Linking structural ensembles from simulations to the analysis of individual cryo-EM images

Dr. Pilar Cossio

March 2017

Biophysics of Tropical Diseases

Max Planck Tandem Group Leader

University of Antioquia

Outline

- **Motivation:** Hybrid methods
- Introduction to cryo-Electron Microscopy
- Cryo-EM of dynamic systems?
  - BioEM: Bayesian inference of individual cryo-EM images
  - Integrating simulations and BioEM for the analysis of dynamic systems
- Conclusions

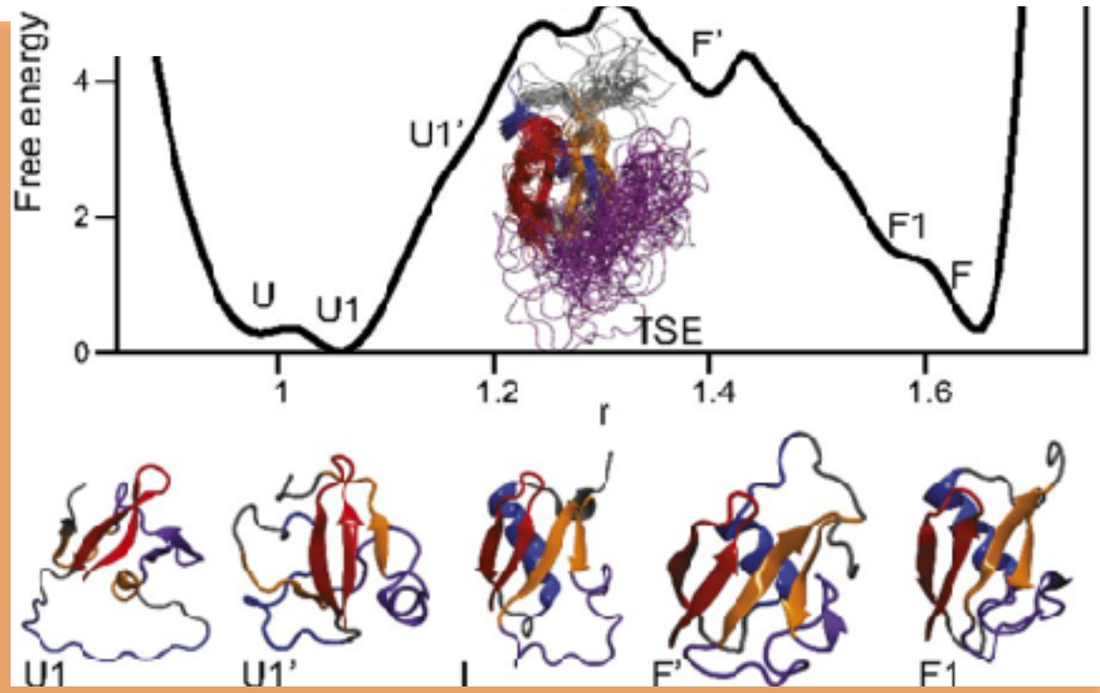
Motivation:

Hybrid/ Integrative Methods

MD Simulations

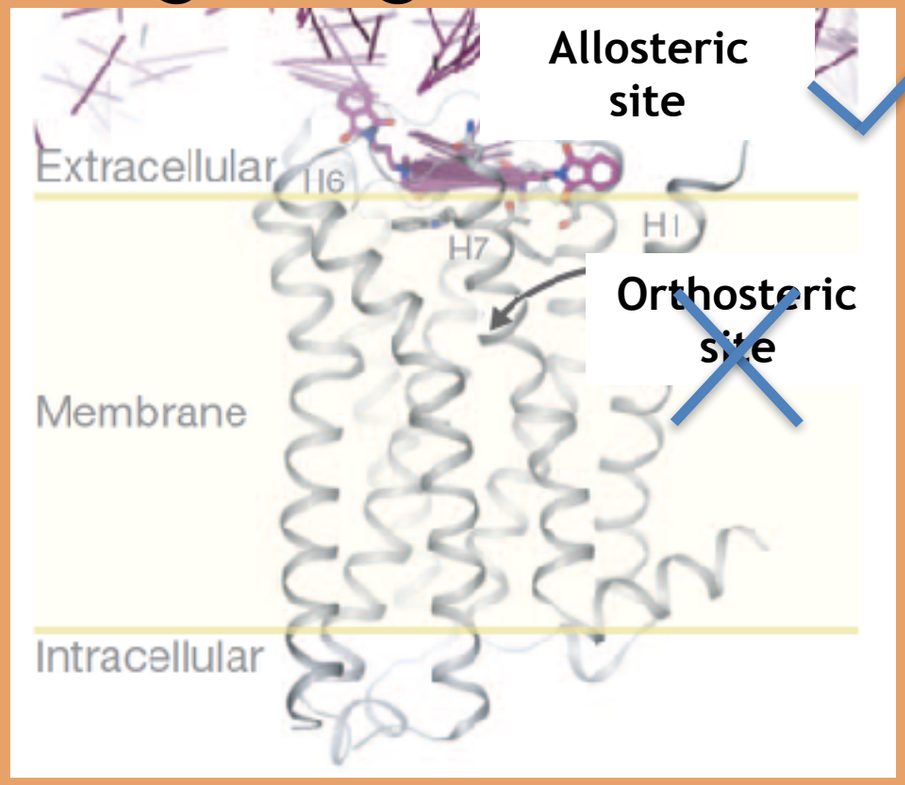
Molecular Dynamics have obtained both the structures and dynamics of proteins:

Protein folding



Ubiquitin folding. *Piana et al.* ***PNAS*** *(2013)* ***110***, *5915-5920.*

Drug Design



Modulation of a G-protein-coupled receptor by allosteric drugs

Dror *et al.* (2013) ***Nature***. 503: 295-299

Limited computation-time, force field accuracy?

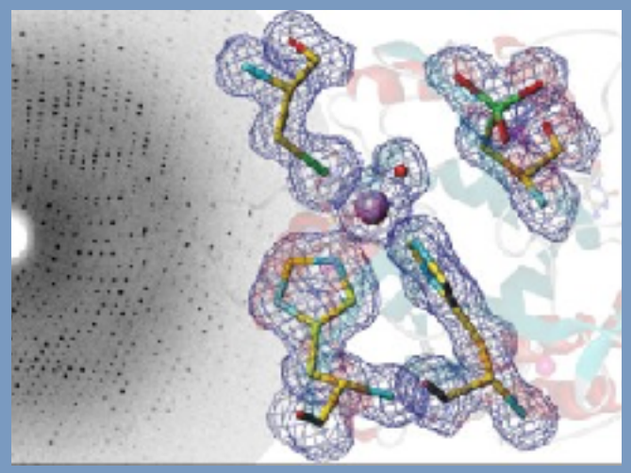
MD simulations can now be used to make predictions.

However, not all systems /phenomena can yet be studied with MD.

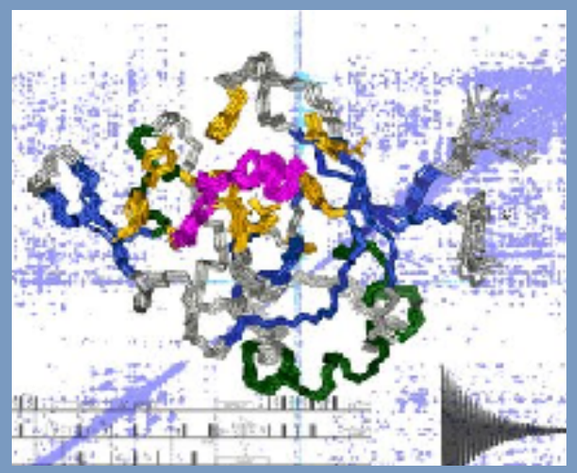
e.g., Large systems with more than 10^7 atoms or times scales greater than *ms*.

Structural Experiments

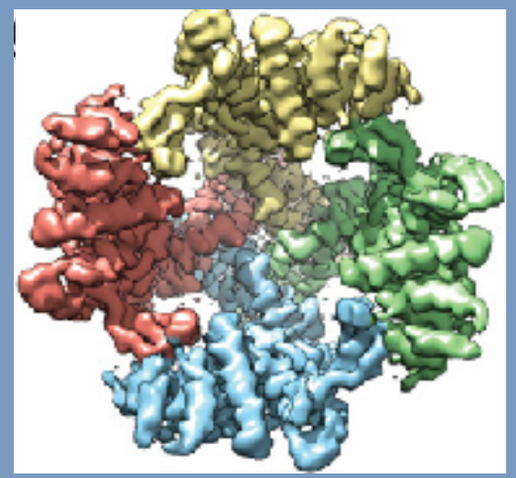
3D structure/map



X-ray crystallography



NMR (nuclear magnetic resonance)

