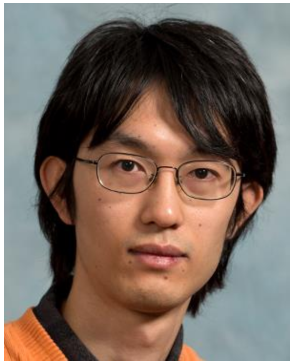
RELION-5

Sjors H.W. Scheres

CCP-EM Icknield
Model building workshop
4 Nov 2024

UKRI
MRC Laboratory of Molecular Biology

The RELION team


Takanori
Nakane


Dari
Kimanius


Johannes
Schwab


Kiarash
Jamali

#teamtomo


Bogdan
Toader


Jasenko
Zivanov


Alister
Burt


Rangana
Warshamanage

Also: past members, Lindahl group & CCP-EM

RELION-5.0 (beta-release)

stdout will go here; double-click this window to open stdout in a separate window

stderr will go here; double-click this window to open stderr in a separate window

https://relion.readthedocs.io/release-5.0

New in RELION-5

- Rethinking regularization
  - Blush

  Nature Methods 21, 1216–1221 (2024)

- Continuous heterogeneity
  - DynaMight

  Nature Methods 21, 1855–1862 (2024)

- Automated model building
  - ModelAngelo

  Nature 628, 450-457 (2024)

- A full tomography pipeline

  FEBS Open Bio, in press (2024)

Blush

Regularisation by Denoising (RED)

Rethinking Regularization

Bayes' theorem

$$P(X|Y) \propto P(Y|X)P(X)$$

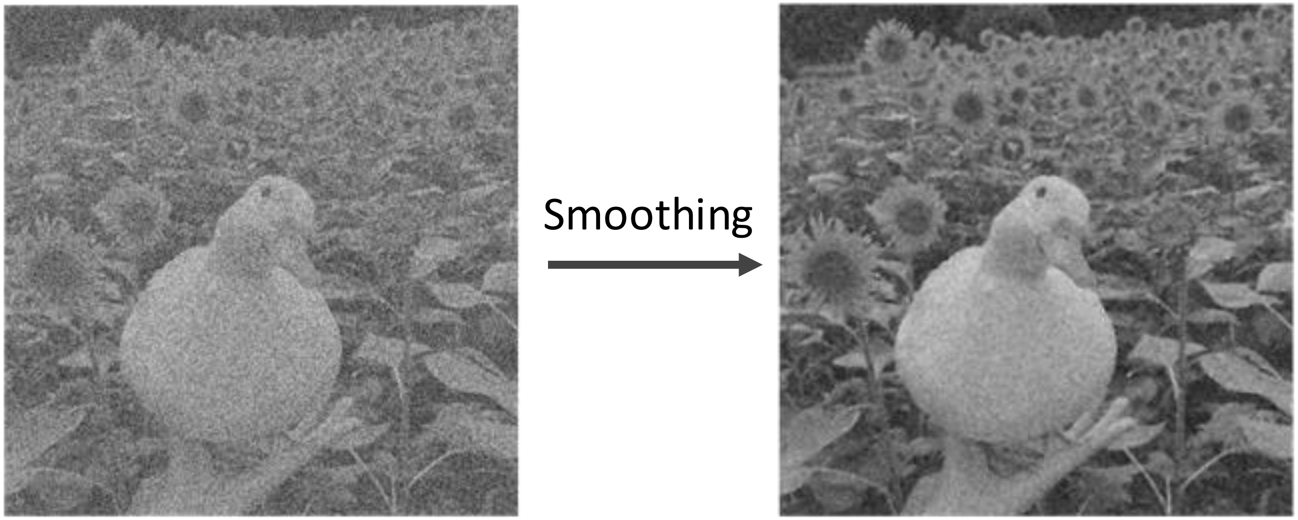
or

$$\log P(X|Y) \propto \underbrace{\log P(Y|X)}_{\text{Likelihood}} + \underbrace{\log P(X)}_{\text{Prior}}$$

Regularization using the Gaussian prior

Molecular densities are smooth

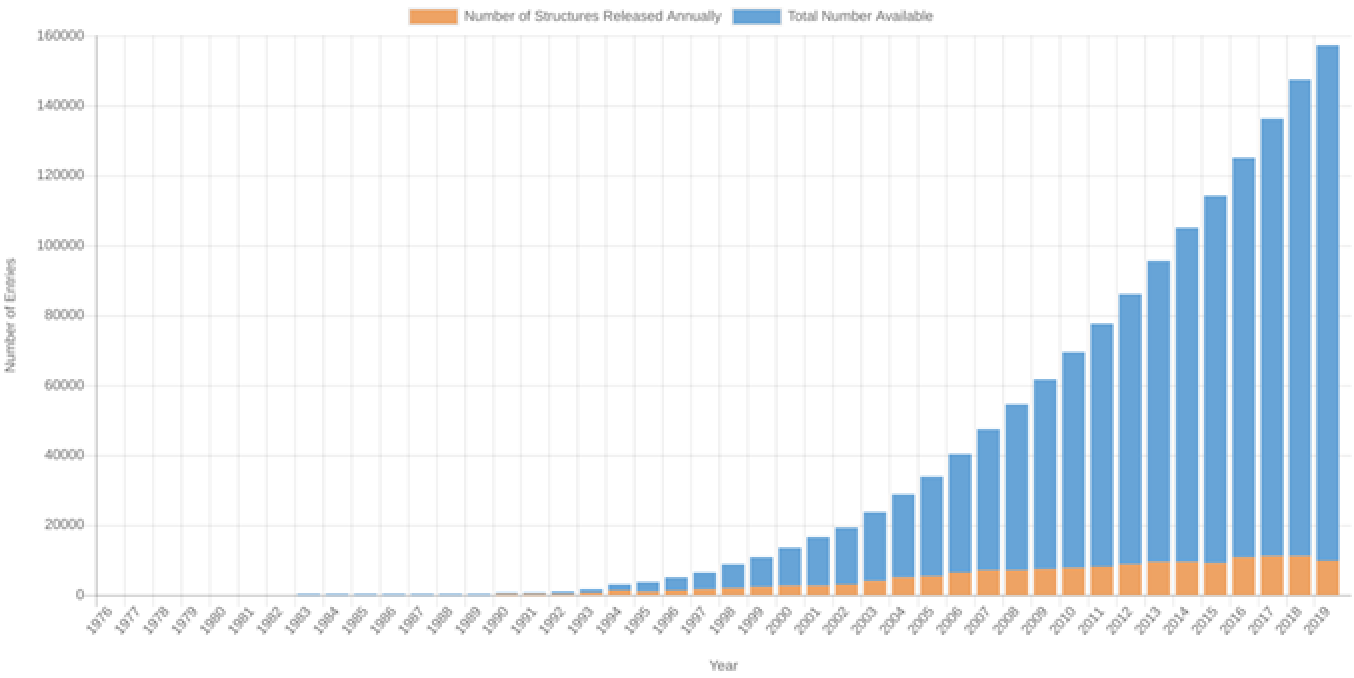
$$\mathrm{P}(x) := C \exp\Big(- \|\tau^{-1} x\|^2 \Big)$$

