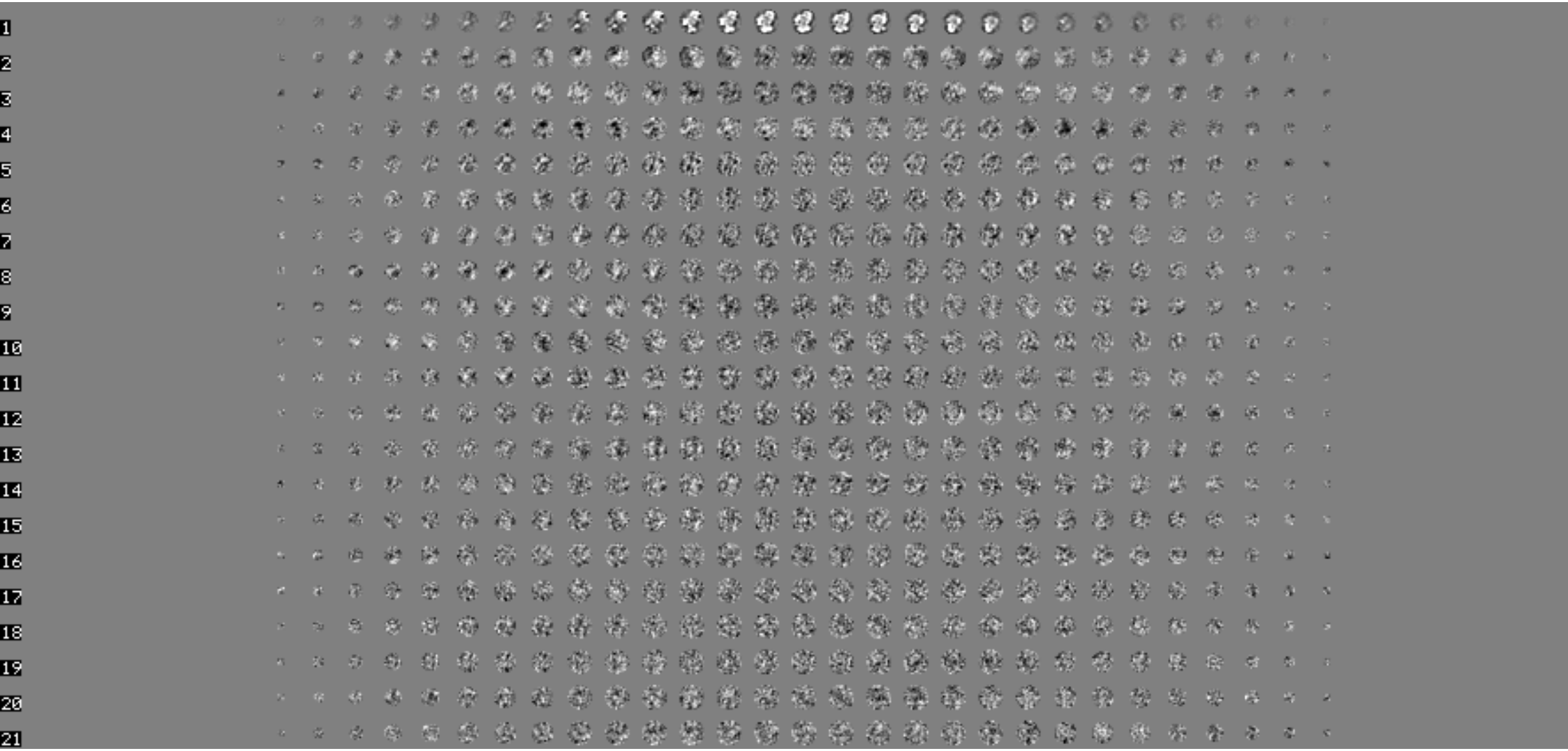
→ equilibrium constant: $K = \frac{k_+}{k_-} = \frac{\{S\}^\sigma \{T\}^\tau}{\{A\}^\alpha \{B\}^\beta}$

The 3D classification procedure 3D-SC has been used since by other groups also:
(Papai *et al.*, *Nature* 2010; Fischer *et al.*, *Nature* 2010). **Simonetti *et al.*, *Nature*, 2008.**

IGBMC

3D MSA

3D eigenimages: (30S)

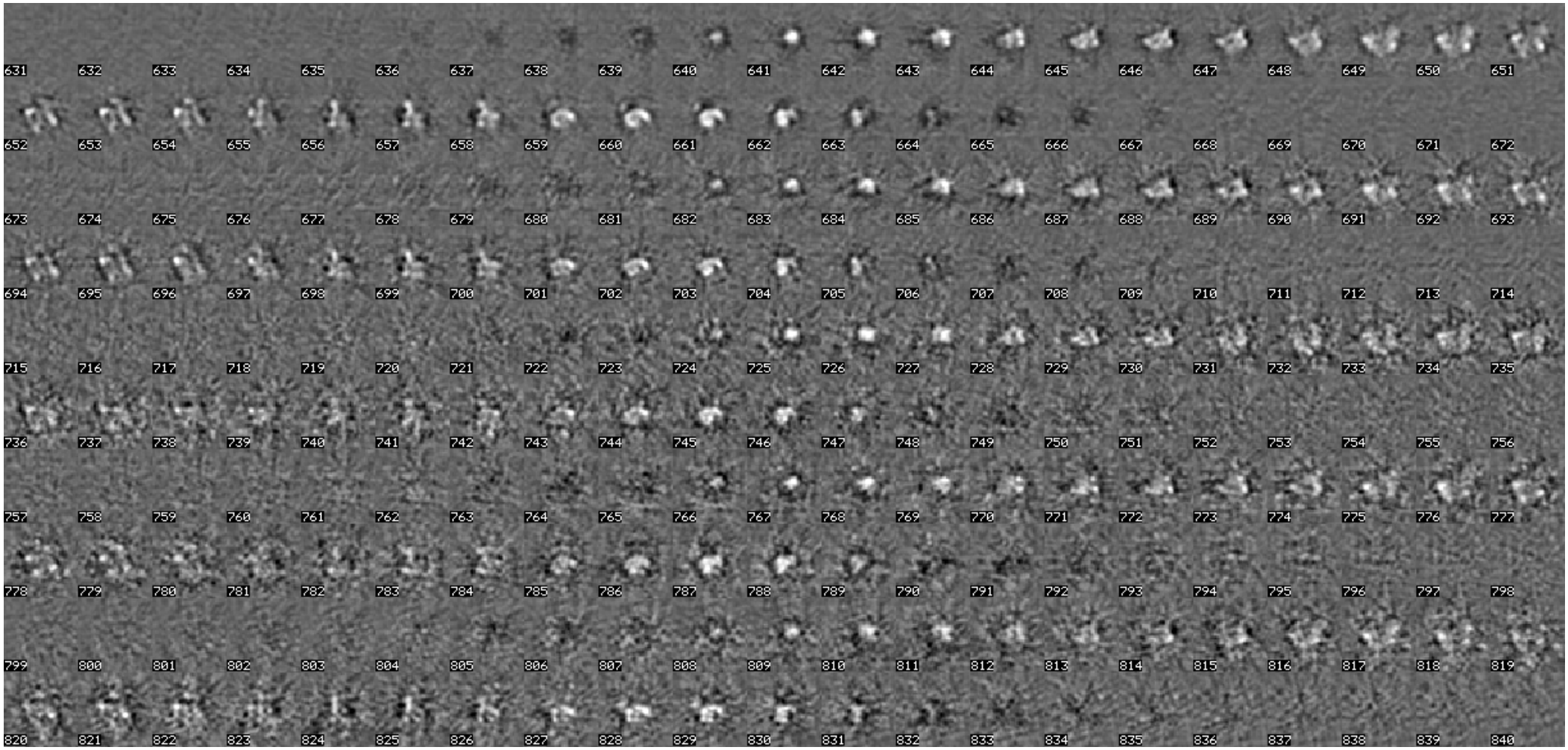


Klaholz, *Open J. Statistics*, 2015.