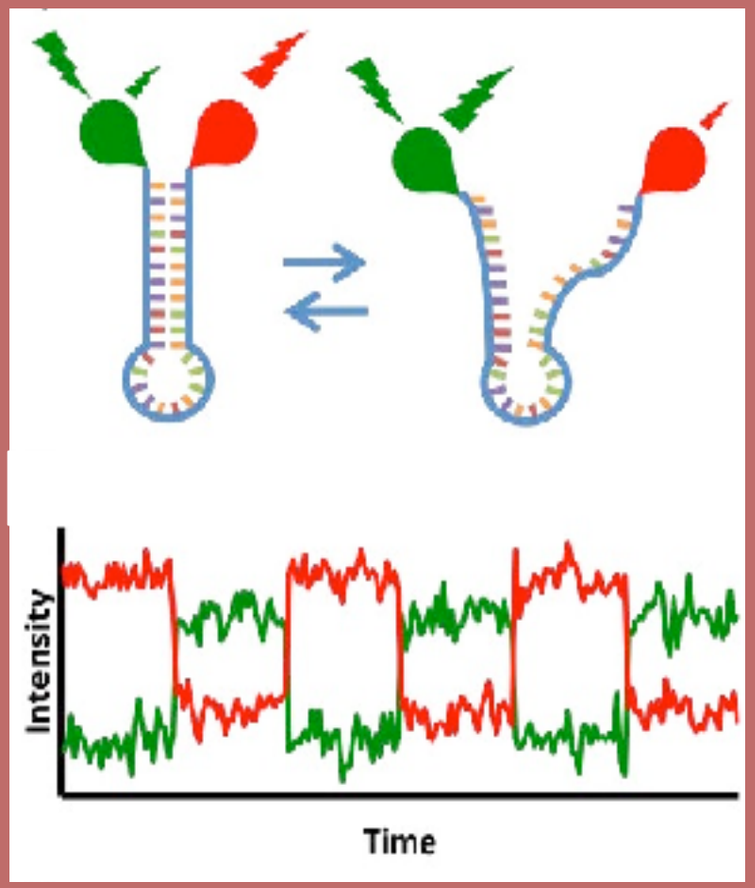


Cryo-electron microscopy

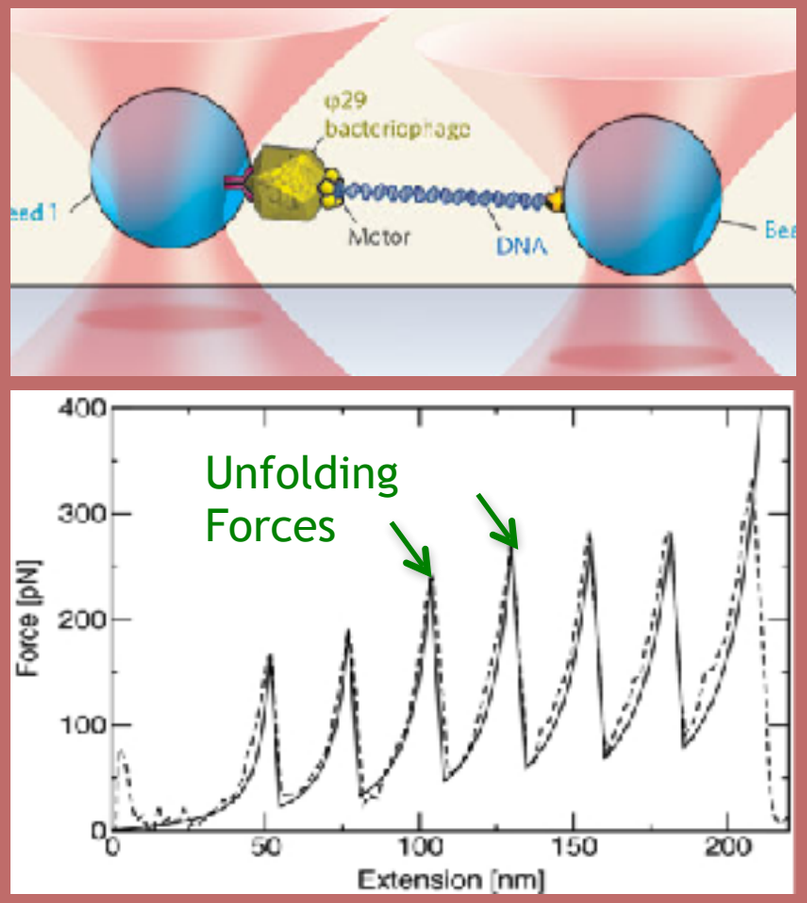
High Resolution (atomic) structures

However, ensemble average: limits the interpretation and extraction of dynamic information.

Single-molecule Experiments

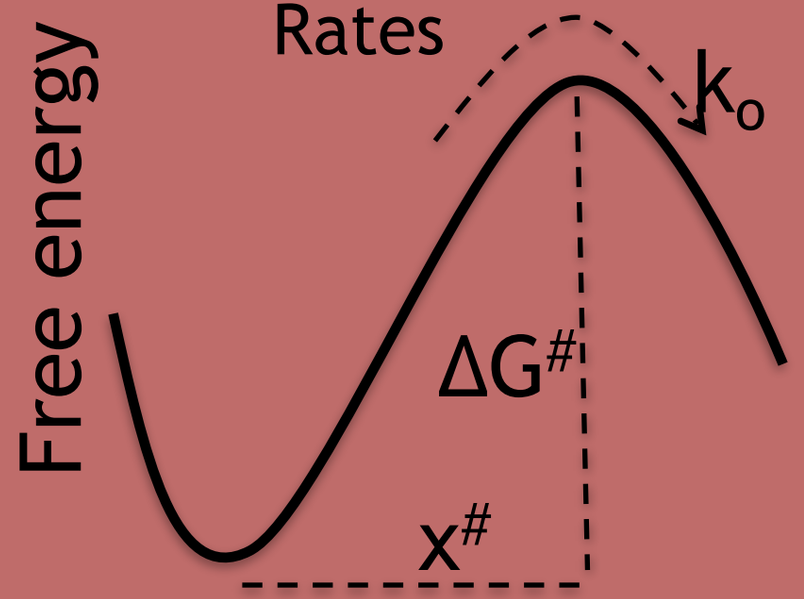


FRET (fluorescence resonance energy transfer)



Force spectroscopy

Dynamics

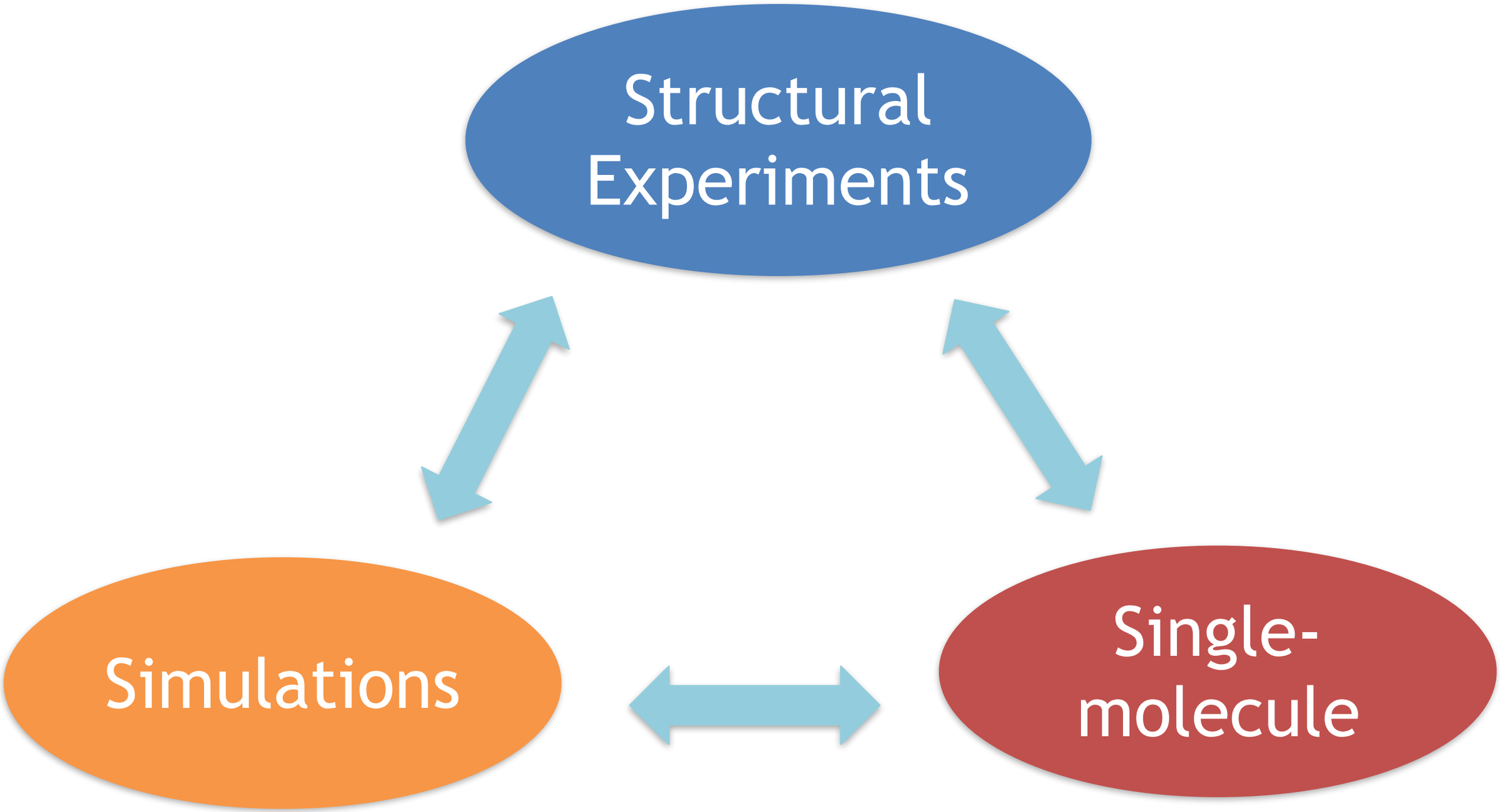


Limited structural information.

Hybrid methods:

Integrate information from both experiments and simulations.

Integrate the methods to understand biomolecules



Requires novel methods:

CellPress

Structure

Meeting Review

Outcome of the First wwPDB Hybrid/Integrative Methods Task Force Workshop

And many more…

Published Online December 2015 Accepted December 2015

Bayesian ensemble refinement by replica simulations and reweighting

Gerhard Hummer[a)] and Jürgen Köfinger

ELSEVIER

Current Opinion in Structural Biology

Volume 42, February 2017, Pages 106–116

Folding & binding • Proteins: Bridging theory and experiment

Principles of protein structural ensemble determination

Massimiliano Bonomi[1,2], Gabriella T. Heller[1,3], Carlo Camilloni[2], Michele Vendruscolo[1]

Show more

http://dx.doi.org/10.1016/j.sbi.2016.12.004

SCIENTIFIC REPORTS

Article | OPEN

The inverted free energy landscape of an intrinsically disordered peptide by simulations and experiments

RESEARCH ARTICLE | COMPUTATIONAL BIOLOGY

Metainference: A Bayesian inference method for heterogeneous systems

Massimiliano Bonomi[1,*,†], Carlo Camilloni[1,†], Andrea Cavalli[1,2] and Michele Vendruscolo[1,*]

+ See all authors and affiliations

Science Advances 22 Jan 2016:
Vol. 2, no. 1, e1501177
DOI: 10.1126/sciadv.1501177