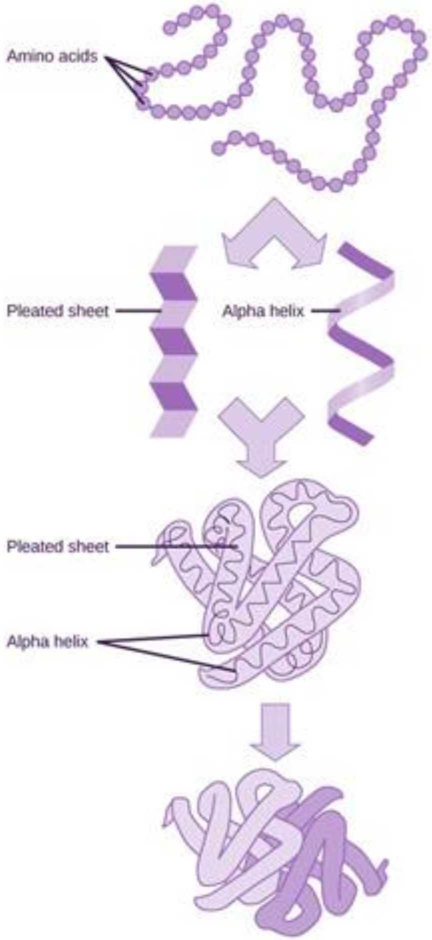


Smoothing

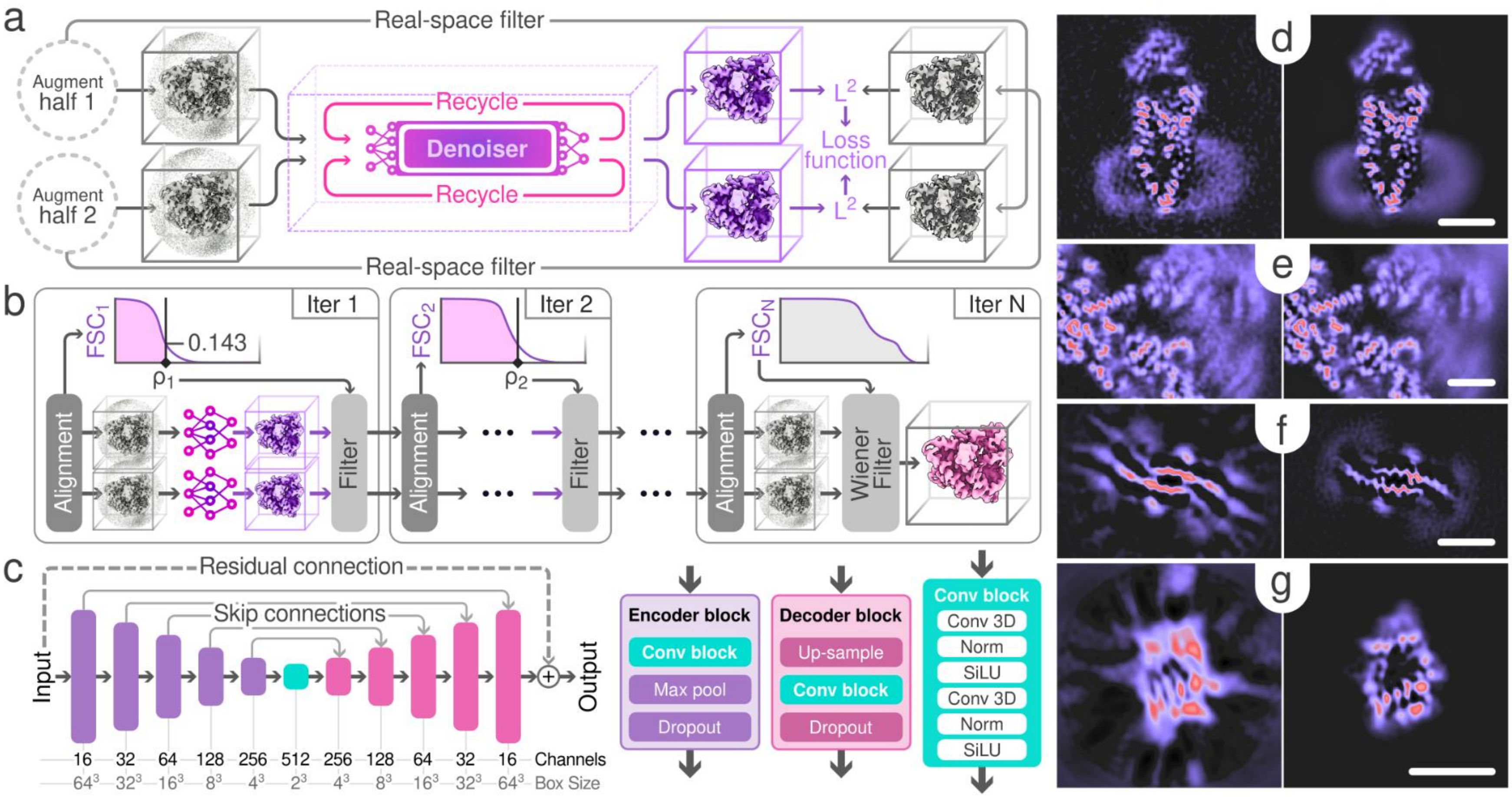
We know more about proteins

Prior structural knowledge

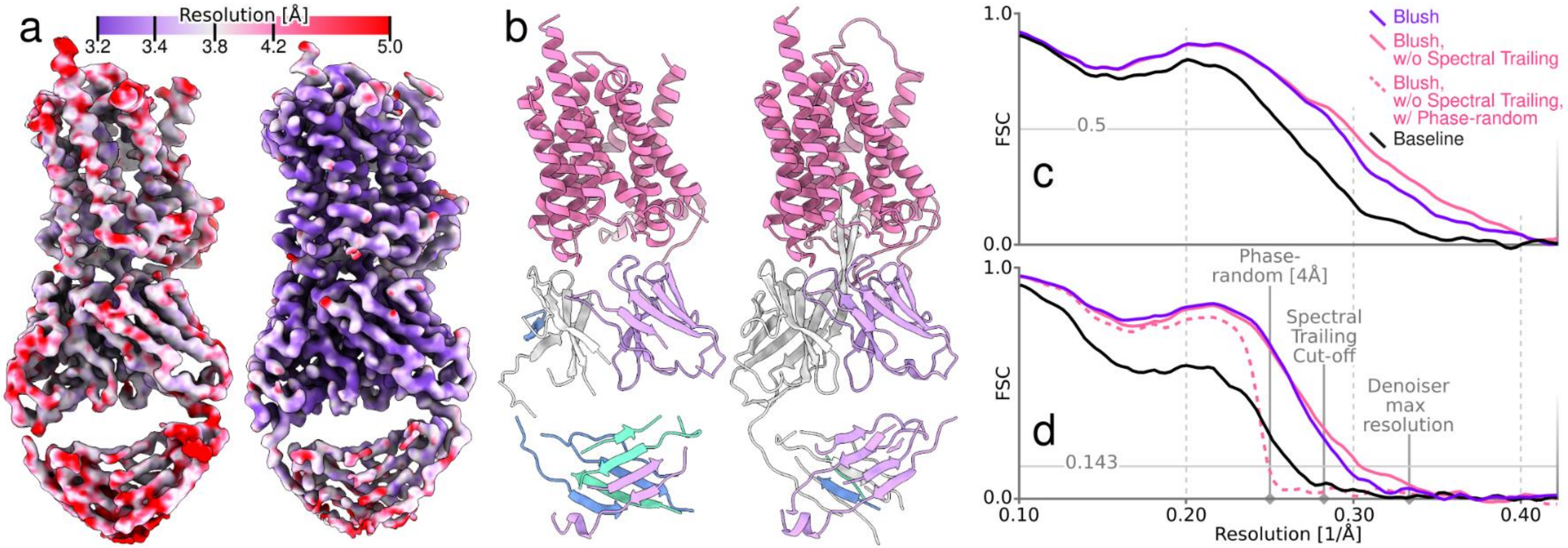




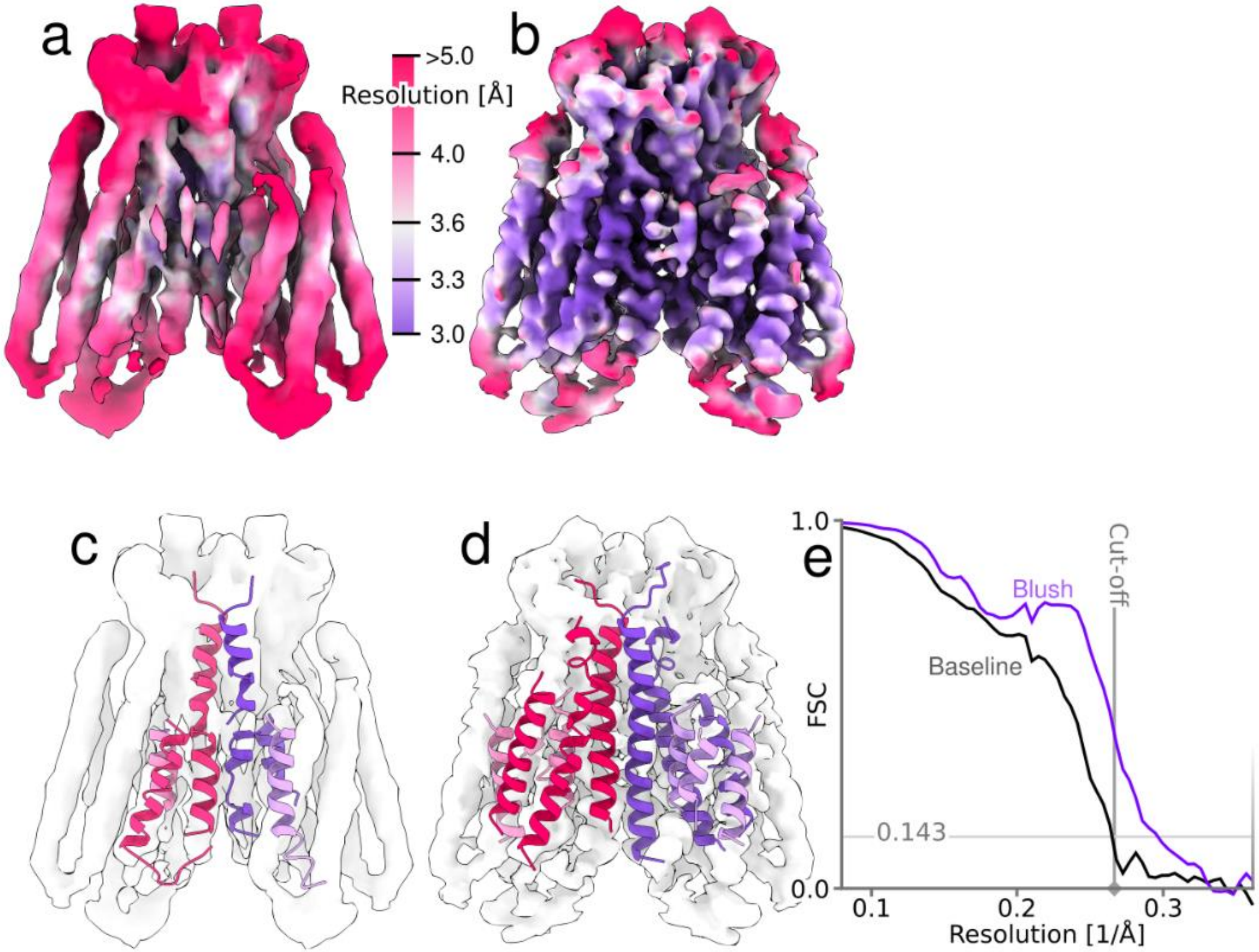
Blush regularisation



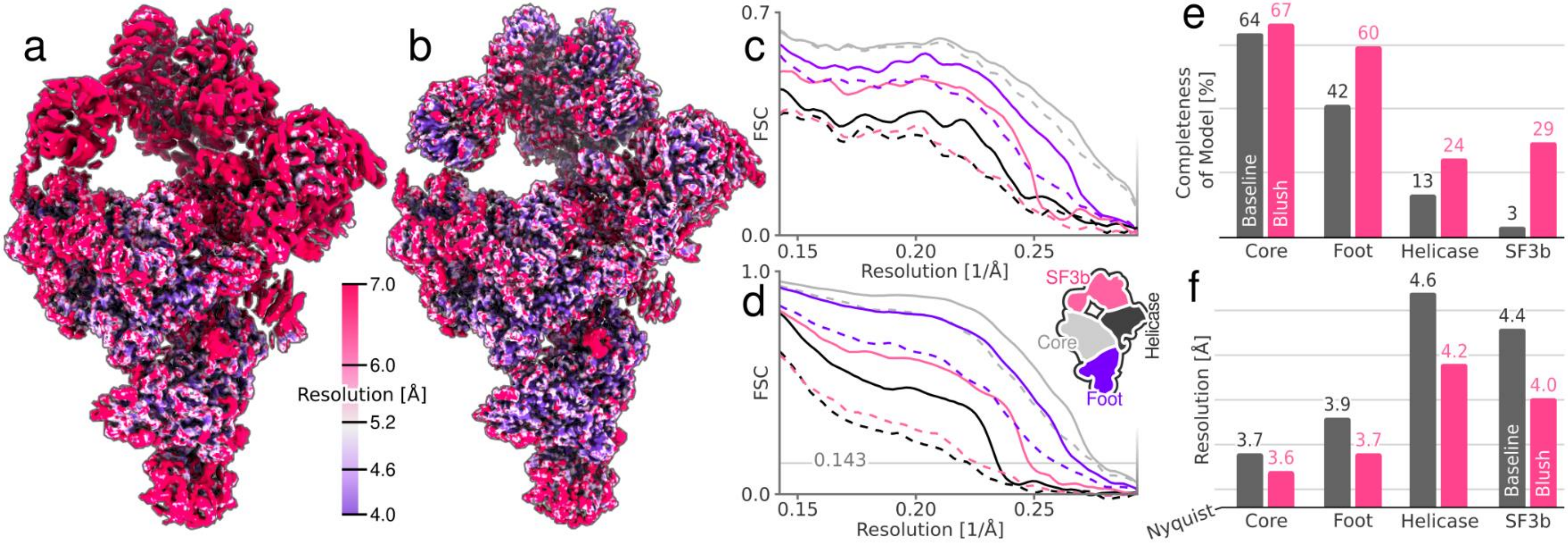
PfCRT (EMPIAR-10330)