IGBMC

3D MSA

3D class averages: (30S)

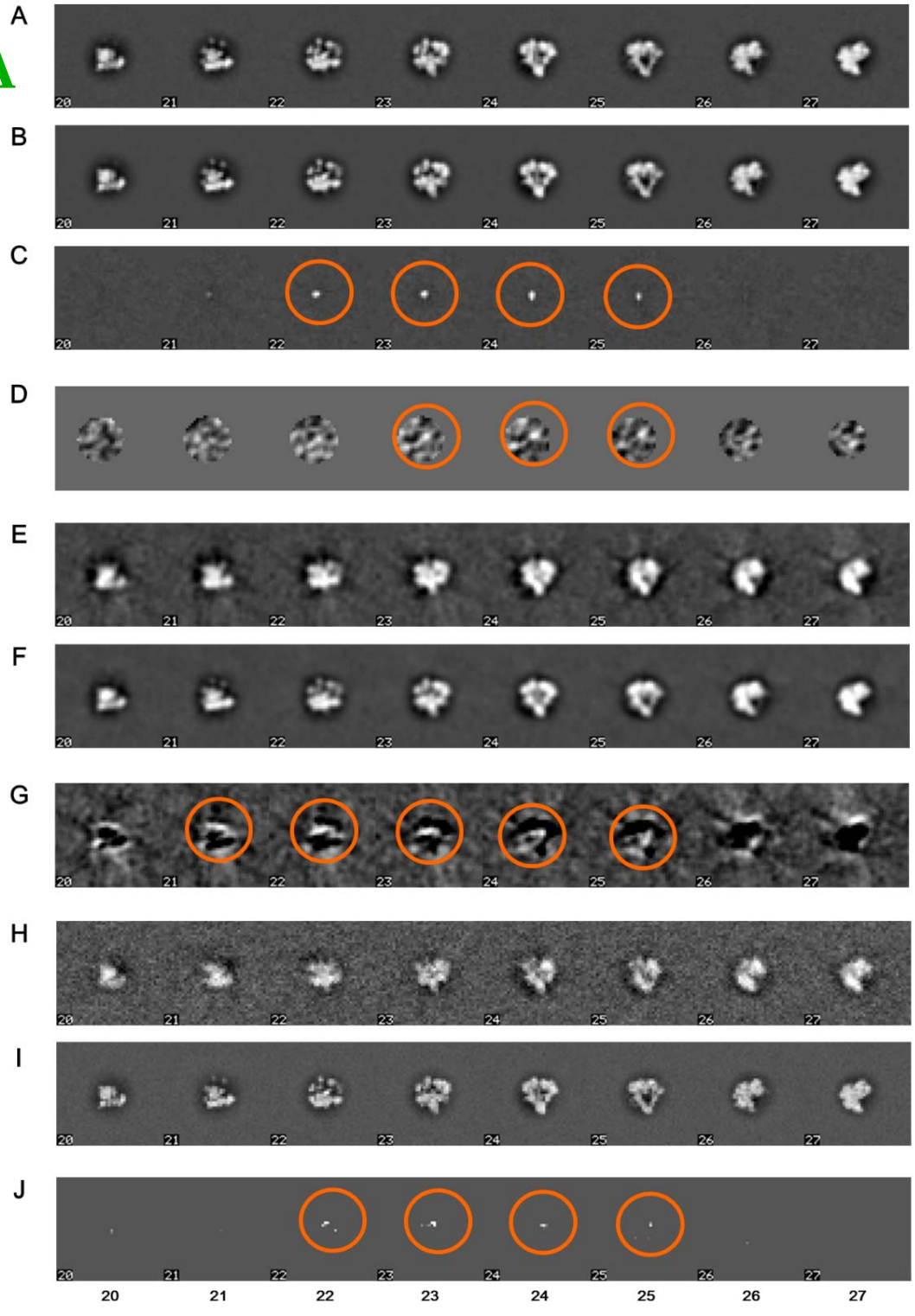


3D #

IGBMC

Klaholz, *Open J. Statistics*, 2015.

3D MSA



small difference found by 3D MSA (<1% of the total mass of the complex; modeled RNA Pol data)

Imagic

Klaholz, *Open J. Statistics*, 2015.

IGBMC

Some examples based on 2D classification or on 3D classification (3D-SC / resampling / bootstrapping):

Klaholz *et al.*, 2004; White *et al.*, 2004 (size variation);

Penczek *et al.*, 2006; Cheng *et al.*, 2007;

Elad *et al.*, 2008; Simonetti *et al.*, 2008;

Wang et al., 2013;

Liao *et al.*, 2015.

focused classification / focused refinement:

concept implemented later into Relion, Frealign etc.

→ analyse regions, subunit structural variations

→ not limited by the number of different structural states in the sample

Determining structures of multiple conformational states in a single sample

3) maximum likelihood based class assignment

→ assign particles to different 3D classes based on maximum likelihood (max. expectation)

(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

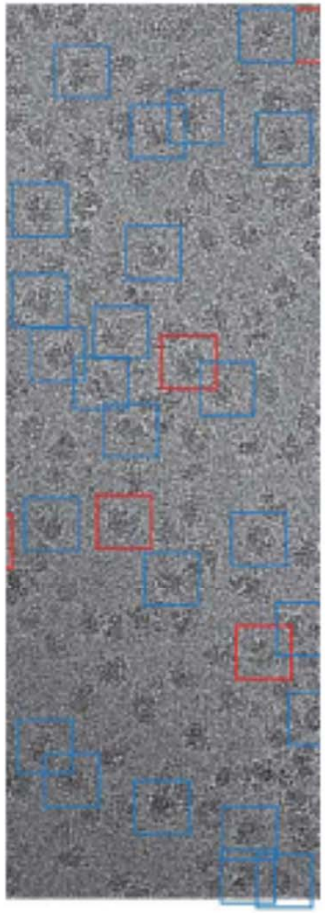
e.g. Scheres *et al.*, *JMB* 2005; *Meth. Enzymol.* 2010;