PLOS | COMPUTATIONAL BIOLOGY

RESEARCH ARTICLE

ENCORE: Software for Quantitative Ensemble Comparison

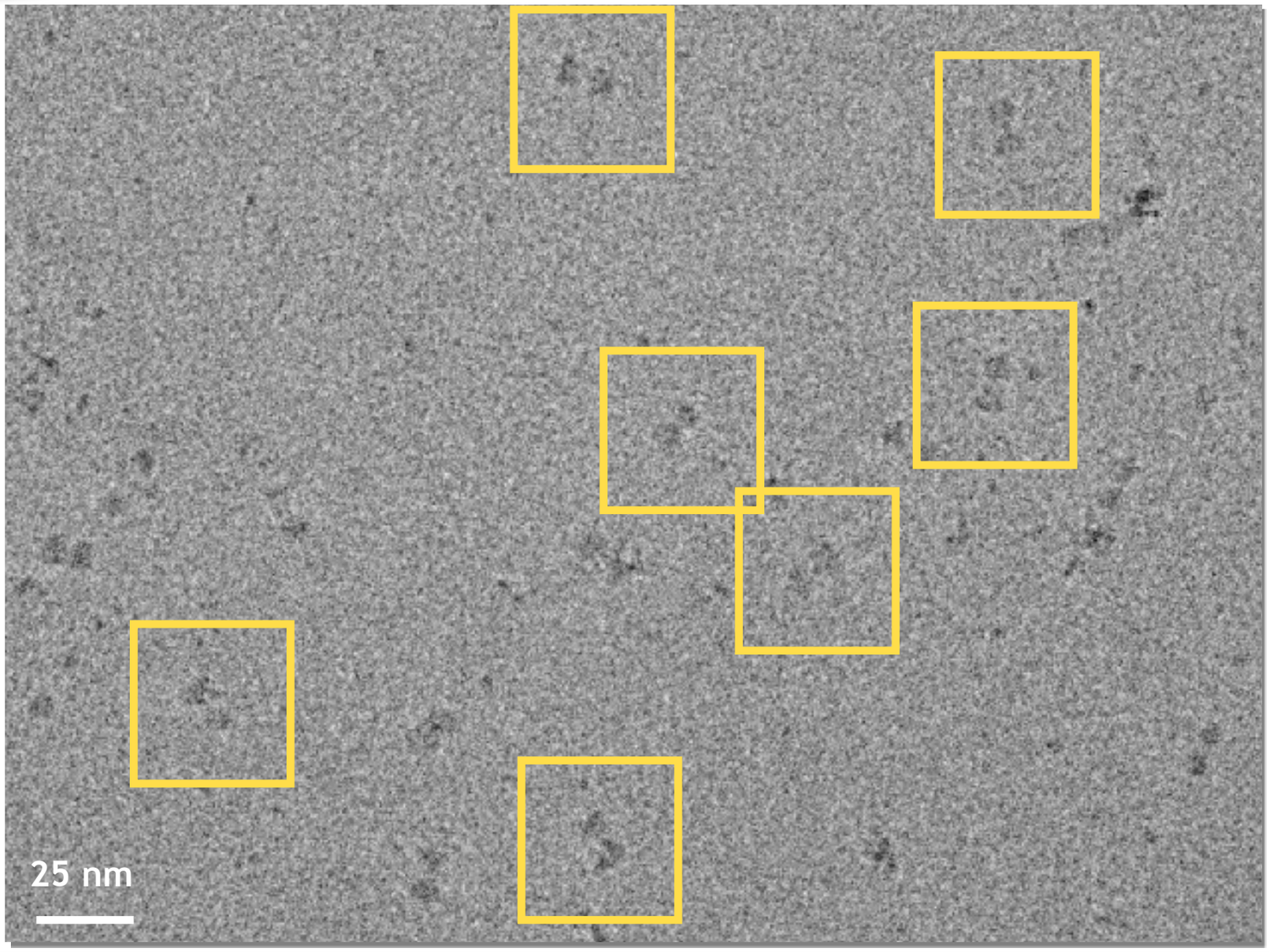
Matteo Tiberti[1,2], Elena Papaleo[2], Tone Bengtsen[2], Wouter Boomsma[2]*, Kresten Lindorff-Larsen[2]*

1 Department of Biotechnology and Biosciences, University of Milano-Bicocca, Milan, Italy, 2 Structural Biology and NMR Laboratory, Department of Biology, University of Copenhagen, Copenhagen, Denmark

cryo-EM

EM Imaging

Cryo-EM: Frozen Biological sample imaged with an electron microscope



A_1A_o ATP-synthase from *Pyrococcus furiosus from Matteo Allegretti.*

Challenge: Images are noisy!

EM Imaging

3D reconstructions

Requirements:

- Relatively large systems.
- Symmetric systems with common features for clustering.
- Non-dynamic systems.
- Hundreds of thousands of particles are needed to obtain a good resolution.

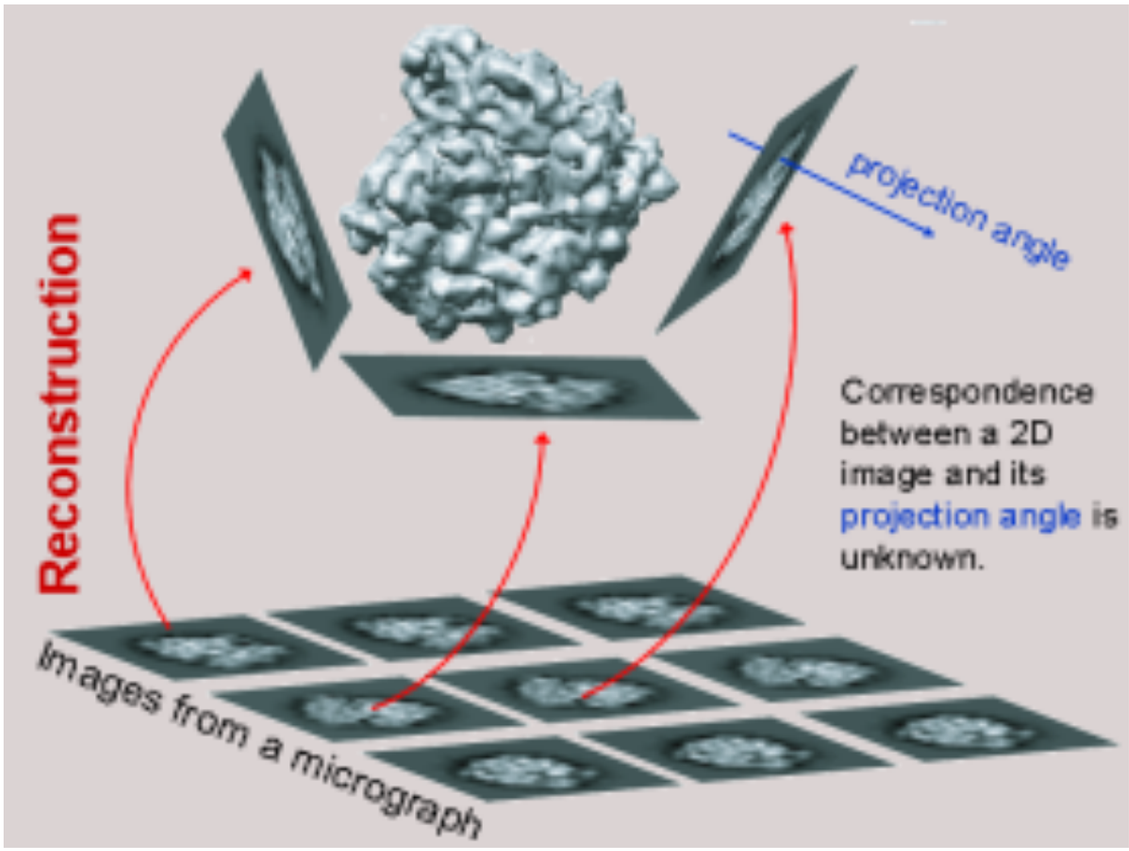


Image from compbio.berkeley.edu

EM Imaging

3D reconstructions

Assumptions:

- All particle-images are in a single conformational state
- The particle-image orientations are randomly distributed
- Sometimes the molecular symmetry.

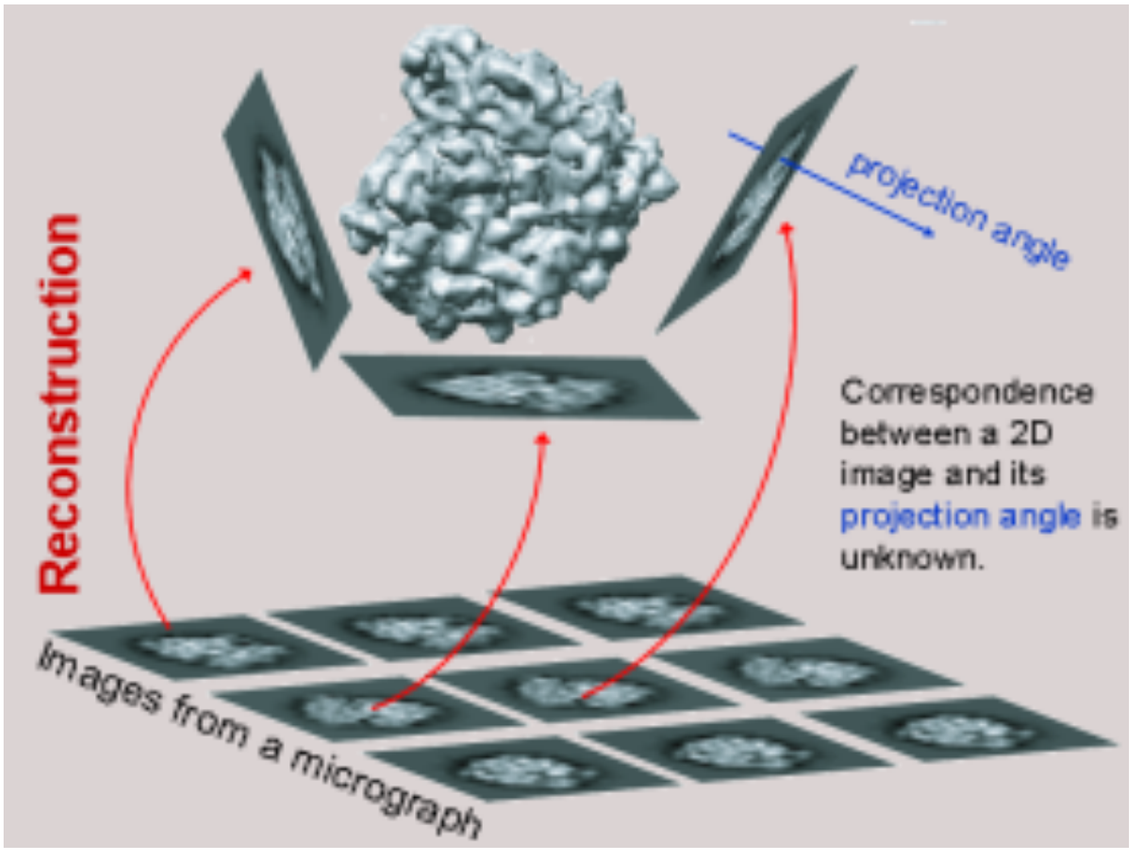
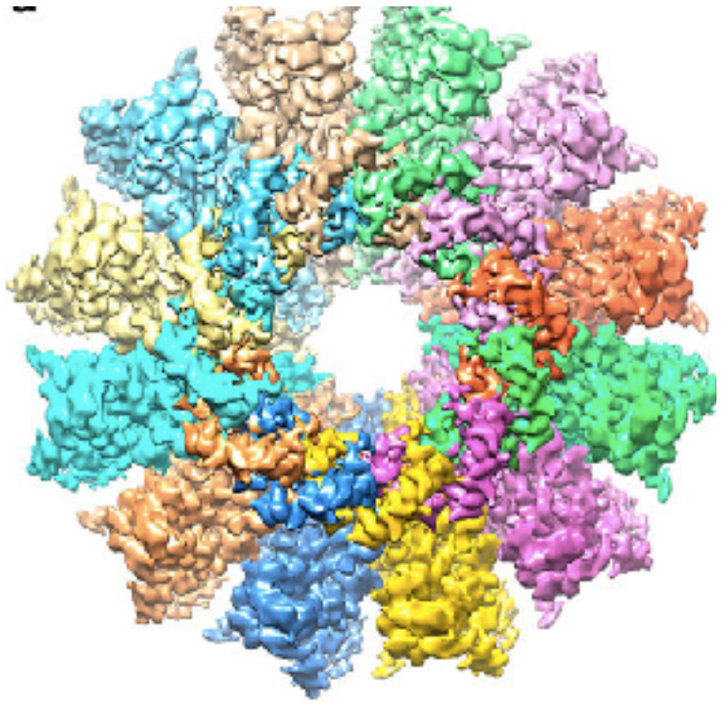