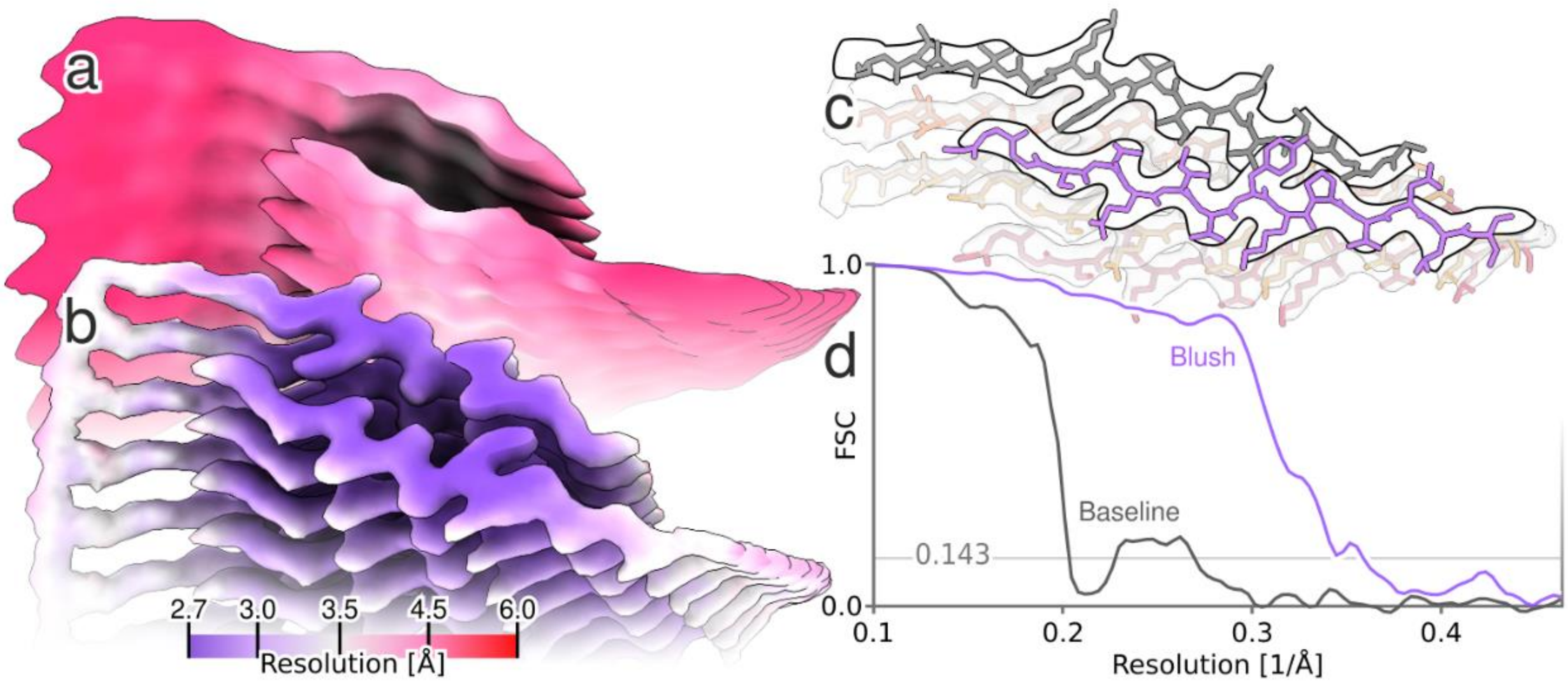
Ste2 (GPCR dimer)



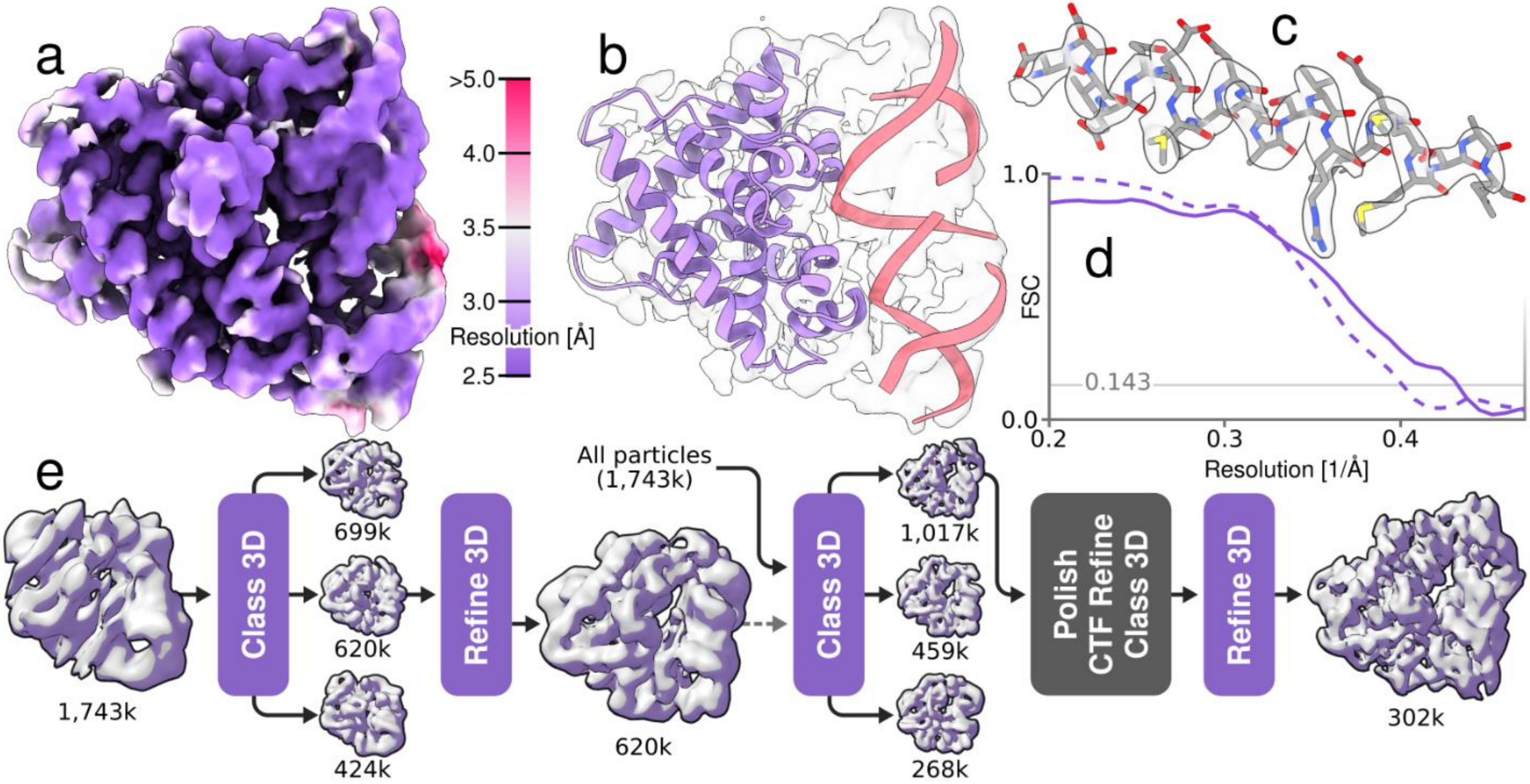
Multibody refinement



A small amyloid (tau FIA)