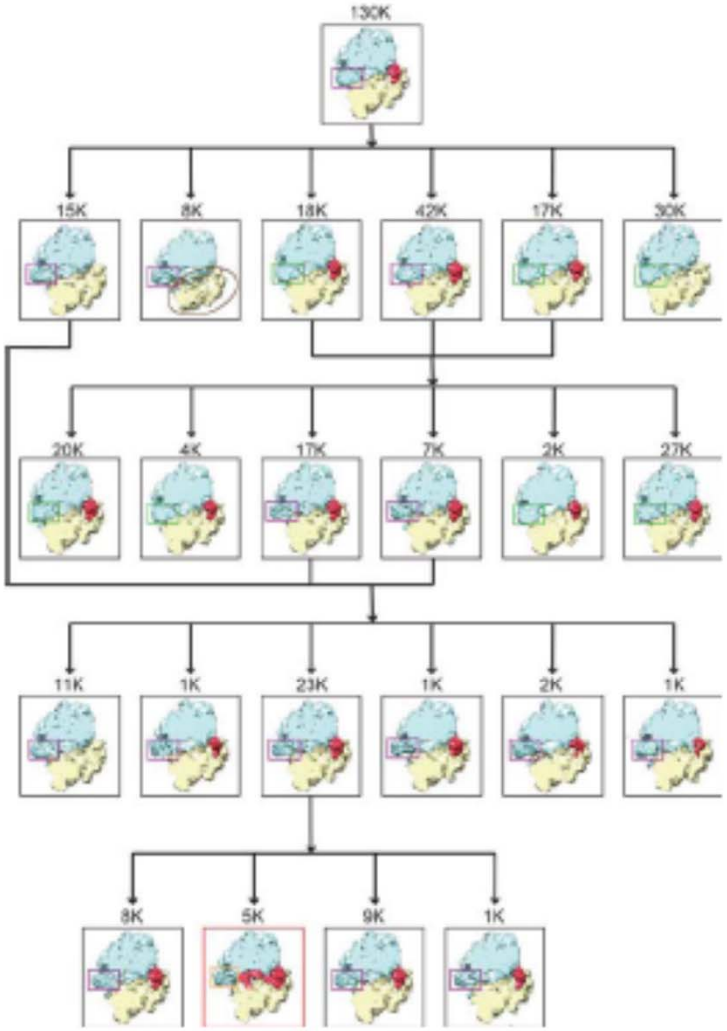
Lyumkis *et al.*, *JSB* 2013

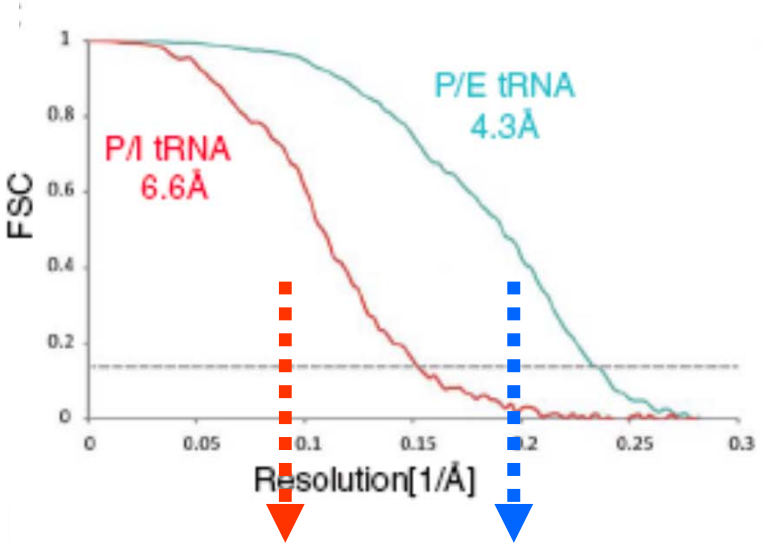
Introduction of the ML concept in cryo-EM: Sigworth, *JSB* 1998;

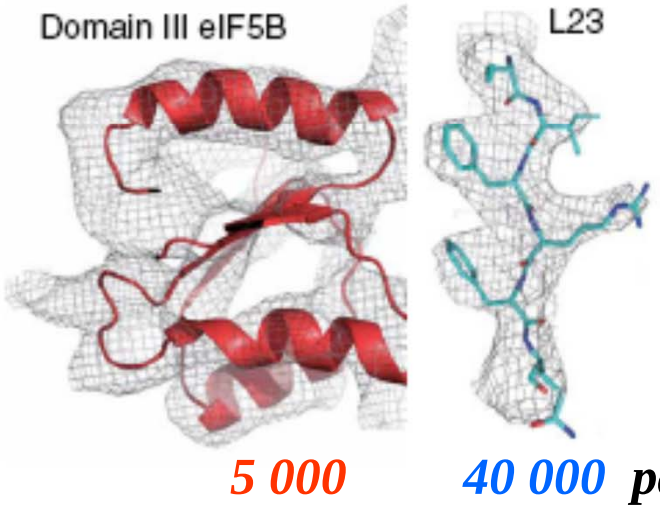
in X-ray crystallography: G. Bricogne, *Acta Cryst A*, 1991.

Examples of ML-based 3D classification









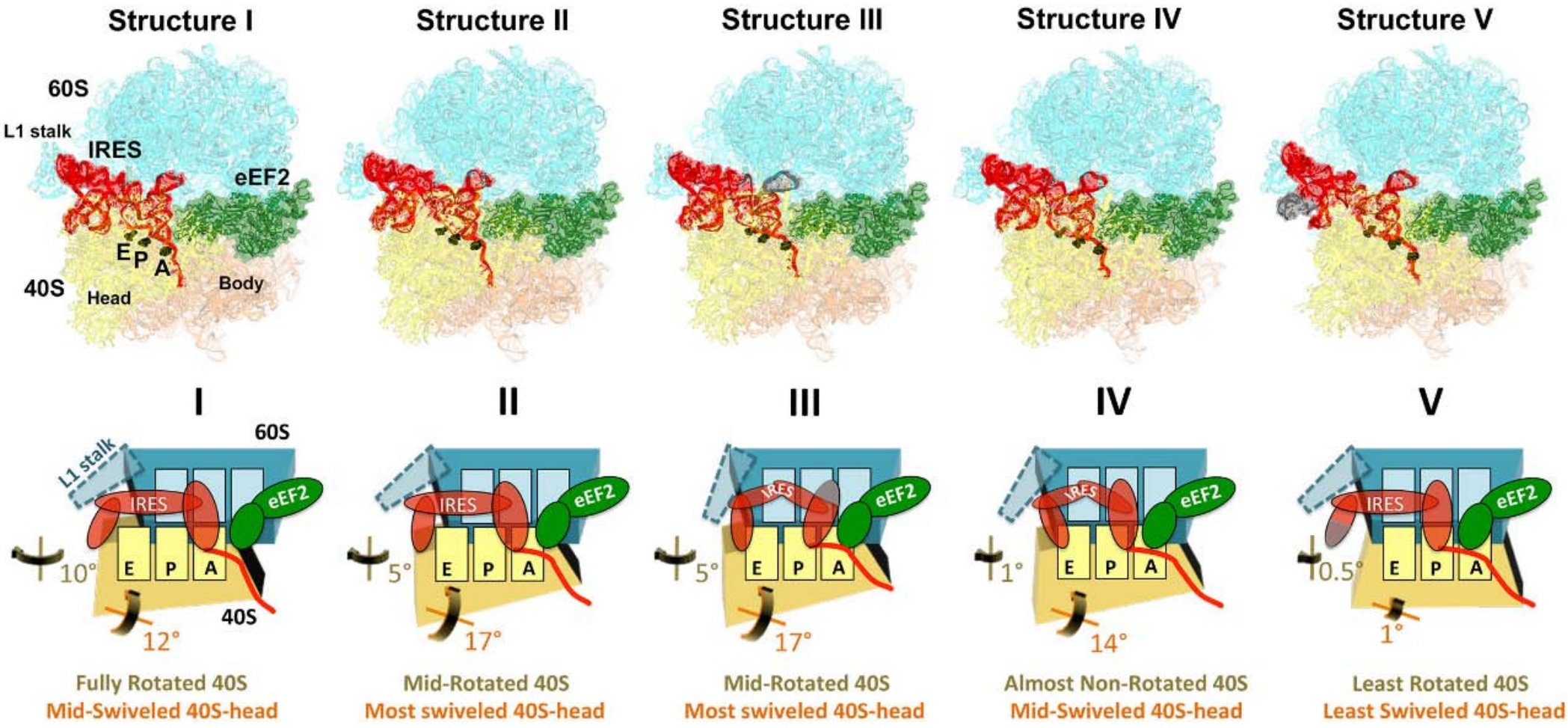
5 000 ***40 000*** ***particles***

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.