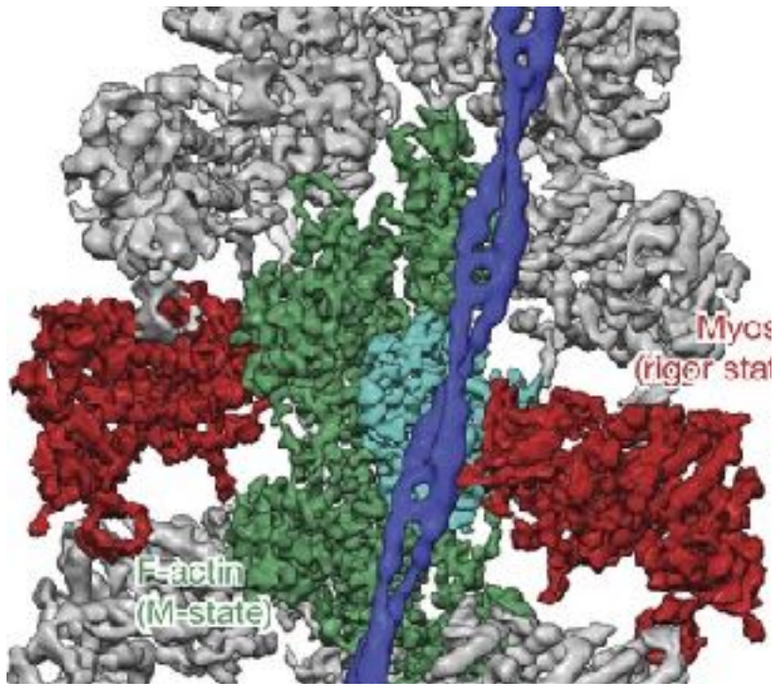
Image from compbio.berkeley.edu

3D reconstructions **near atomic resolution:**

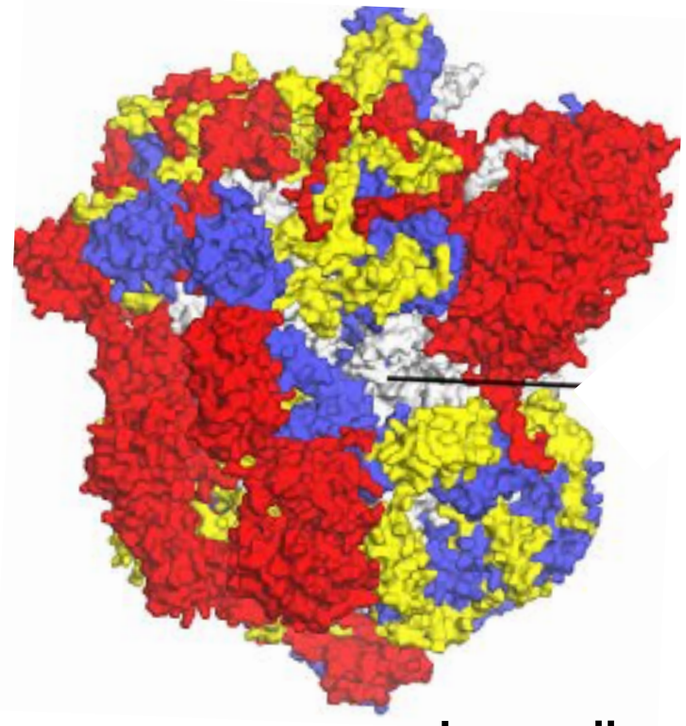
EM Imaging

Bacteriophage T4,
Nature Commun,
7548 (2015)





actomyosin complex,
Nature, 534, 724 -728
(2016)



Large ribosomal subunit, *Science*, **348**, 95-98 (2014).

And many more…

Cryo-EM has revolutionized structural biology!

What to do when the EM reconstruction methods fail?

BioEM: *Bayesian inference of individual EM images*

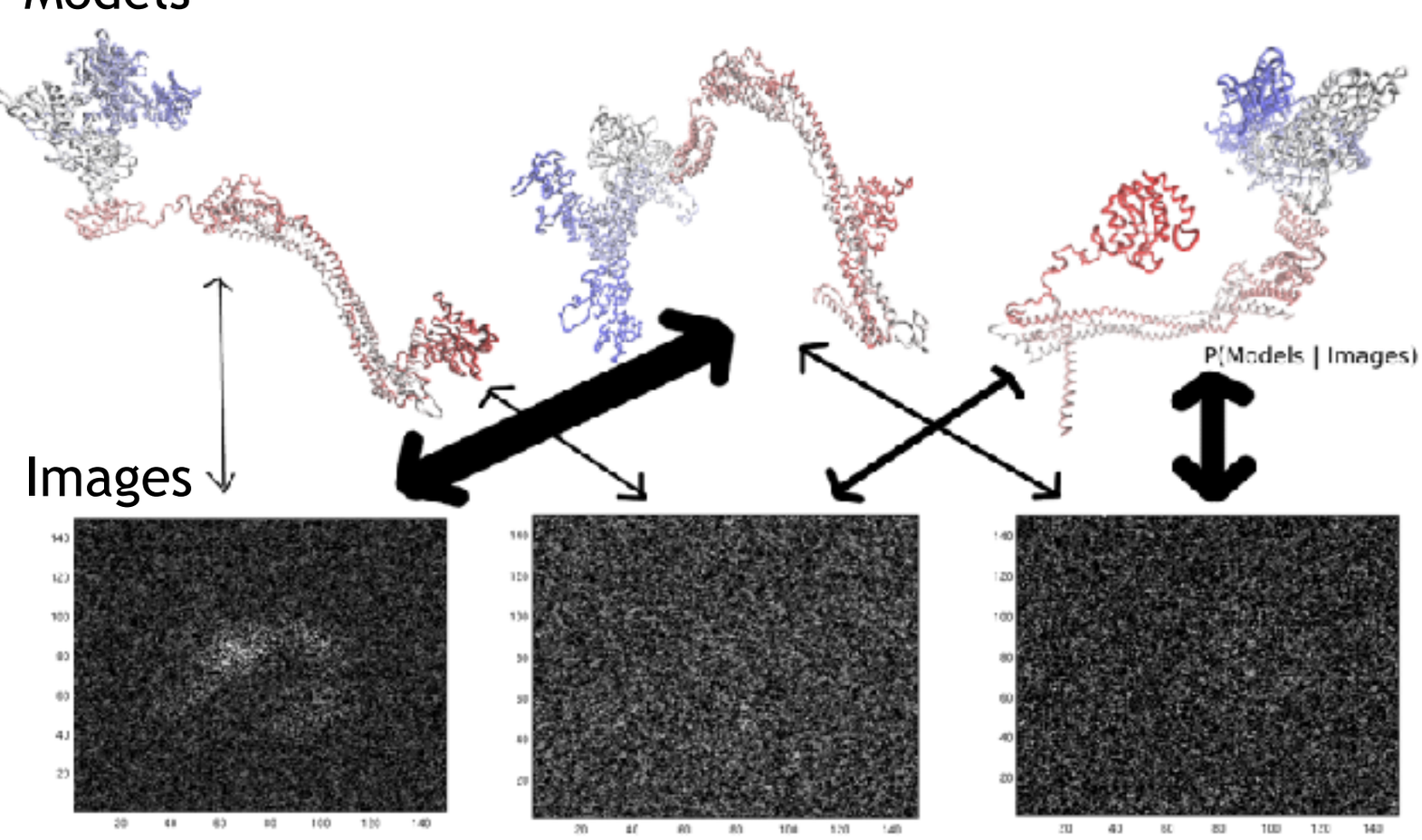
EM Imaging

of dynamic/flexible and asymmetric biomolecules?

Dr. Gerhard Hummer

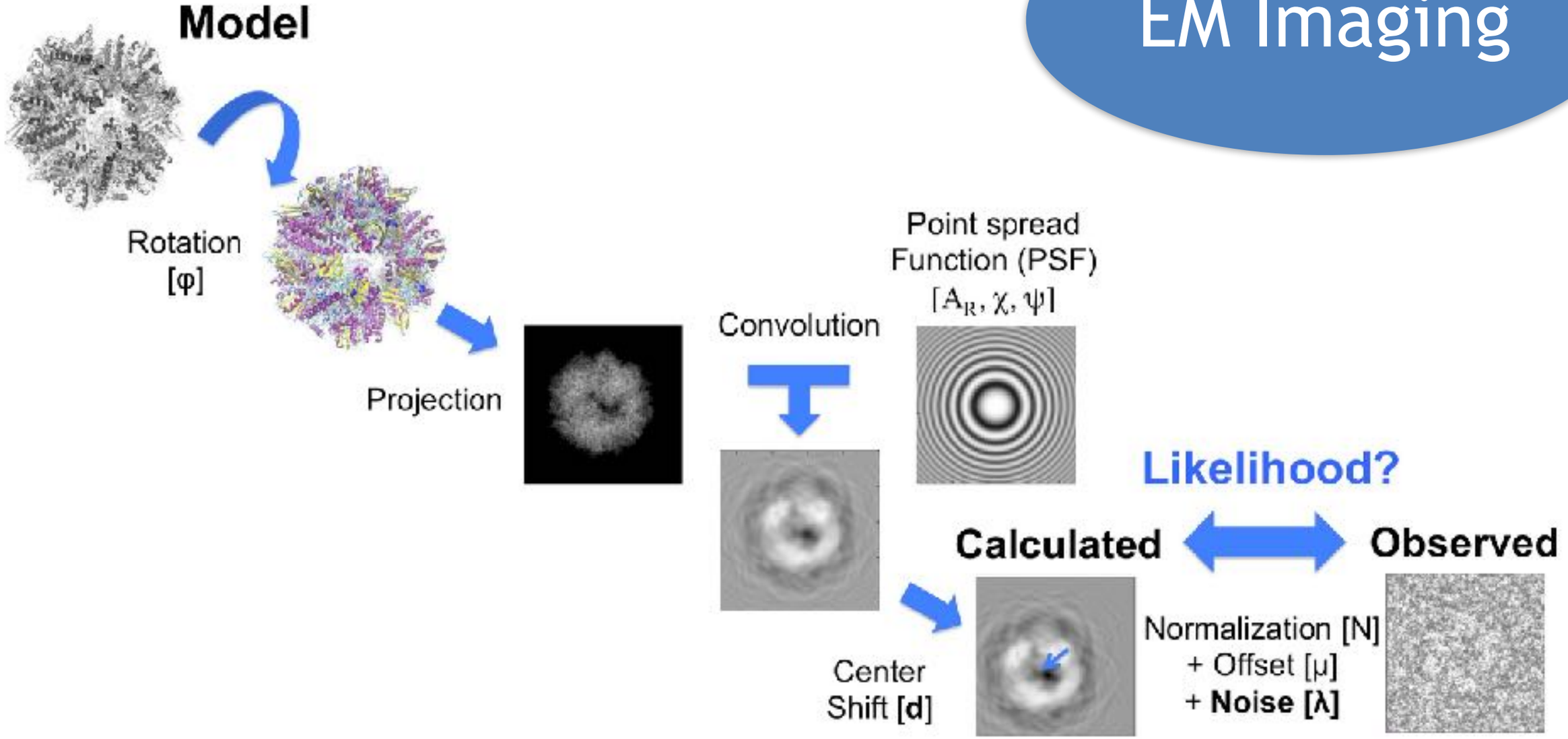
mpibp
max-planck-institut für biophysik

Bayesian inference of electron microscopy (BioEM): obtain the probability of each model given a set of EM images.