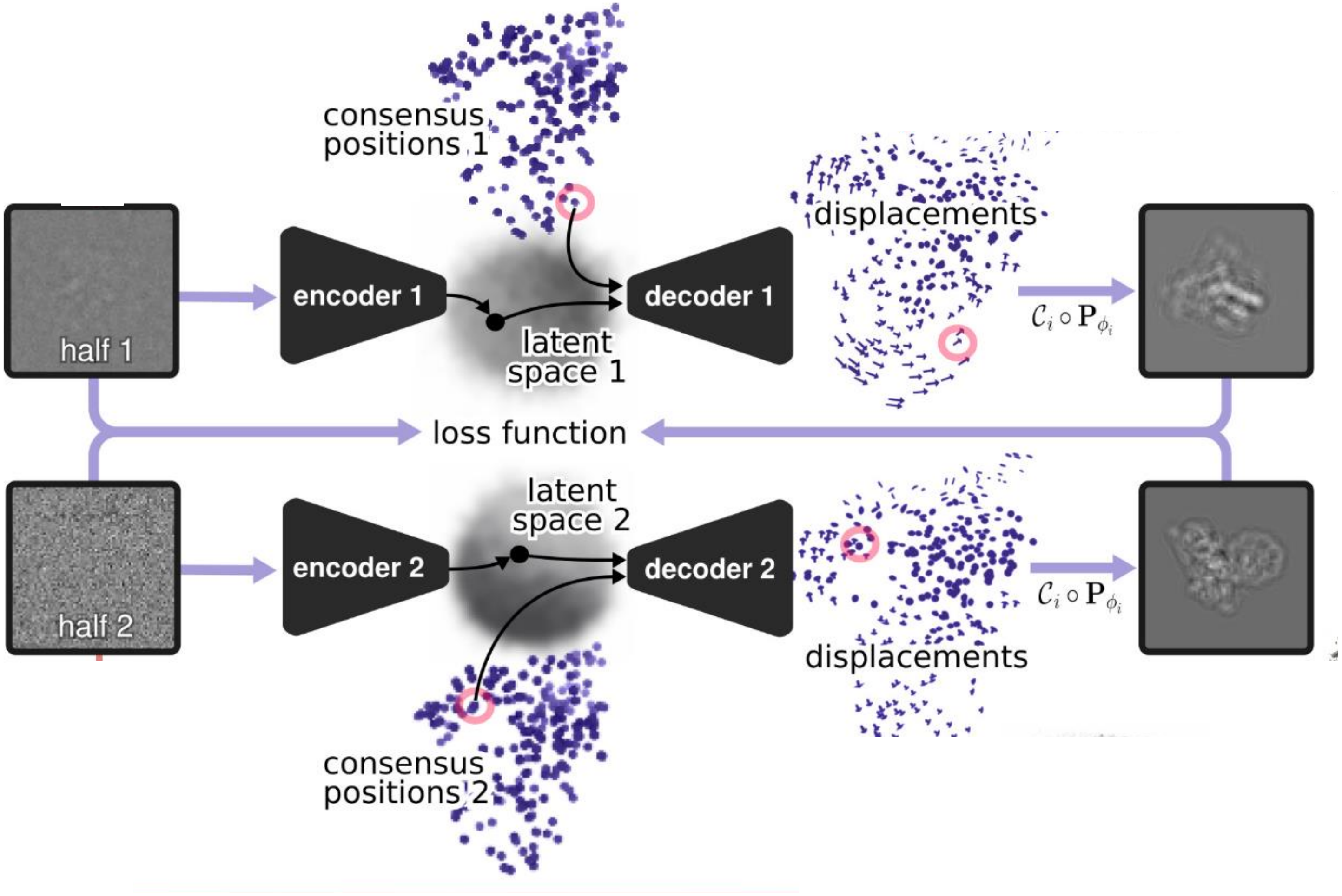
A 40 kDa protein-nucleic acid complex



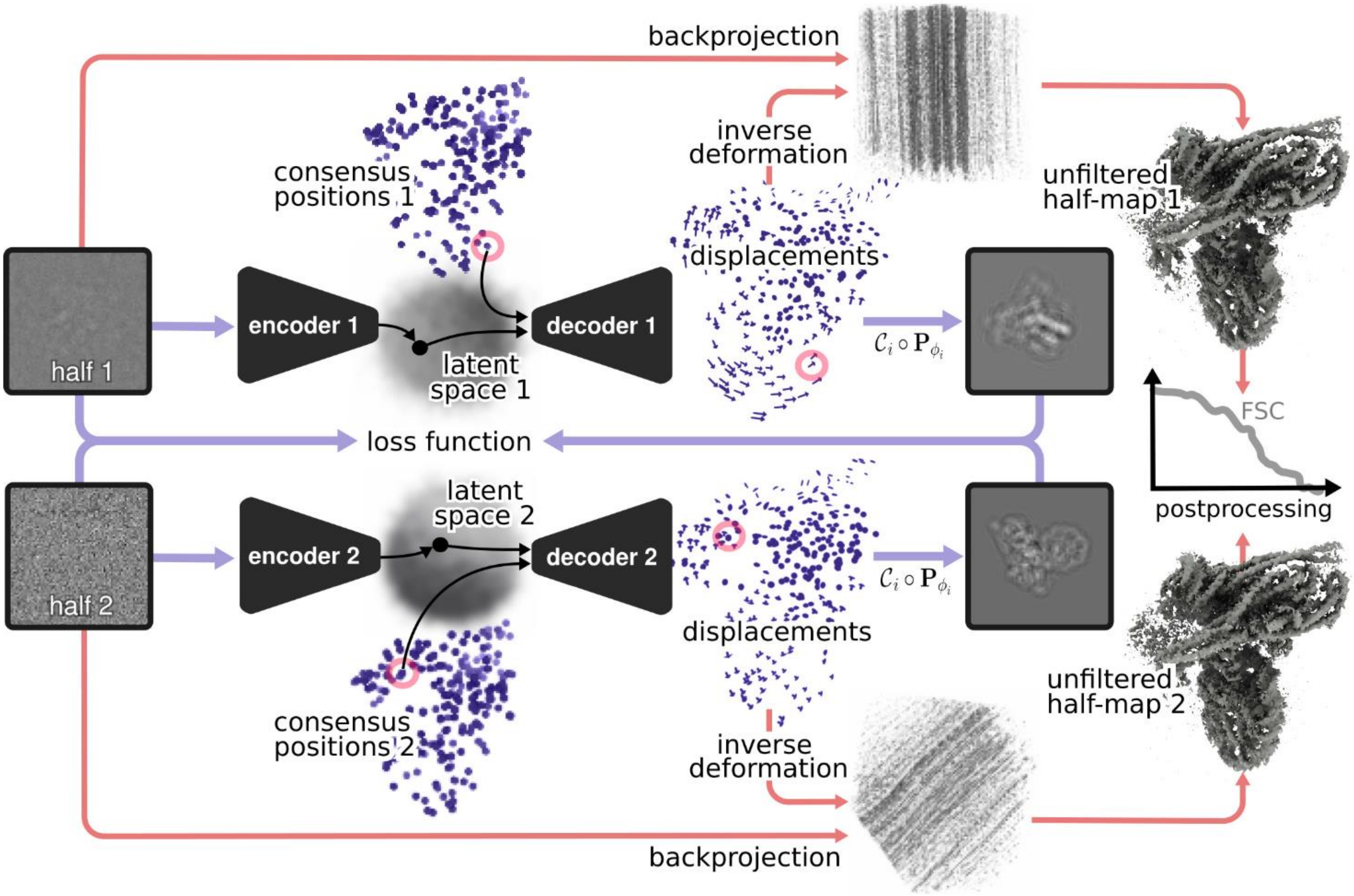
DynaMight:

“modelling structural **Dyna**mics that **Might** improve your map”