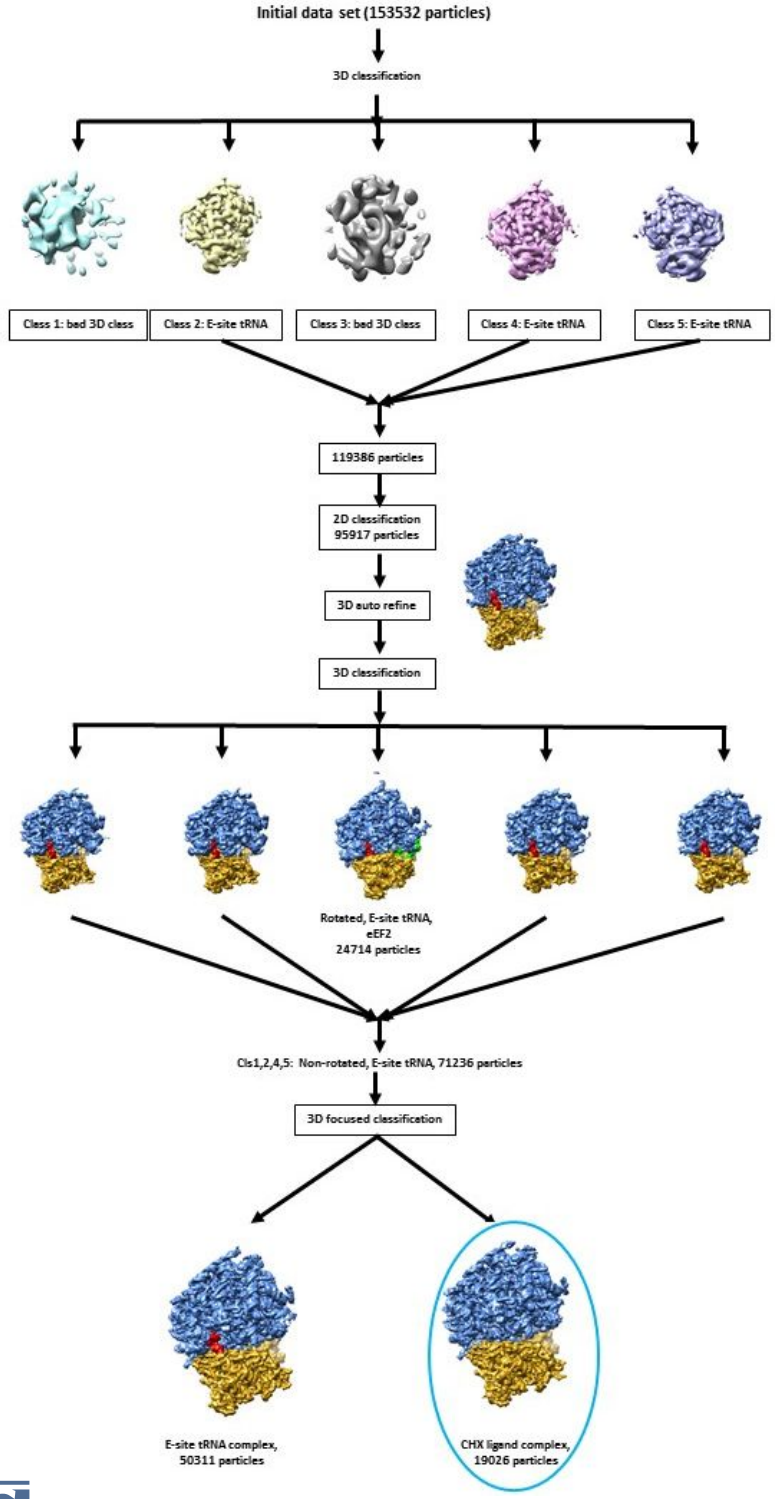
IGBMC

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



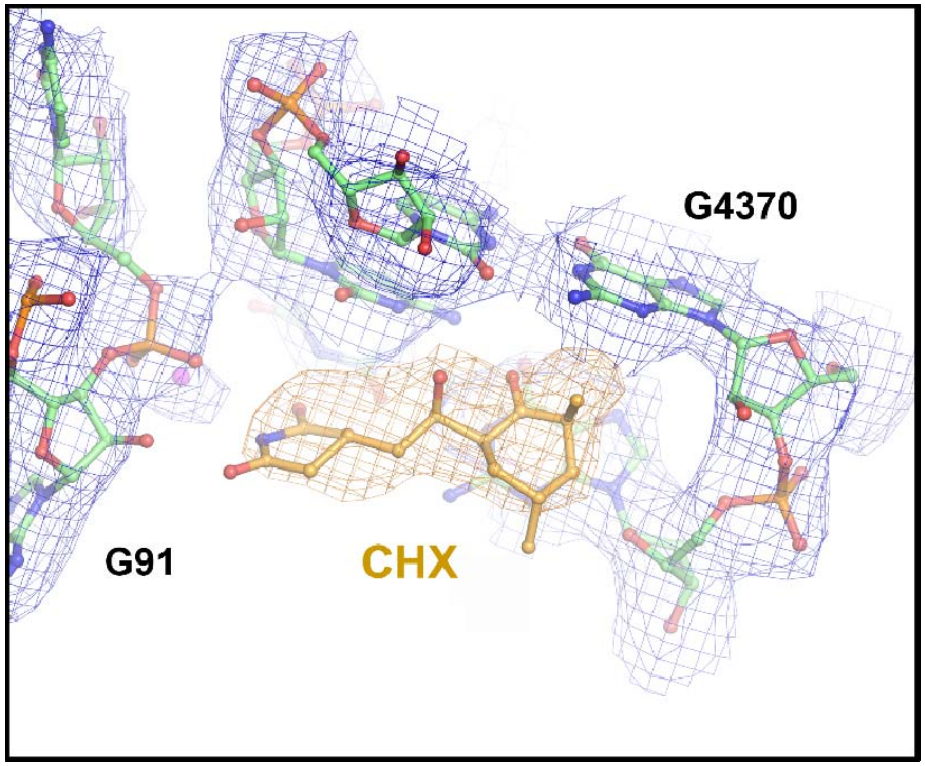
Abeyrathne *et al.*, *eLife* 2016

IGBMC



e.g. ML-based focused classification

sorting scheme for human 80S/antibiotic complex



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

IGBMC

Summary:

Determining structures of multiple conformational states in a single sample

Possibilities to address heterogeneity:

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 class assignment

4) deep learning methods (coming)

see also review by White *et al.*, *Biomed. Res. Int.* 2017, 1032432.

Summary:

Determining structures of multiple conformational states in a single sample

Conclusions & tips:

- do not assume a single state in your sample / multi-subunit complex

- consider lower symmetry (viruses etc.) to see differences between subunits

- if to use local / focused classification/refinement: use slightly larger region

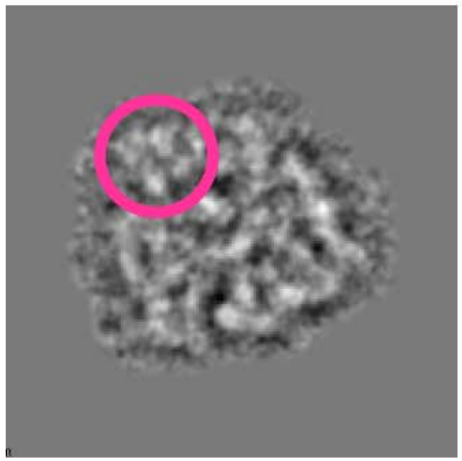
- also useful in sub-tomogram averaging and 3D classifications

- consider: any sub-ensembles will not be entirely homogeneous due to the statistical nature of the procedures (resampling and/or ML)