Cossio, Hummer. (2013) **J. Struct. Biol.** 184: 427-37.

EM Imaging



Likelihood function:

$$L(I^{Obs} \mid I^{Cal}) = \exp(-\sum_{pix}(I^{Obs} - I^{Cal}(\theta))^2 / 2\lambda^2)$$

Parameters

Noise

Bayesian Analysis: Integrate the likelihood over all possible parameters and include prior information too.

For an individual image

$$P(M|I^{obs}) = \int L(I^{obs}|I^{cal}) p_M p(\theta) d\theta$$

Priors

Parameters

For multiple images $\mathscr{I} = \{I_1^{obs}, I_2^{obs}, ...\}$

$$P(M|\mathscr{I}) = \prod_{I^{obs}} P(M|I^{obs})$$

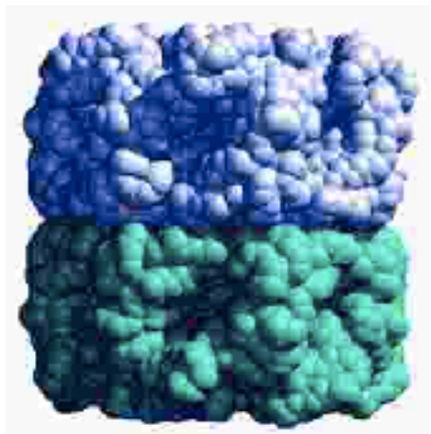
EM Imaging

$$P(M_A|\mathscr{I})/P(M_B|\mathscr{I})$$

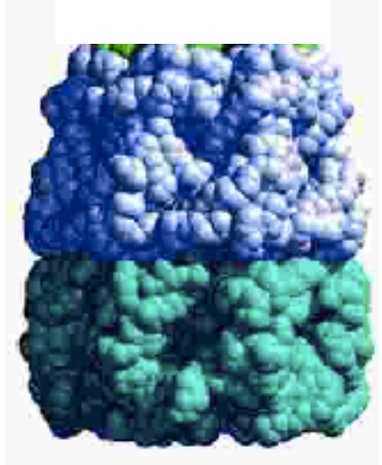
Model Ranking/ Comparison

GroEL Chaperonin: *a test system*

APO



GroEL+ATP



EM (same data)

EM (same data)

X-ray

X-ray

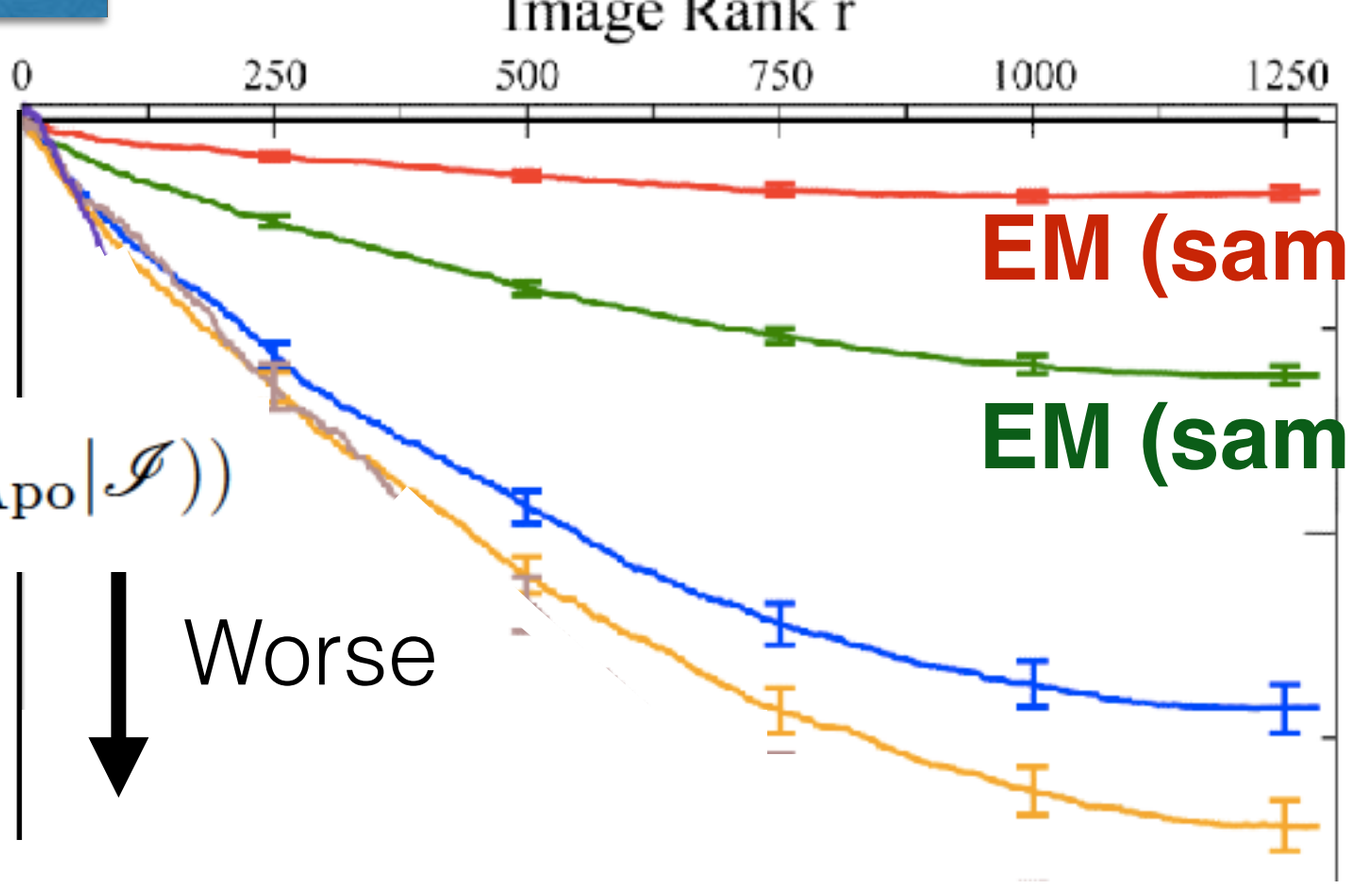
X-ray

cryo-EM images in the APO state

Cumulative Evidence

$$\sum \ln\left(P(M|\mathscr{I})/P(\mathrm{Xray}_{\mathrm{Apo}}|\mathscr{I})\right)$$

Worse



EM (same data)

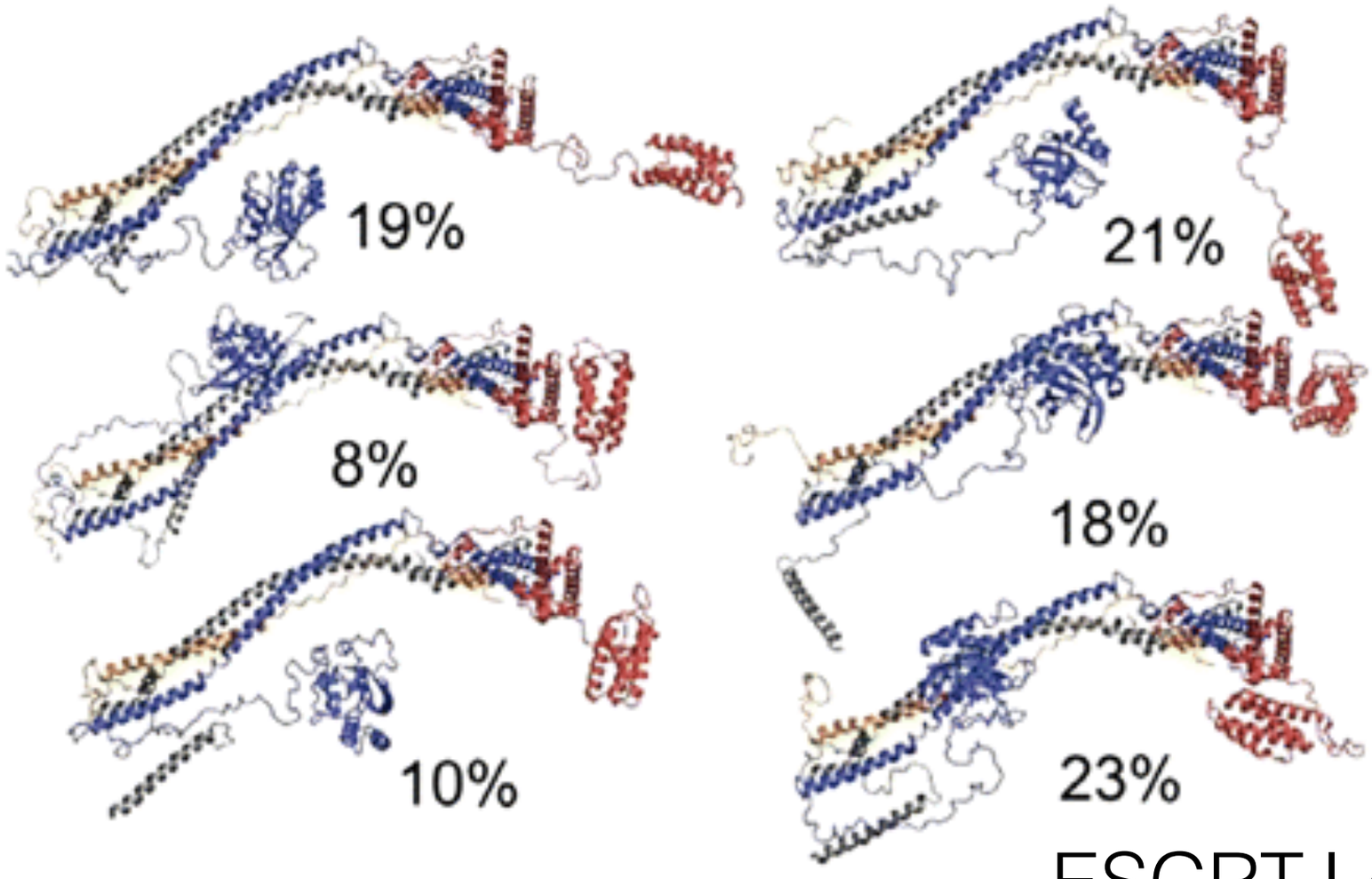
EM (same data)

X-ray

X-ray

Linking to structural ensembles from simulations

Information from cryo-EM images of structural ensembles (e.g. flexible biomolecules)?