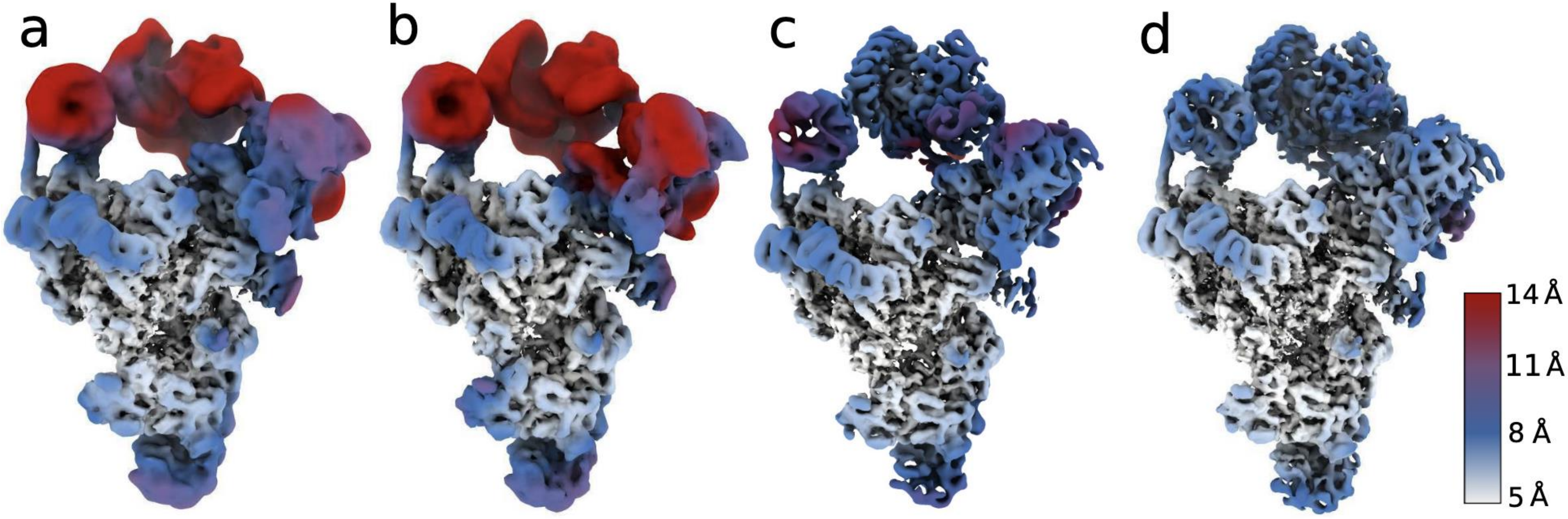
Learnt 3D deformation fields



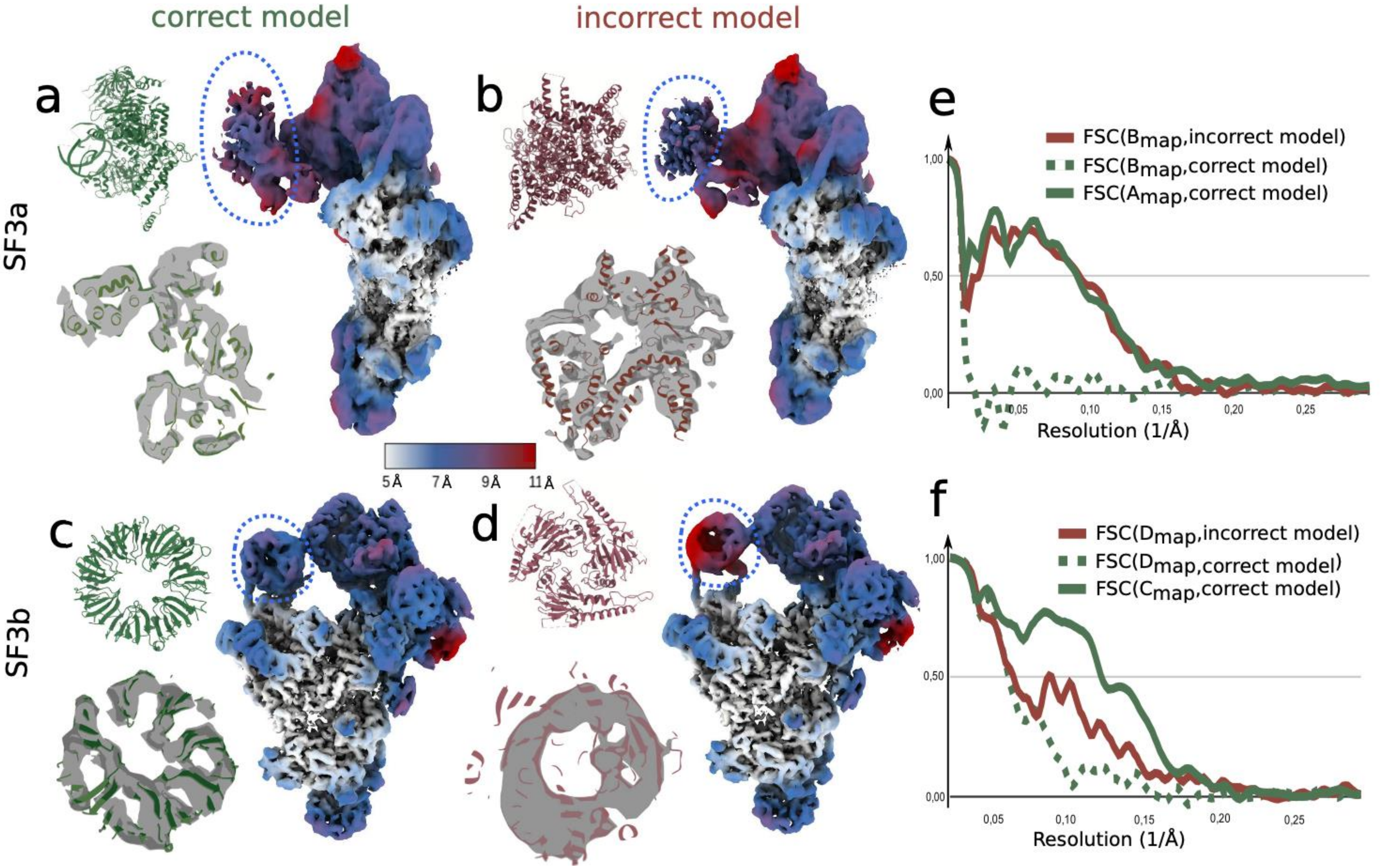
Learnt 3D deformation fields



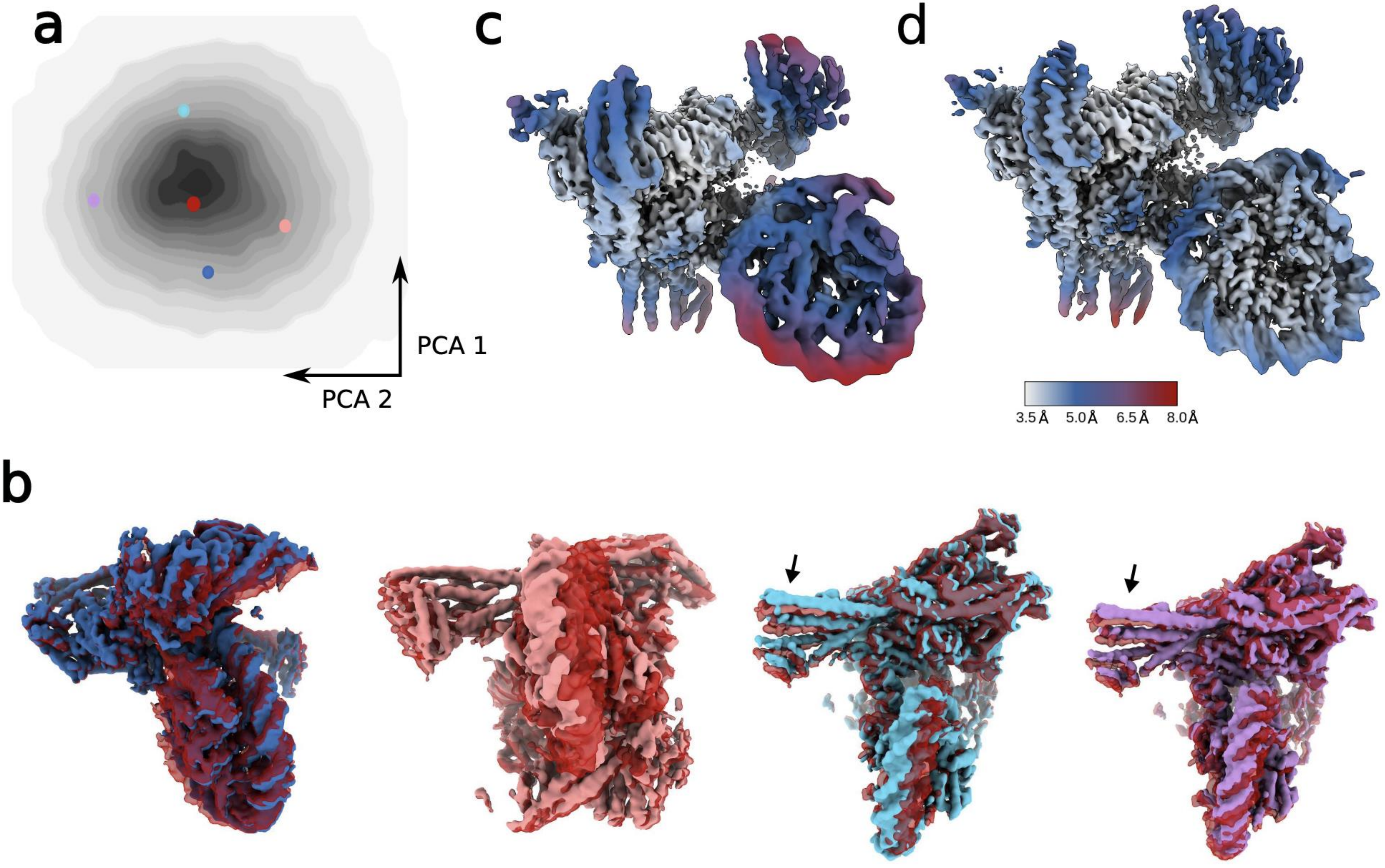
Regularisation of deformations



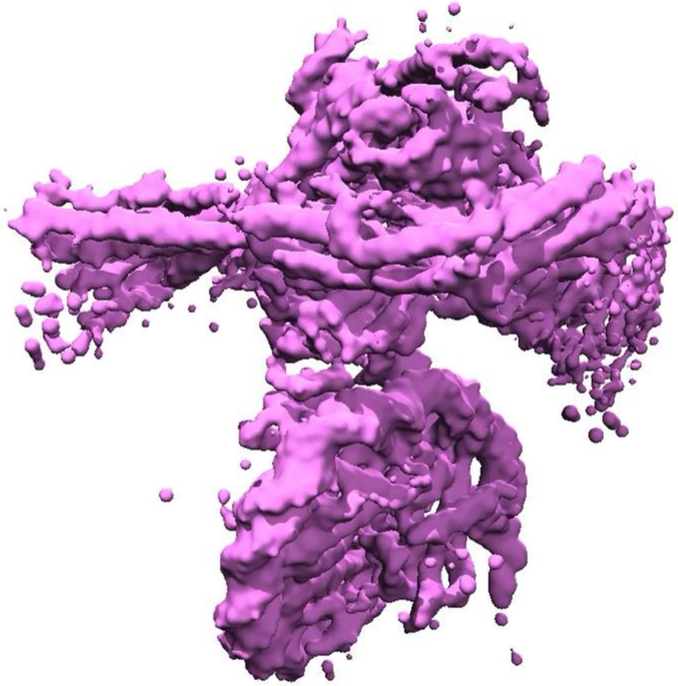
The pitfalls of model bias