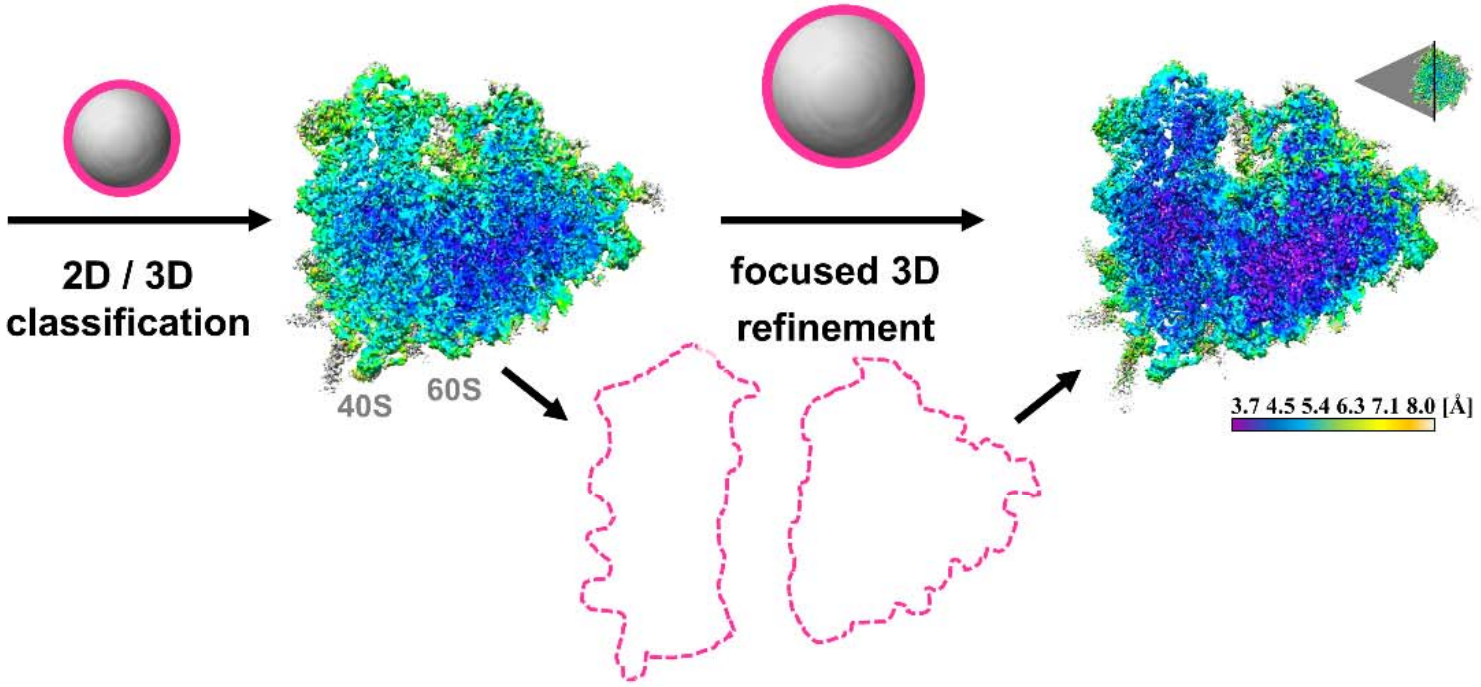
- after 3D classification: go for focused refinement →

IGBMC

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



Concept of local 2D MSA / focused classification



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

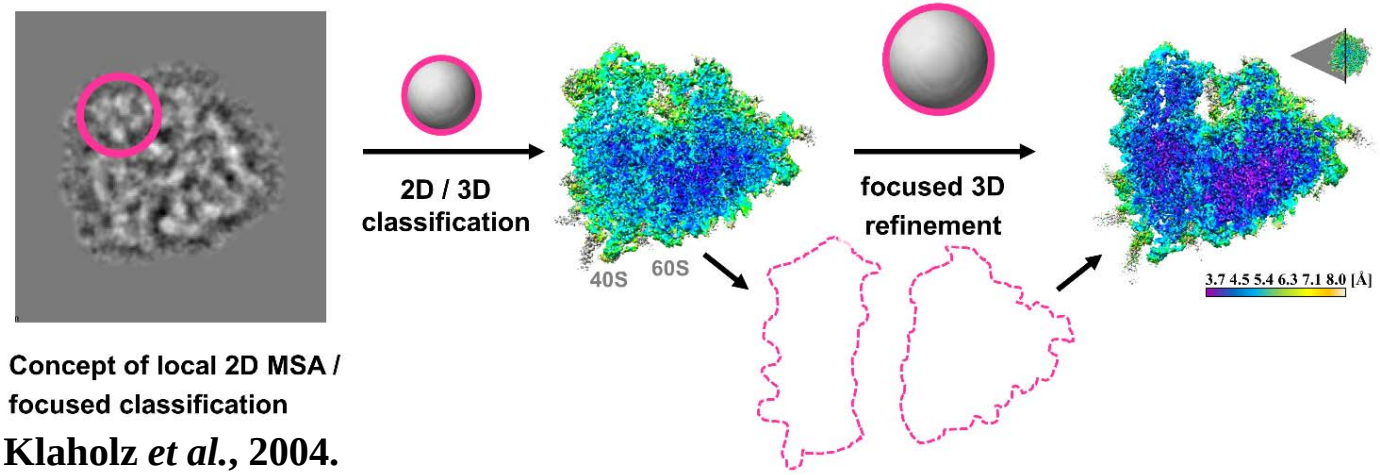
see also:
Klaholz *et al.*, *Nature* 2004;
White *et al.*, *JSB* 2004;
Penczek *et al.*, *JSB* 2006;
Wong *et al.*, *Elife* 2014;
…

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.**

Advanced image processing to improve cryo-EM reconstructions and map interpretation

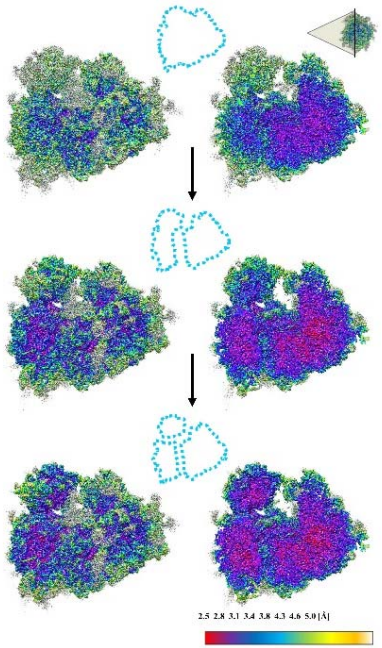
Local MSA / focused 2D/3D classification & focused refinement:



Concept of local 2D MSA / focused classification

Klaholz *et al.*, 2004.