ESCRT-I complex

Boura et al. PNAS,**108**, 9437–9442 (2011)

Posterior BioEM probability of sets of models $\mathscr{M} = \{M_1, M_2, ...\}$

$$P(\mathscr{M}|\mathscr{I}) = \prod_{I^{obs}} \sum_{M} w_M P(M|I^{obs})$$

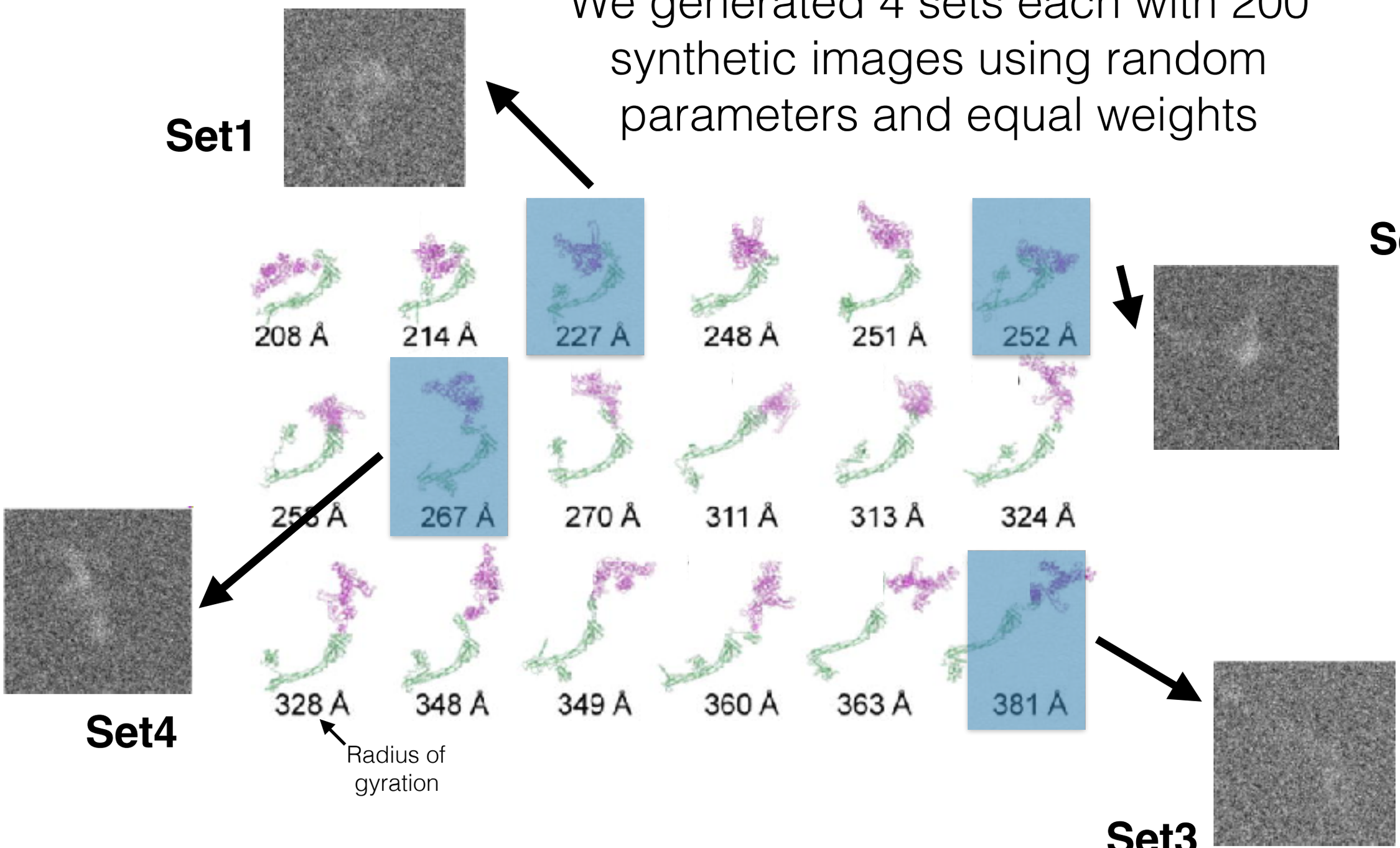
where the model weights are normalized $\sum_M w_M = 1$

-) *Maximum entropy:* optimize the weights (w_M) of each model to fit best the data.

-) *Minimum ensemble:* minimum number of structures that best represent the data.

Minimum ensemble method validated with the ESCRT I-II supercomplex*

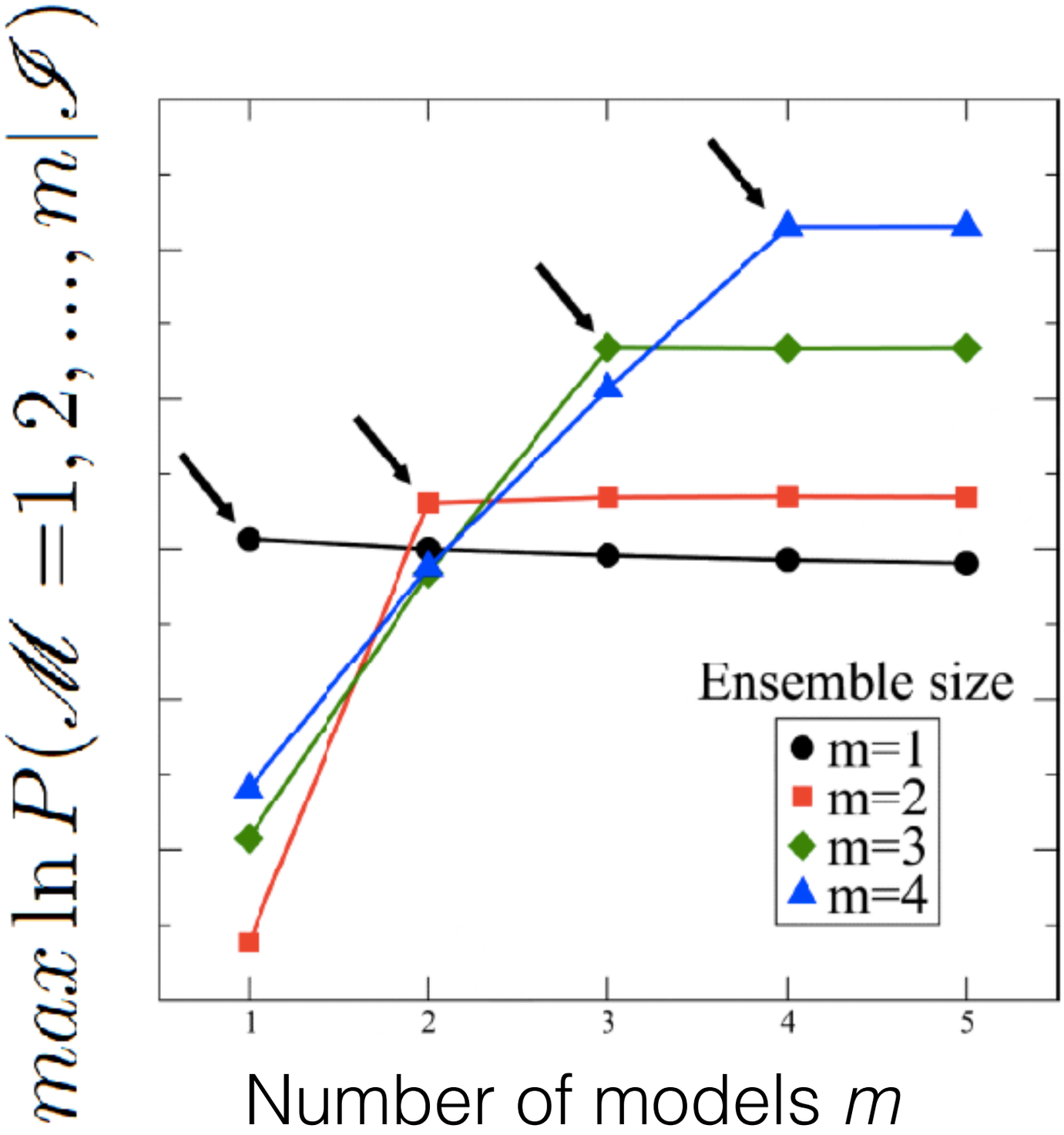
We generated 4 sets each with 200 synthetic images using random parameters and equal weights



*Boura et al. *Structure.* **20**: 874–886 (2012).

Minimum Ensemble?

- The posterior increases sharply until the number of models reaches the actual ensemble size, as indicated by arrows.
- The minimum number of members of the ensemble is that by which adding an extra member does not increase the posterior



However, the calculation of BioEM posteriors for large numbers of particles and models is computationally demanding.