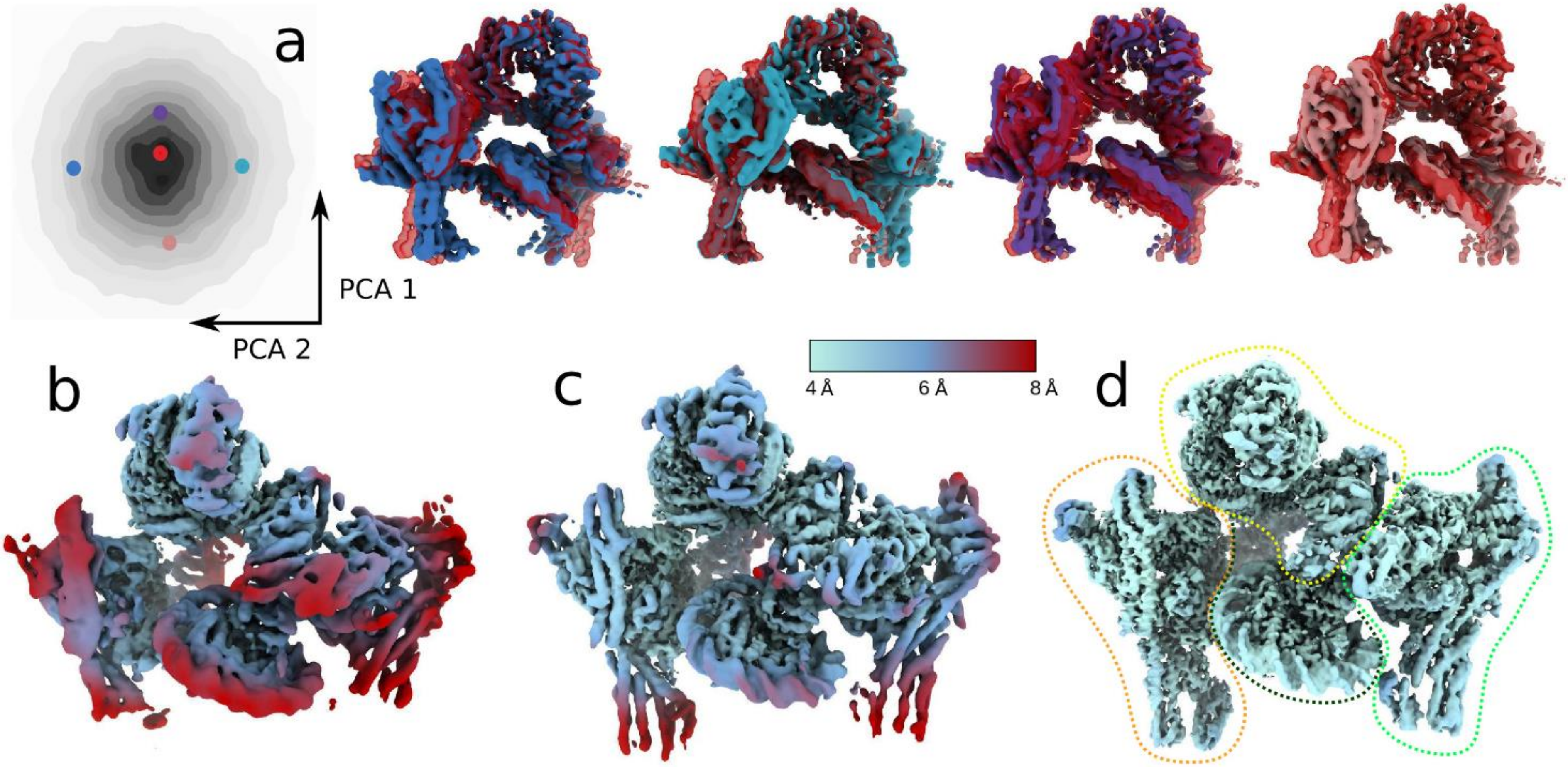
The yeast CCAN:Cenp-A complex



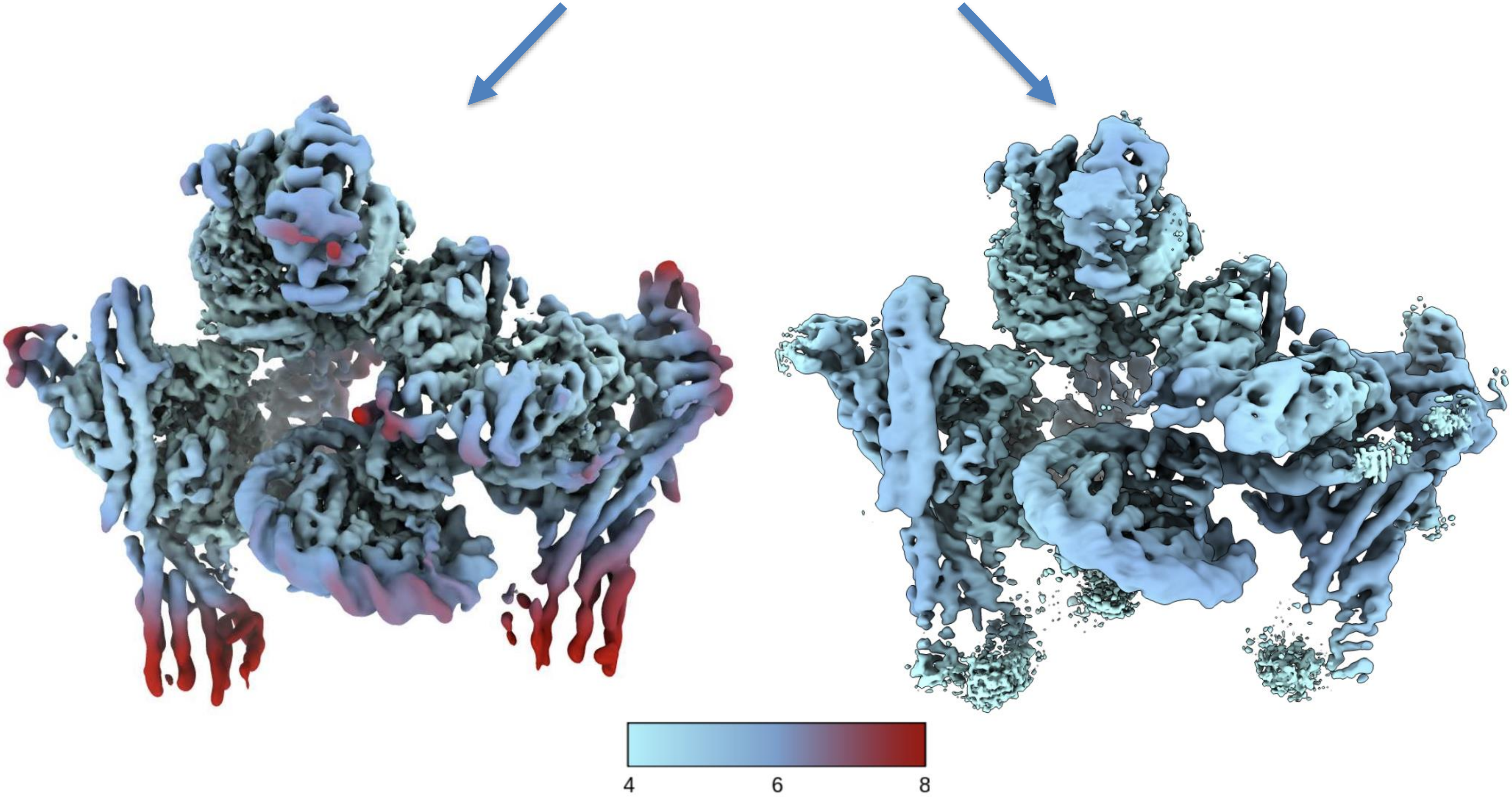
Protein dynamics



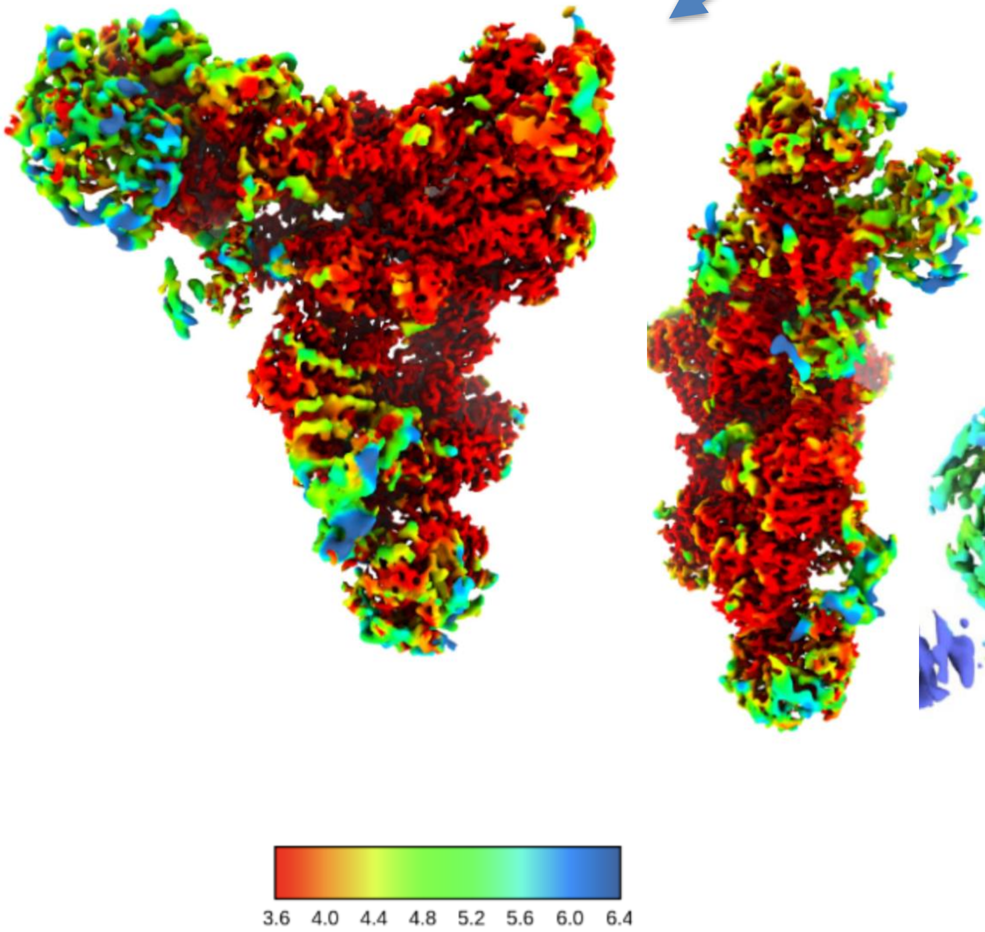
Yeast kinetochore (II)

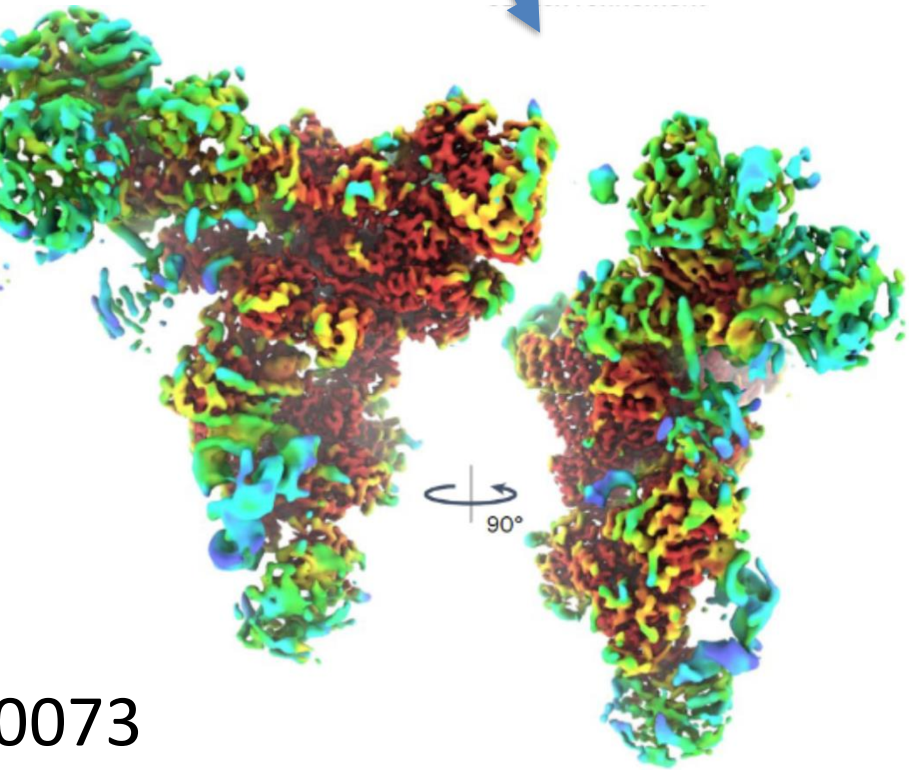


Dynamight vs 3Dflex (I)



Dynamight vs 3Dflex (II)