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

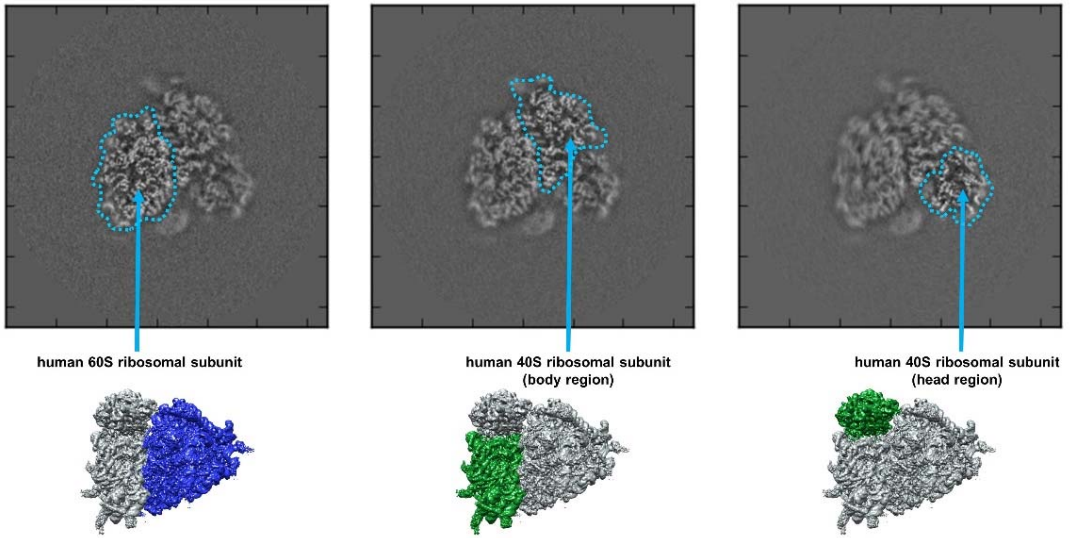


Focused refinement

global 80S refinement

refinement of the individual
60S and 40S ribosomal subunits

refinement of the individual
60S subunit, 40S body and 40S head regions

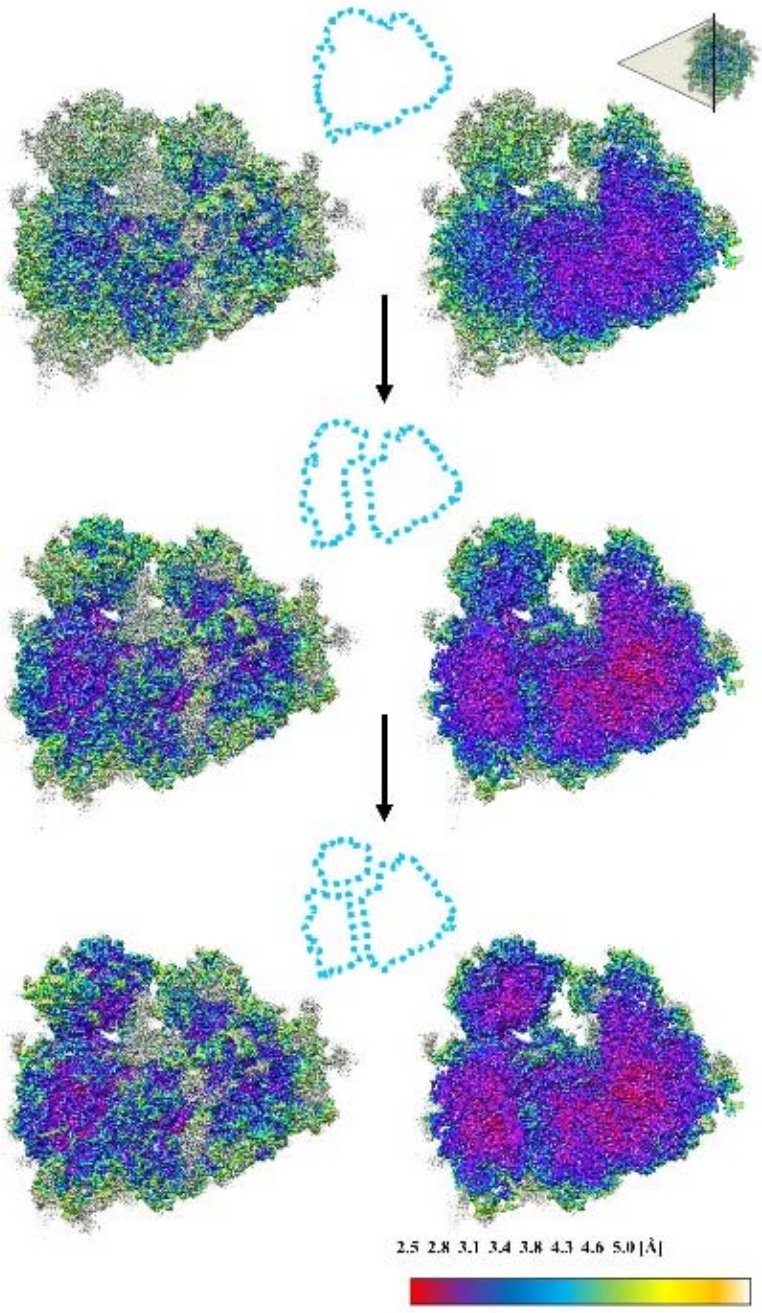


Natchiar *et al.*, *Nature* 2017.

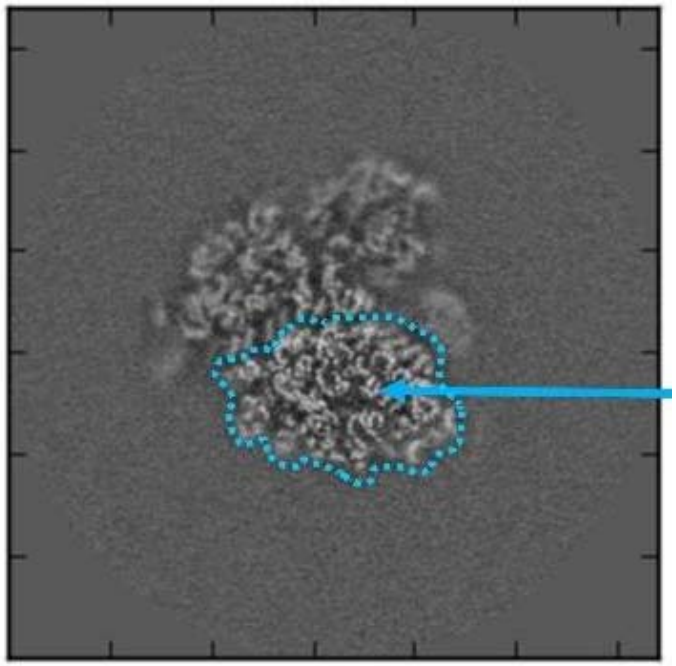
IGBMC

Focused refinement

(with subtraction, Relion)



global 80S refinement



refinement of the individual
60S and 40S ribosomal subunits

refinement of the individual
60S subunit, 40S body and 40S head regions

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

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

Specific tips on focused refinement:

- works best after 3D classification / sorting

→ makes sure that it corresponds to a conformational / functional state, therefore the PDB requires to deposit the low-resolution map before focused refinement

→ provides a composite map

- further improves with partial signal subtraction (Bai *et al.*, 2015)
- localized reconstruction of subunits on viruses (Ilca *et al.*, 2015)
- subunit subtraction and focused refinement on GroEL (Roh *et al.*, 2017)
- re-centering focused region helps (alignment quality); e.g. Blees *et al.*, 2017
- dynamic signal subtraction (Schoebel *et al.*, 2017)
- multi-body refinement (Nakane *et al.*, 2018)

Some references

- Benzécri JP, in: S. Watanabe (Ed.), Methodologies of Pattern Recognition, Academic Press, New York, 1969, pp. 35-74.
- van Heel, M. & Frank, J. (1981). Use of multivariate statistics in analyzing the images of biological macromolecules. Ultramicroscopy, 6, 187-194.

- van Heel, M. (1984). Multivariate statistical classification of noisy images (randomly oriented biological macromolecules). Ultramicroscopy, 13, 165-183. http://dx.doi.org/10.1016/0304-3991(84)90066-4

- Borland, L. & van Heel, M. (1990). Classification of image data in conjugate representation spaces. J. Opt.Soc. Am. A7, 601-610. http://dx.doi.org/10.1364/JOSAA.7.000601

- van Heel, M., Portugal, R., Schatz, M. (2009). DOI: 10.13140/2.1.5192.2729. In book: An electronic text book: Electron microscopy in Life Science, Publisher: 3D-EM Network of Excellence, Editors: A. Verkley, Elena Orlova.