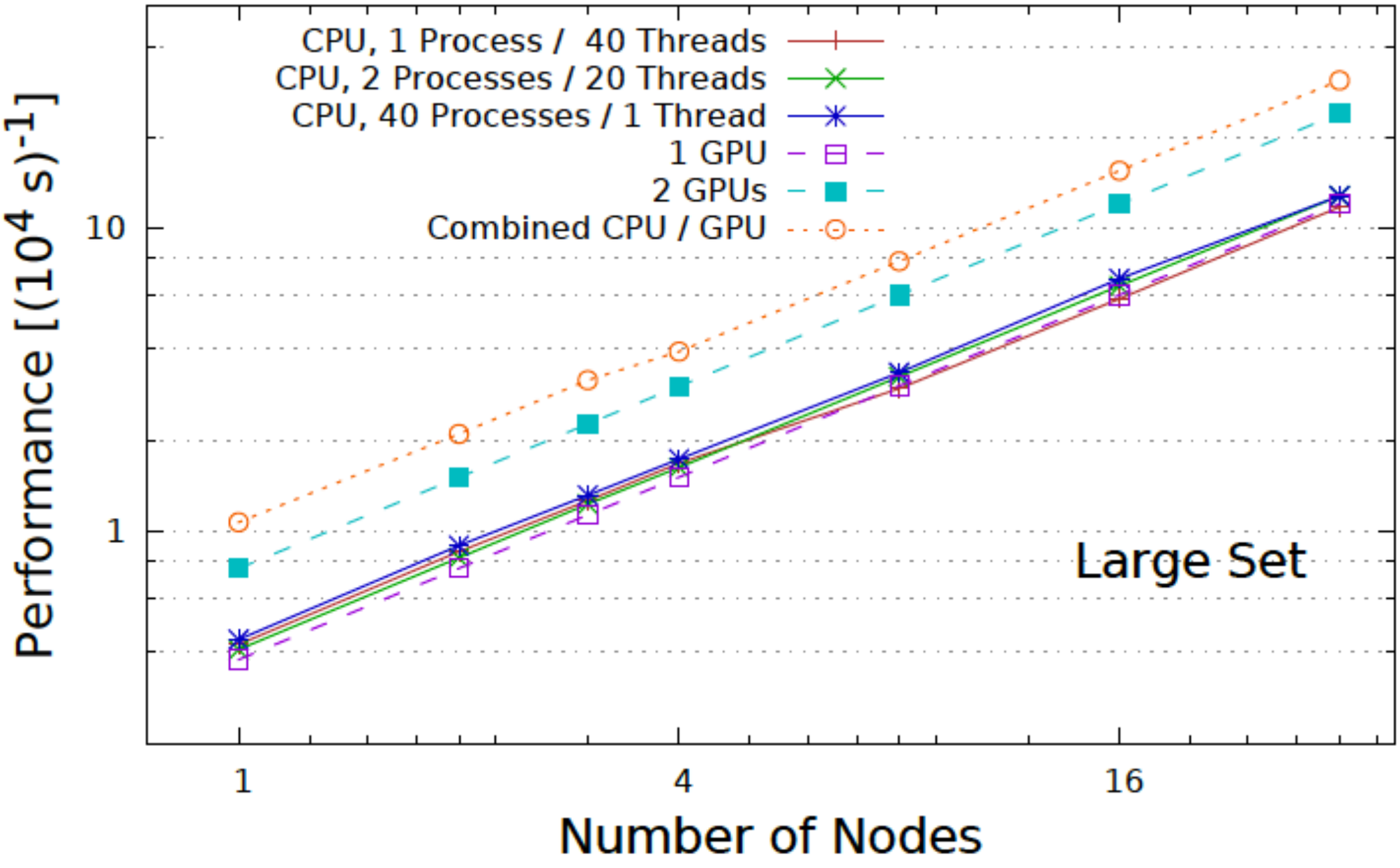
Building a fast code

BioEM + GPUs*

https://gitlab.rzg.mpg.de/MPIBP-Hummer/BioEM

* **Cossio**, et al. (2017) *Compu. Phys. Commun. 210*, 163-171.

BioEM perfomance over CPUs and GPUs



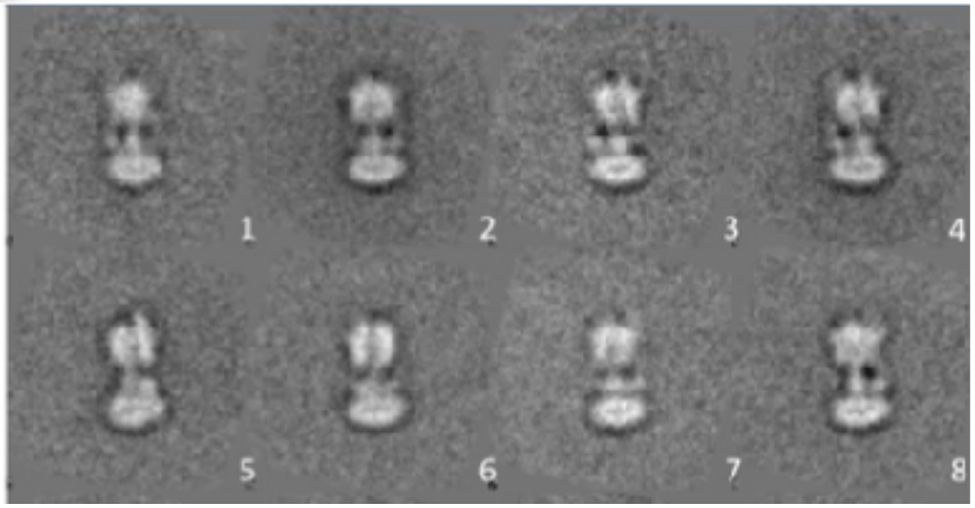
* **Cossio**, et al. (2017) *Compu. Phys. Commun. 210*, 163-171.

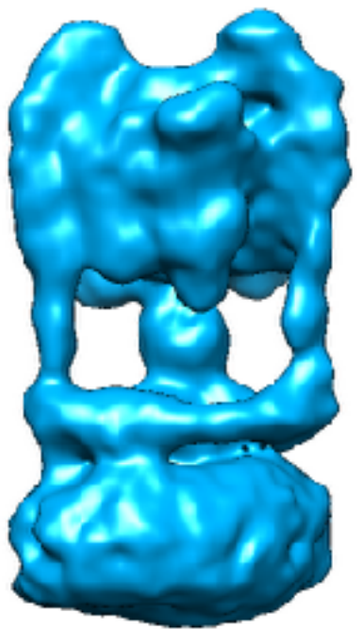
BioEM Applications

Imaging Application

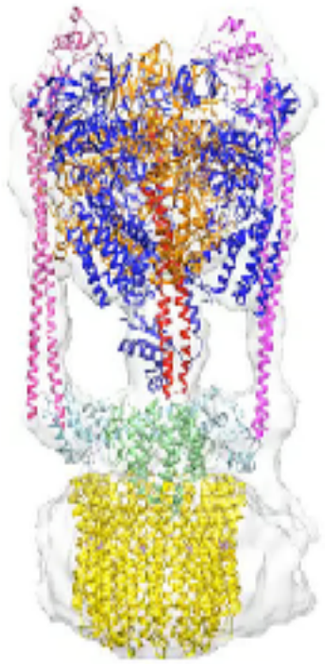
What is the most probable c-ring stoichiometry *Archaea* ATP-synthase?

51000 particles
13 Å resolution

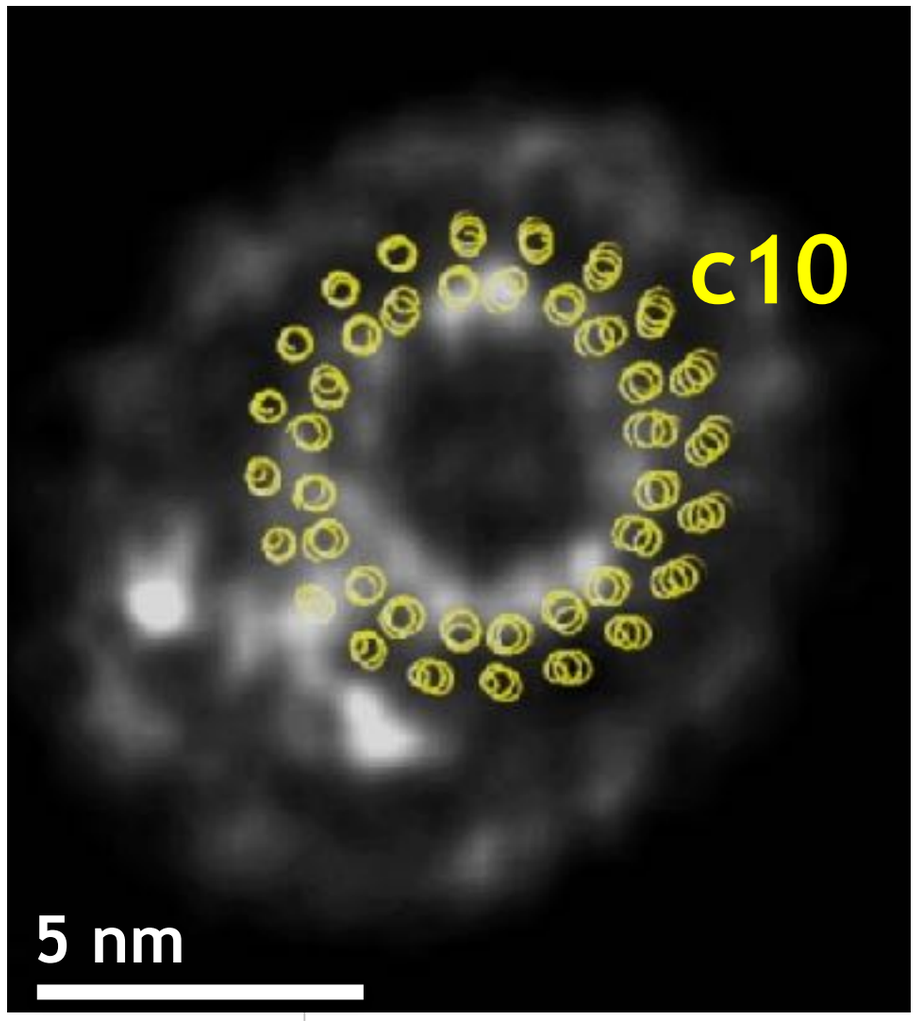




Unpublished data



5 nm



A_1A_o ATP-synthase from Pyrococcus furiosus

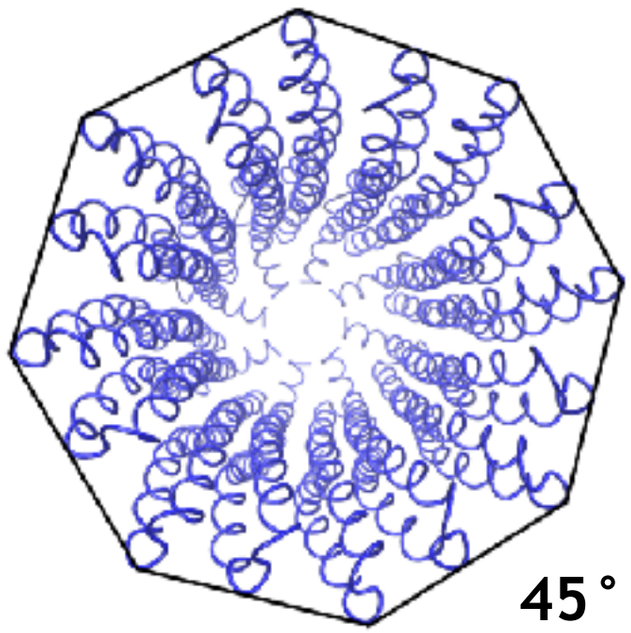
Collaboration with Prof. Dr. Kühlbrandt, Dr. Vonck, Dr. Allegreti. Max Planck Biophysics