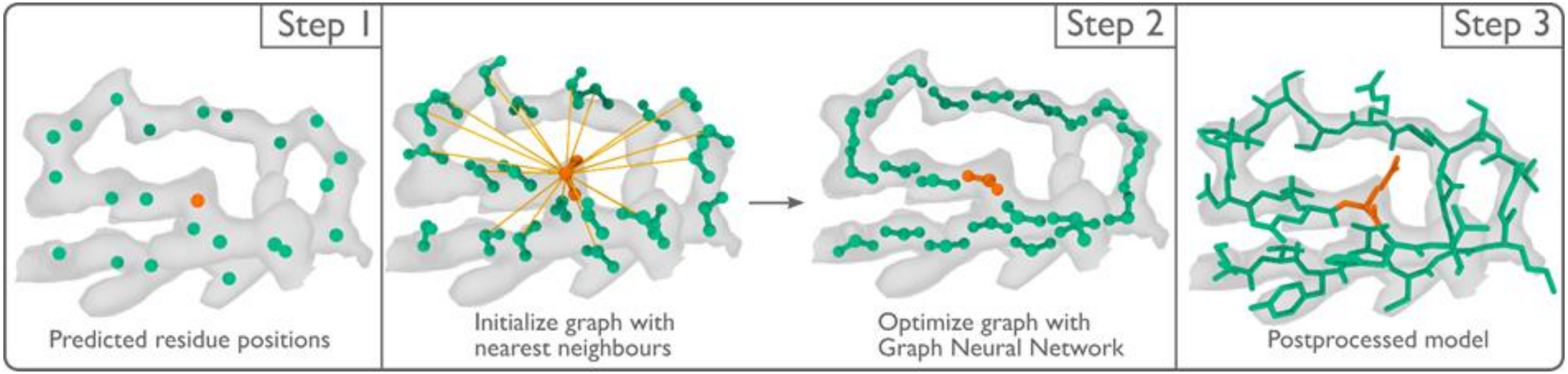




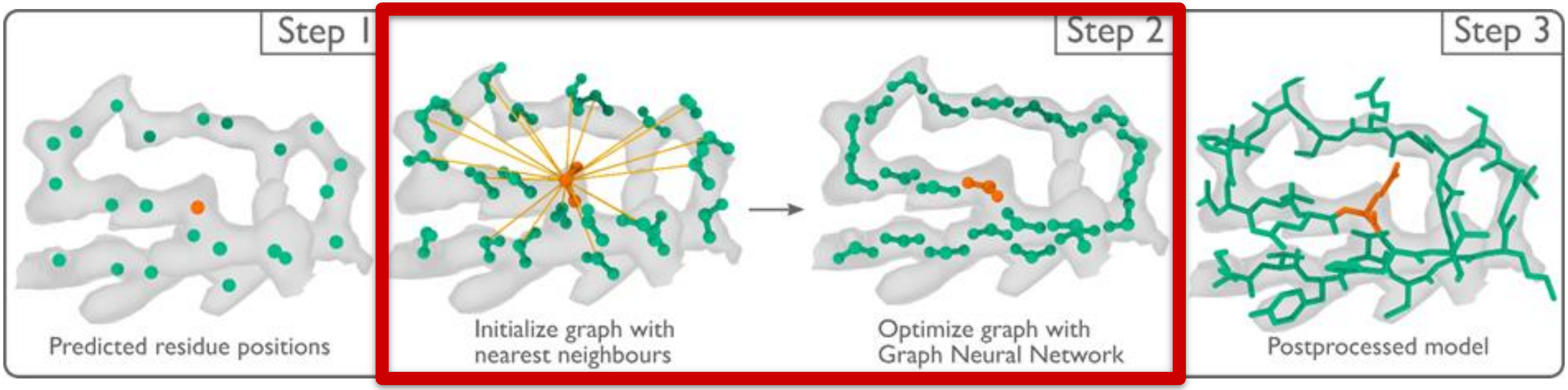
EMPIAR-10073

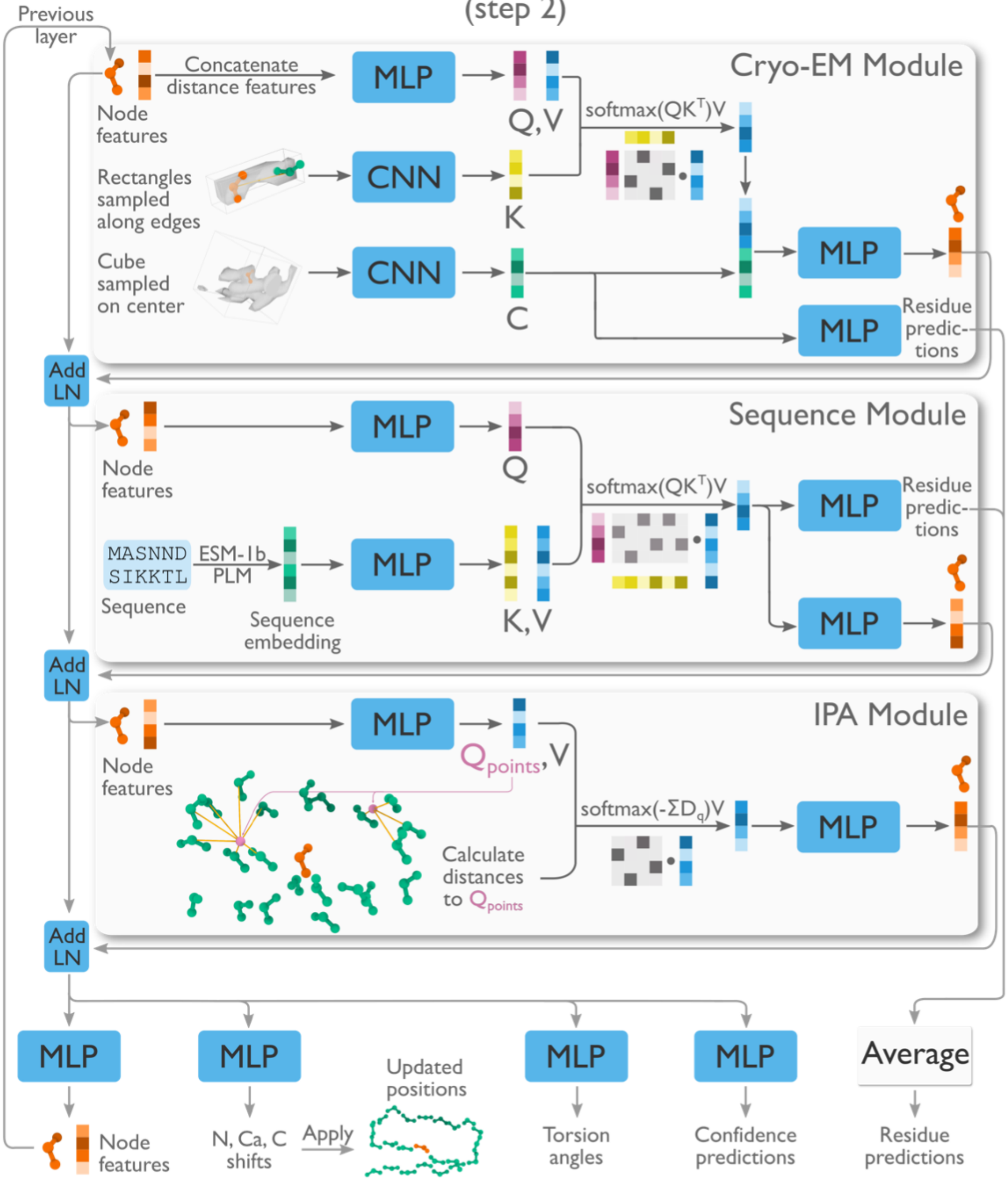
ModelAngelo

ModelAngelo

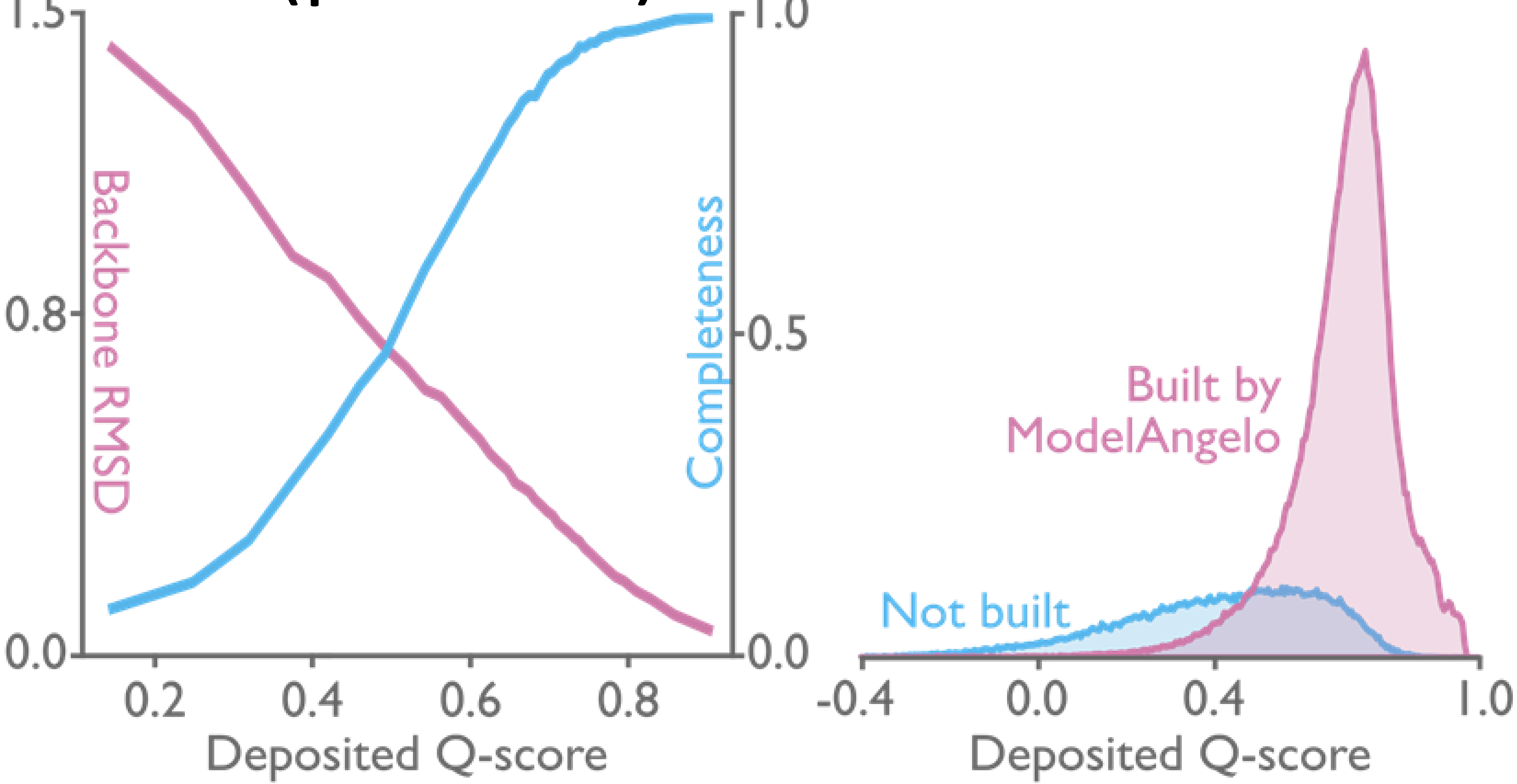


ModelAngelo

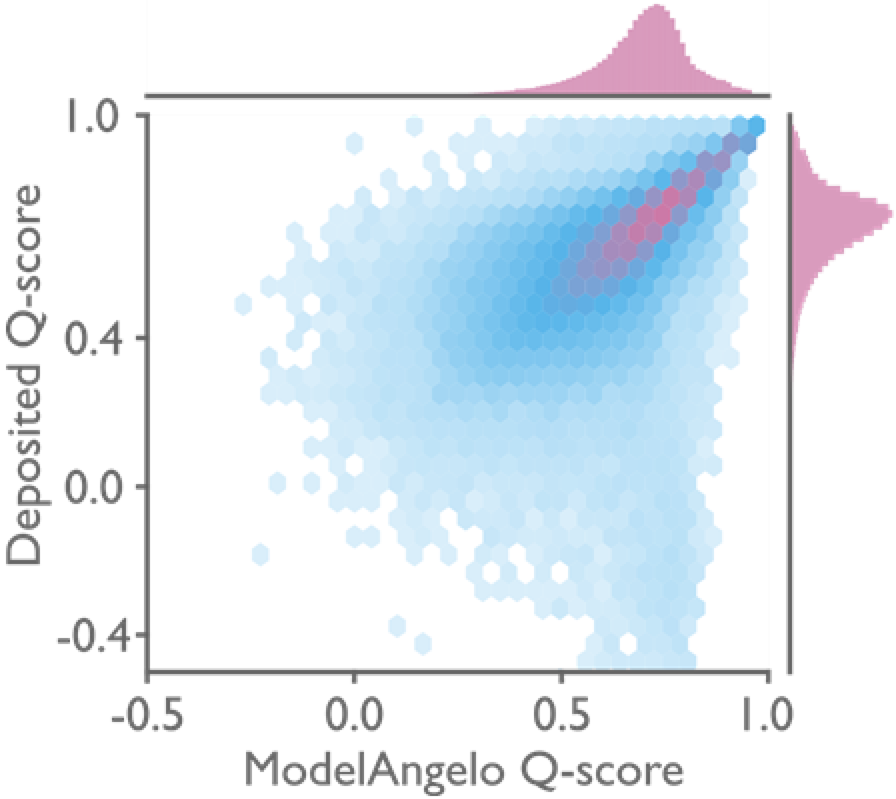


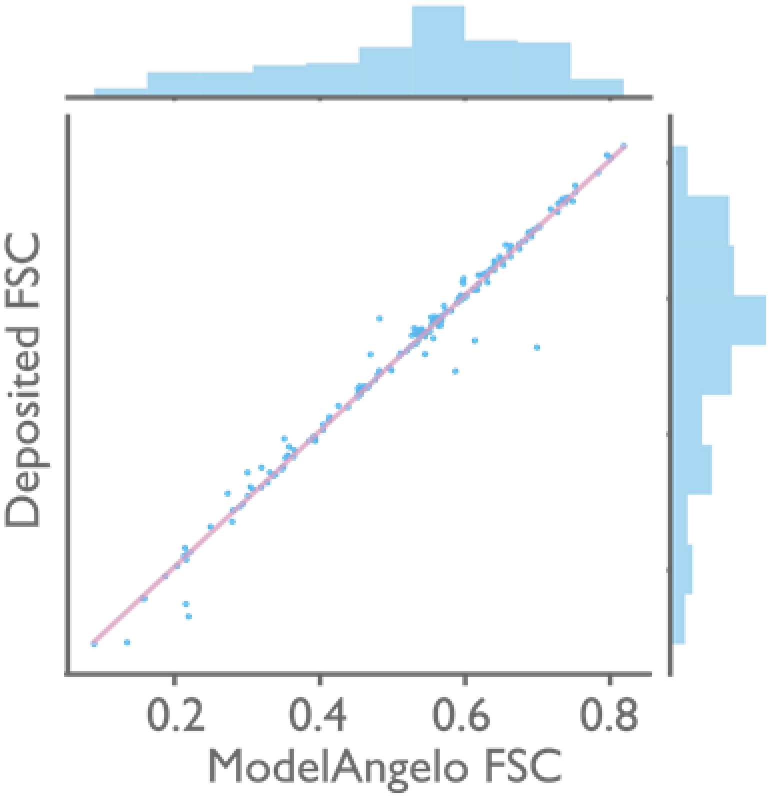

(step 2)
Previous layer
Cryo-EM Module
Node features
Concatenate distance features
MLP
Q,V
softmax(QK^T)V
Rectangles sampled along edges
CNN
K
Cube sampled on center
CNN
C
MLP
MLP
Residue predic-tions
Add LN
Sequence Module
Node features
MLP
Q
softmax(QK^T)V
MLP
Residue predic-tions
MASNND SIKKTL
Sequence
ESM-1b PLM
Sequence embedding
MLP
K,V
MLP
Add LN
IPA Module
Node features
MLP
Q_points, V
softmax(-ΣD_q)V
MLP
Calculate distances to Q_points
Add LN
MLP
Node features
MLP
N, Ca, C shifts
Apply
Updated positions
MLP
Torsion angles
MLP
Confidence predictions
Average
Residue predictions