- Leschziner AE, Nogales E. (2007). Visualizing flexibility at molecular resolution: analysis of heterogeneity in single-particle electron microscopy reconstructions. Annu Rev Biophys Biomol Struct., 36, 43-62. http://dx.doi.org/10.1146/annurev.biophys.36.040306.132742

- Brink J, Ludtke SJ, Kong Y, Wakil SJ, Ma J, Chiu W. Experimental verification of conformational variation of human fatty acid synthase as predicted by normal mode analysis. Structure. 2004 Feb; 12(2):185-91. DOI: 10.1016/j.str.2004.01.015

- Gao H, Valle M, Ehrenberg M, Frank J. (2004). Dynamics of EF-G interaction with the ribosome explored by classification of a heterogeneous cryo-EM dataset. J. Struct. Biol., 147, 283-90. http://dx.doi.org/10.1016/j.jsb.2004.02.008

- Sigworth FJ. (1998). A maximum-likelihood approach to single-particle image refinement. J Struct Biol., 122, 328-39. http://dx.doi.org/10.1006/jsbi.1998.4014

- Scheres SH, Valle M, Nuez R, Sorzano CO, Marabini R, Herman GT, Carazo JM. (2005). Maximum-likelihood multi-reference refinement for electron microscopy images. J. Mol. Biol. 22, 139-49. http://dx.doi.org/10.1016/j.jmb.2005.02.031

- Scheres SH. (2010). Classification of structural heterogeneity by maximum-likelihood methods. Methods Enzymol., 482, 295-320. http://dx.doi.org/10.1016/S0076-6879(10)82012-9

- Lyumkis D, Brilot AF, Theobald DL, Grigorieff N. (2013). Likelihood-based classification of cryo-EM images using FREALIGN. J Struct Biol., 183, 377-88. http://dx.doi.org/10.1016/j.jsb.2013.07.005

IGBMC

Some references (cont.)

- Amunts A, Brown A, Toots J, Scheres SH, Ramakrishnan V. (2015). The structure of the human mitochondrial ribosome. Science, 348, 95-8. http://dx.doi.org/10.1126/science.aaa1193

- Khatter H, Myasnikov AG, Natchiar SK, Klaholz BP. (2015). Structure of the human 80S ribosome. Nature, 30, 640-5. http://dx.doi.org/10.1038/nature14427

- Ohi, M., Li, Y., Cheng, Y., Walz, T. (2004). Negative Staining and Image Classification - Powerful Tools in Modern Electron Microscopy. Biol Proced Online, 6, 23-34. http://dx.doi.org/10.1251/bpo70

- Fu, J., Gao, H., Frank, J. (2006). Unsupervised classification of single particles by cluster tracking in multi-dimensional space. J. Struct. Biol., 157, 226-39. http://dx.doi.org/10.1016/j.jsb.2006.06.012

- Herman GT, Kalinowski M. (2008). Classification of heterogeneous electron microscopic projections into homogeneous subsets. Ultramicroscopy, 108, 327-38. http://dx.doi.org/10.1016/j.ultramic.2007.05.005

- Shatsky M, Hall RJ, Nogales E, Malik J, Brenner SE. (2010). Automated multi-model reconstruction from single-particle electron microscopy data. J Struct Biol., 170, 98-108. http://dx.doi.org/10.1016/j.jsb.2010.01.007

- Tang G, Peng L, Baldwin PR, Mann DS, Jiang W, Rees I, Ludtke SJ. (2007). EMAN2: an extensible image processing suite for electron microscopy. J Struct Biol., 157, 38-46. http://dx.doi.org/10.1016/j.jsb.2006.05.009

- Elmlund H, Elmlund D, Bengio S. (2013). PRIME: probabilistic initial 3D model generation for single-particle cryo-electron microscopy. Structure, 21, 1299-306. http://dx.doi.org/10.1016/j.str.2013.07.002

- Liao HY, Hashem Y, Frank J. (2015). Efficient estimation of three-dimensional covariance and its application in the analysis of heterogeneous samples in cryo-electron microscopy. Structure, 23, 1129-37. http://dx.doi.org/10.1016/j.str.2015.04.004

- Wang Q, Matsui T, Domitrovic T, Zheng Y, Doerschuk PC, Johnson JE. (2013). Dynamics in cryo EM reconstructions visualized with maximum-likelihood derived variance maps. J Struct Biol., 181, 195-206. http://dx.doi.org/10.1016/j.jsb.2012.11.005

- Klaholz, B. P., Myasnikov, A. G., & van Heel, M. (2004). Visualization of release factor 3 on the ribosome during termination of protein synthesis. Nature, 427, 862-5. http://dx.doi.org/10.1038/nature02332

- Penczek, P. A, Frank, J., Spahn, C. M. (2006a). A method of focused classification, based on the bootstrap 3D variance analysis, and its application to EF-G-dependent translocation. J. Struct. Biol., 154, 184-94. http://dx.doi.org/10.1016/j.jsb.2005.12.013

- White, H. E., Saibil HR, Ignatiou A, Orlova EV. (2004). Recognition and separation of single particles with size variation by statistical analysis of their images. J Mol Biol., 13, 453-60. http://dx.doi.org/10.1016/j.jmb.2003.12.015

- Orlova EV, Saibil HR. (2010). Methods for three-dimensional reconstruction of heterogeneous assemblies. Methods Enzymol., 482, 321-41. http://dx.doi.org/10.1016/S0076-6879(10)82013-0

IGBMC

Some references (cont.)

- Elad N, Clare DK, Saibil HR, Orlova EV. (2008). Detection and separation of heterogeneity in molecular complexes by statistical analysis of their two-dimensional projections. J Struct Biol., 162, 108-20. http://dx.doi.org/10.1016/j.jsb.2007.11.007

- De Carlo S, Carles C, Riva M, Schultz P. (2003). Cryo-negative staining reveals conformational flexibility within yeast RNA polymerase I. J Mol Biol., 329, 891-902. http://dx.doi.org/10.1016/S0022-2836(03)00510-2

- Cheng A, Yeager M. (2007). Bootstrap resampling for voxel-wise variance analysis of three-dimensional density maps derived by image analysis of two-dimensional crystals. J Struct Biol., 158, 19-32. http://dx.doi.org/10.1016/j.jsb.2006.10.003

- Myasnikov, A. G., Marzi, S., Simonetti, A., Giuliodori, A. M., Gualerzi, C. O., Yusupova, G., Yusupov, M., Klaholz, B. P. (2005). Conformational transition of initiation factor 2 from the GTP- to GDP-bound state visualized on the ribosome. Nat Struct Mol Biol, 12, 1145-1149. http://dx.doi.org/10.1038/nsmb1012

- Penczek, P. A., Yang, C., Frank, J., Spahn, C. M. (2006b). Estimation of variance in single-particle reconstruction using the bootstrap technique. J. Struct. Biol., 154, 168-83. http://dx.doi.org/10.1016/j.jsb.2006.01.003

- Zhang W, Kimmel M, Spahn CM, Penczek PA. (2008). Heterogeneity of large macromolecular complexes revealed by 3D cryo-EM variance analysis. Structure, 16, 1770-6. http://dx.doi.org/10.1016/j.str.2008.10.011

- Quenouille, M.H. (1949). Approximate tests of correlation in time series. J. Roy. Statist. Soc. B., 11, 68–84. http://dx.doi.org/10.1017/s0305004100025123

- Efron, B. (1979). Bootstrap methods: Another look at the jackknife. The Annals of Statistics, 7, 1-26. http://dx.doi.org/10.1214/aos/1176344552

- Efron, B. (1981). Nonparametric estimates of standard error: The jackknife, the bootstrap and other methods. Biometrika, 68, 589-599. http://dx.doi.org/10.1093/biomet/68.3.589

- Simon J.L. (1969). Basic Research Methods in Social Sciences: the Art of Empirical Investigation. Random House, New York.

- Simon J.L. (1997). Resampling: The new statistics (2nd edition). Duxbury, MA, Thompson International.

- Good, P. (2005). Introduction to Statistics Through Resampling Methods and R/S-PLUS. Wiley. http://dx.doi.org/10.1002/9780471722502

Some references (cont.)