Imaging Application

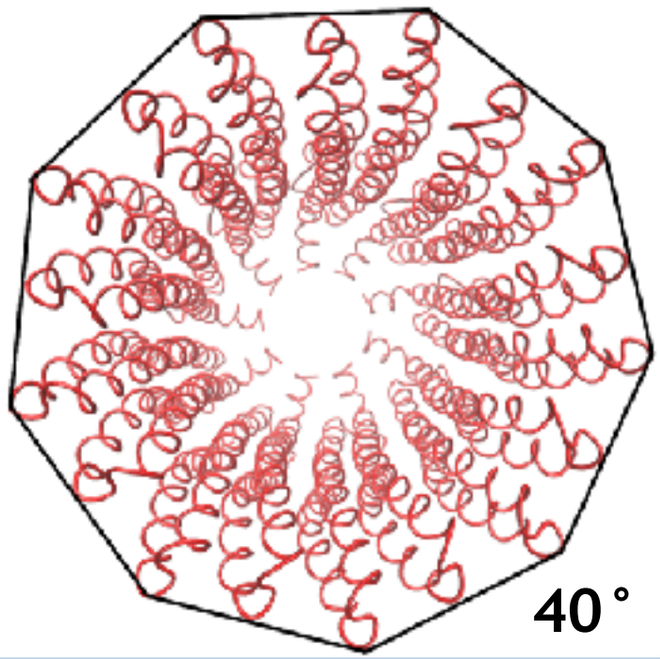
C-ring Models

Monomer from c10 crystal

c(8) - ring

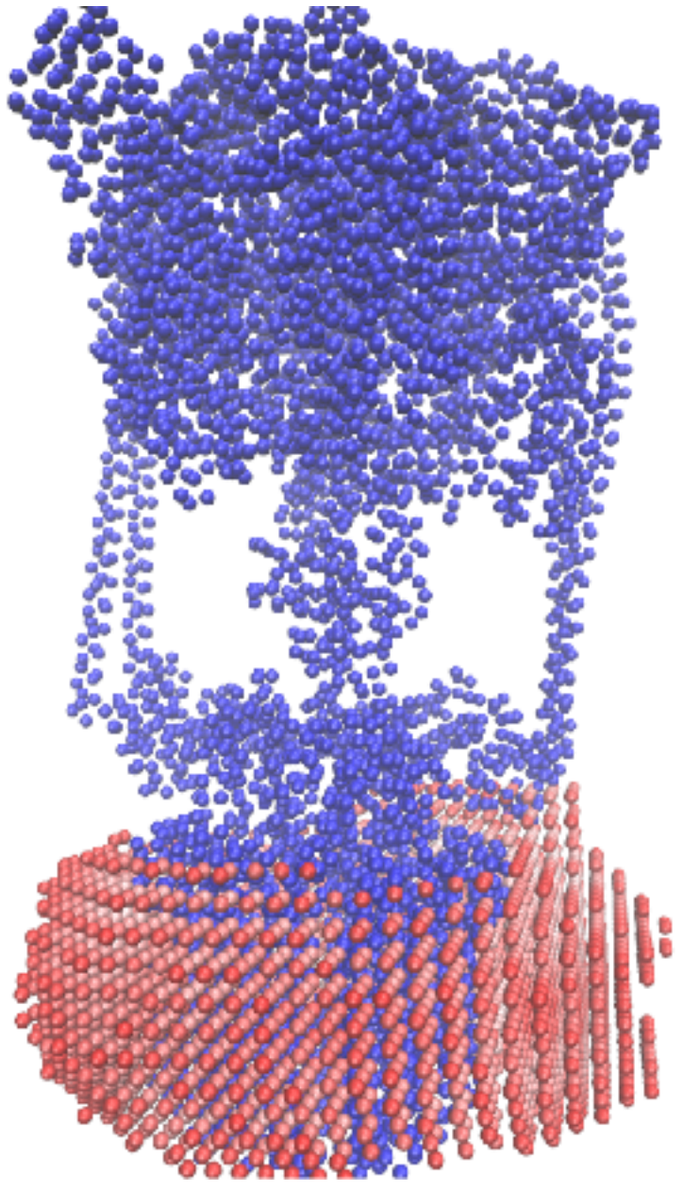


c(9) - ring



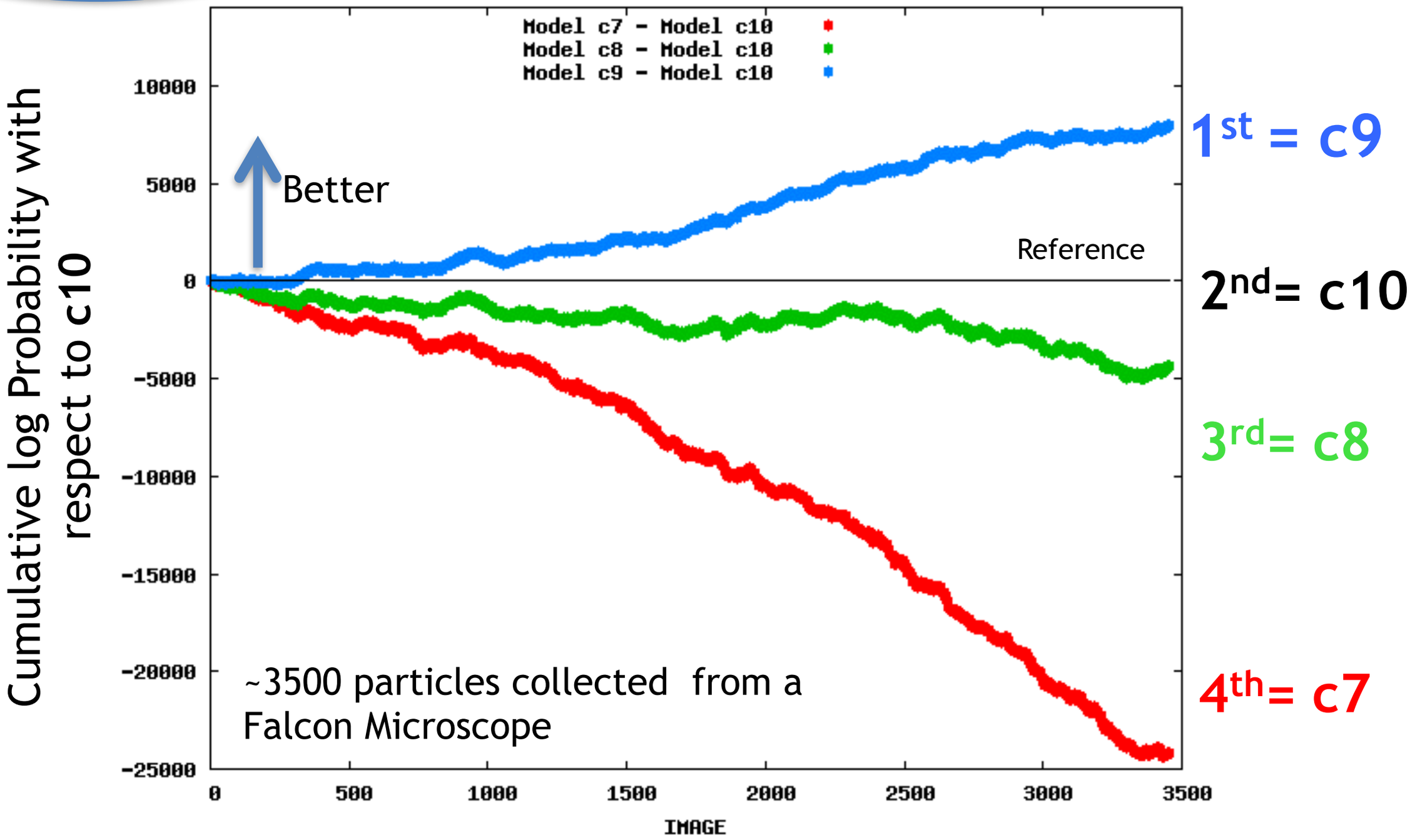
Soluble part
+ c-ring (7,8, 9,10 sym.)

+ detergent from 3D reconstruction



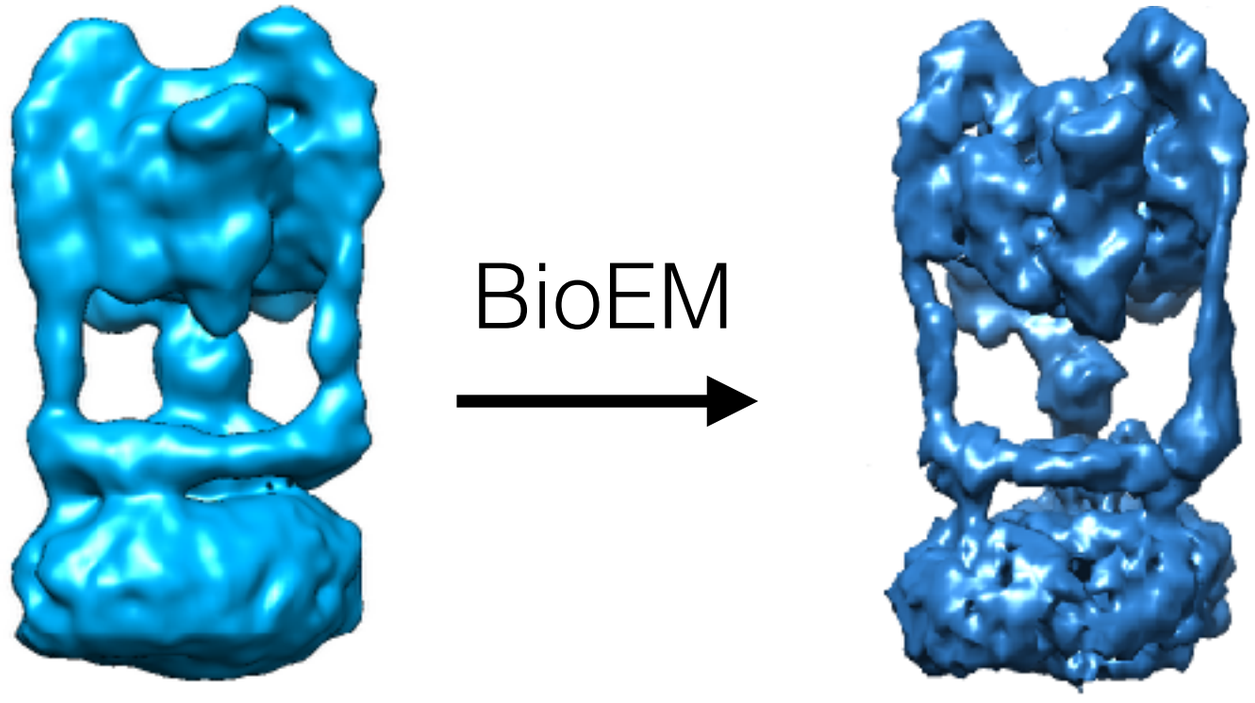
Imaging Application

BioEM Probability Discrimination:



Current & Future work

- 3D Reconstruction Refinement: Improve the resolution of a 3D map using BioEM



- Ensemble refinement with coarse-grained simulations using the BioEM minimal ensemble method

- BioEM gradient-based simulations/refinement: use the BioEM posterior as a biasing force

UNIVERSIDAD DE ANTIOQUIA
1803

MAX-PLANCK-GESELLSCHAFT

Max Planck Tandem Group
Biophysics of Tropical Diseases

Welcome to Colombia:

Lonely Planet names Colombia 2nd best country to visit in 2017

October 26, 2016

Guests are Welcome!

Acknowledgements

Dr. Gerhard Hummer

Prof. Dr. Alessandro Laio

Thank you for your attention