- Simonetti, A., Marzi, S., Myasnikov, A. G., Fabbretti, A., Yusupova, G., Yusupov, M., Gualerzi, C. O., Klaholz, B. P. (2008). Structure of the 30S translation initiation complex. Nature, 455, 416-20. http://dx.doi.org/10.1038/nature07192

- Kettenberger, H, Armache KJ, Cramer P. (2004). Complete RNA polymerase II elongation complex structure and its interactions with NTP and TFIIS. Mol Cell, 16, 955-65. http://dx.doi.org/10.1016/j.molcel.2004.11.040

- Haynor, D.R., Woods, S.D. (1989). Resampling estimates of precision in emission tomography, IEEE Trans. Med. Imaging, 8, 337–343. http://dx.doi.org/10.1109/42.41486

- Maitra, R., (1998). An approximate bootstrap technique for variance estimation in parametric images. Med. Image Anal., 2, 379–382. http://dx.doi.org/10.1016/S1361-8415(98)80018-2

- Voorhees RM, Fernández IS, Scheres SH, Hegde RS. (2014). Structure of the mammalian ribosome-Sec61 complex to 3.4 Å resolution. Cell, 157, 1632-43. http://dx.doi.org/10.1016/j.cell.2014.05.024

- Bai XC, Fernandez IS, McMullan G, Scheres SH. (2013). Ribosome structures to near-atomic resolution from thirty thousand cryo-EM particles. eLife.;2:e00461. doi: 10.7554/eLife.00461. http://dx.doi.org/10.7554/eLife.00461

- Scheres SH, Gao H, Valle M, Herman GT, Eggermont PP, Frank J, Carazo JM. (2007). Disentangling conformational states of macromolecules in 3D-EM through likelihood optimization. Nat Methods, 4, 27-9. http://dx.doi.org/10.1038/nmeth992

- Orlov I, Rochel N, Moras D, Klaholz BP. (2012). Structure of the full human RXR/VDR nuclear receptor heterodimer complex with its DR3 target DNA. EMBO J., 31, 291-300. http://dx.doi.org/10.1038/emboj.2011.445

- Papai G, Tripathi MK, Ruhlmann C, Layer JH, Weil PA, Schultz P. (2010). TFIIA and the transactivator Rap1 cooperate to commit TFIID for transcription initiation. Nature, 465, 956-60. http://dx.doi.org/10.1038/nature09080

- Simonetti A, Marzi S, Billas IM, Tsai A, Fabbretti A, Myasnikov AG, Roblin P, Vaiana AC, Hazemann I, Eiler D, Steitz TA, Puglisi JD, Gualerzi CO, Klaholz BP. (2013). Involvement of protein IF2 N domain in ribosomal subunit joining revealed from architecture and function of the full-length initiation factor. Proc Natl Acad Sci U S A, 110, 15656-61. http://dx.doi.org/10.1073/pnas.1309578110

- Fischer N, Konevega AL, Wintermeyer W, Rodnina MV, Stark H. (2010). Ribosome dynamics and tRNA movement by time-resolved electron cryomicroscopy. *Nature*, 466, 329-33. http://dx.doi.org/10.1038/nature09206

- Behrmann E, Loerke J, Budkevich TV, Yamamoto K, Schmidt A, Penczek PA, Vos MR, Bürger J, Mielke T, Scheerer P, Spahn CM. (2015). Structural snapshots of actively translating human ribosomes. Cell, 161, 845-57. http://dx.doi.org/10.1016/j.cell.2015.03.052

- DePristo M.A., de Bakker P.I.W., & T.L. Blundell (2004). Heterogeneity and inaccuracy in protein structures solved by X-ray crystallography, Structure, 12, 831–838. http://dx.doi.org/10.1016/j.str.2004.02.031

IGBMC

Some references (cont.)

- Levin, E. J., Kondrashov DA, Wesenberg GE, Phillips GN Jr. (2007). Ensemble refinement of protein crystal structures: validation and application. Structure, 15, 1040-52. http://dx.doi.org/10.1016/j.str.2007.06.019

- Klaholz, B. P., Pape, T., Zavialov, A. V., Myasnikov, A. G., Vestergaard, B., Orlova, E., Ehrenberg, M. & van Heel, M. (2003). Structure of the Escherichia coli ribosomal termination complex with release factor 2. Nature, 421, 90-94. http://dx.doi.org/10.1038/nature01225

- Marzi, S., Myasnikov, A. G., Serganov, A., Ehresmann, C., Romby P., Yusupov, M. & Klaholz, B. P. (2007). Structured mRNAs regulate translation initiation by binding to the platform of the ribosome. Cell, 130, 1019–1031. http://dx.doi.org/10.1016/j.cell.2007.07.008

- O. von Loeffelholz, S. K. Natchiar, N. Djabeur, A. G. Myasnikov, H. Kratzat, J.-F. Ménétret, I. Hazemann & B. P. Klaholz. Focused classification and refinement in high-resolution cryo-EM structural analysis of ribosome complexes. Curr. Opin. Struct. Biol., 2017, 46, 140-148. http://dx.doi.org/10.1016/j.sbi.2017.07.007

- S. K. Natchiar, A. G. Myasnikov, H. Kratzat, I. Hazemann & B. P. Klaholz. Visualization of chemical modifications in the human 80S ribosome structure. Nature, 2017, 551, 472-477. http://dx.doi.org/10.1038/nature24482

- I. Orlov, A. G. Myasnikov, L. Andronov, S. K. Natchiar, H. Khatter, B. Beinsteiner, J-F. Ménétret, I. Hazemann, K. Mohideen, K. Tazibt, R. Tabaroni, H. Kratzat, N. Djabeur, T. Bruxelles, F. Raivoniaina, L. di Pompeo, M. Torchy, I. Billas, A. Urzhumtsev & B. P. Klaholz. The integrative role of cryo electron microscopy in molecular and cellular structural biology. Biol Cell., 2017, 109, 81-93. http://dx.doi.org/10.1111/boc.201600042

- Bai, X.-c., Rajendra, E., Yang, G., Shi, Y., Scheres, S. H., Dec. 2015. Sampling the conformational space of the catalytic subunit of human-secretase. eLife 4, e11182.

- Nakane T, Kimanius D, Lindahl E, Scheres SH. Characterisation of molecular motions in cryo-EM single-particle data by multi-body refinement in RELION. Elife. 2018 Jun 1;7. pii: e36861. http://dx.doi.org/10.7554/eLife.36861.

- Blees A, Januliene D, Hofmann T, Koller N, Schmidt C, Trowitzsch S, Moeller A, Tampé R. Structure of the human MHC-I peptide-loading complex. Nature. 2017 Nov 23;551(7681):525-528. http://dx.doi.org/10.1038/nature24627.

Ilca, S. L., Kotecha, A., Sun, X., Poranen,M. M., Stuart, D. I., Huiskonen, J. T., Localized reconstruction of subunits from electron cryomicroscopy images of macromolecular complexes. *Nature Communications* 6, 8843, 2015.

- Schoebel, S., Mi, W., Stein, A., Ovchinnikov, S., Pavlovicz, R., DiMaio, F., Baker, D., Chambers, M. G., Su, H., Li, D., Rapoport, T. A., Liao, M. Cryo-EM structure of the protein-conducting ERAD channel Hrd1 in complex with Hrd3. *Nature* 548 (7667), 352–355, 2017.

- Roh SH, Hryc CF, Jeong HH, Fei X, Jakana J, Lorimer GH, Chiu W. Subunit conformational variation within individual GroEL oligomers resolved by Cryo-EM. Proc Natl Acad Sci U S A. 2017;114, 8259-8264. http://dx.doi.org/10.1073/pnas.1704725114.

IGBMC