Results (proteins)

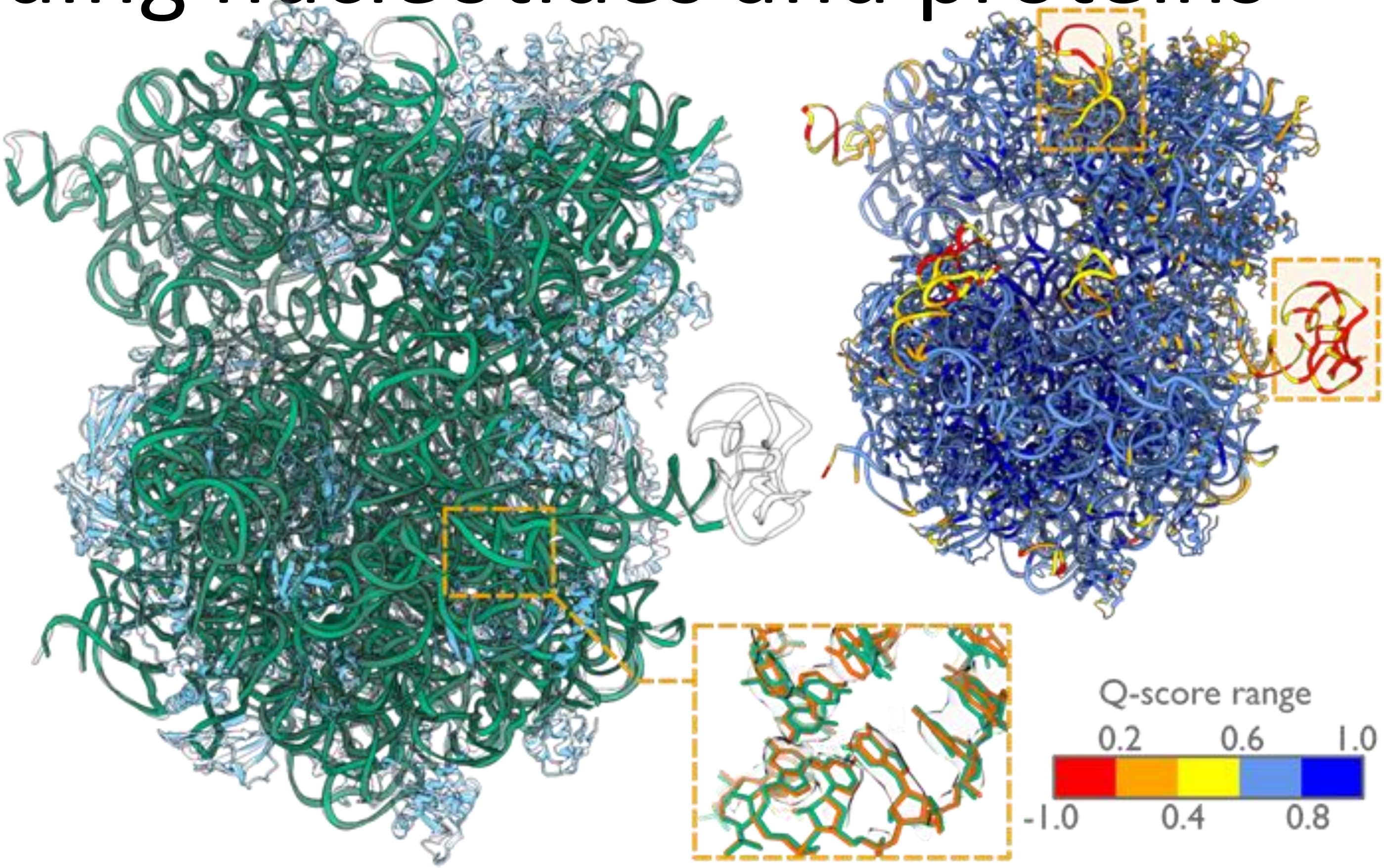


Results (proteins)



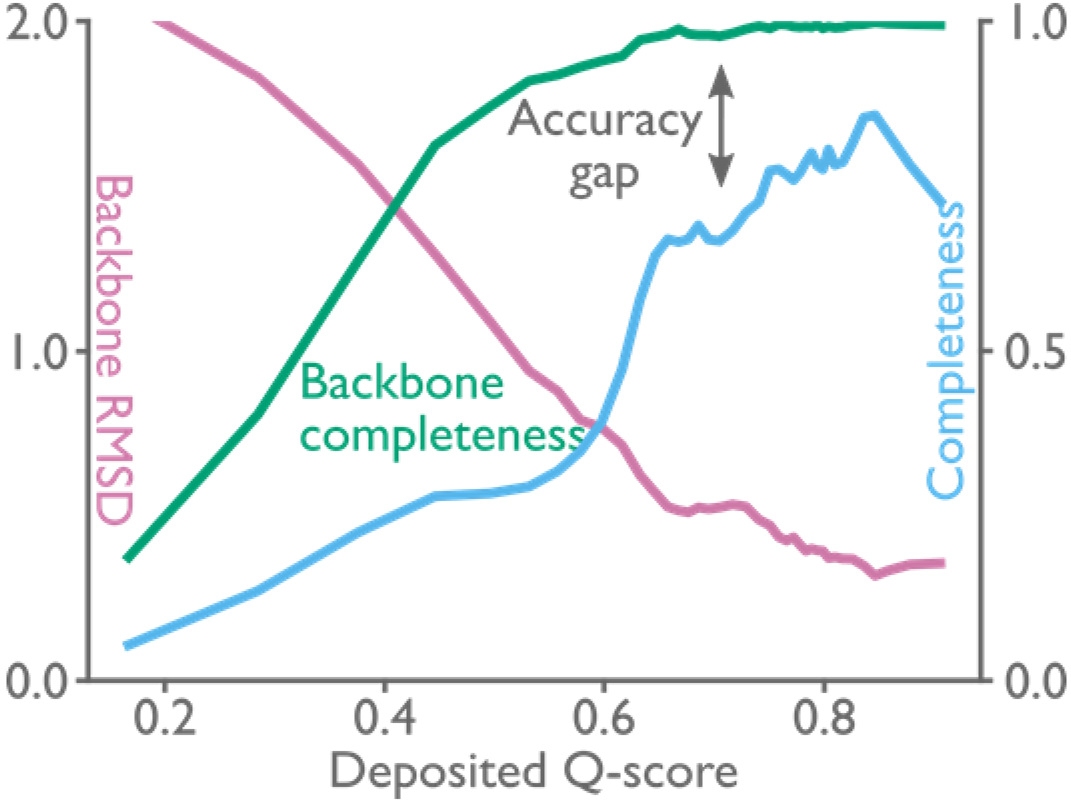


Building nucleotides and proteins

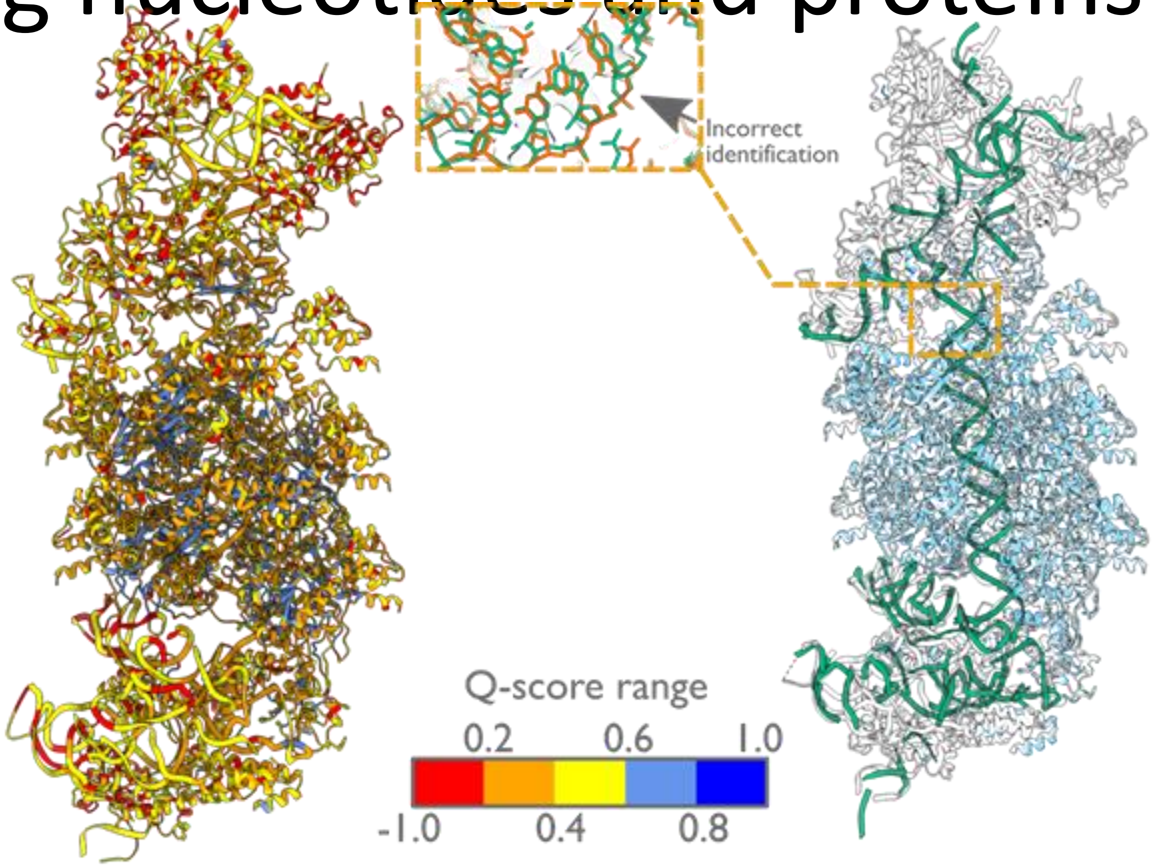


PDB:7S1G

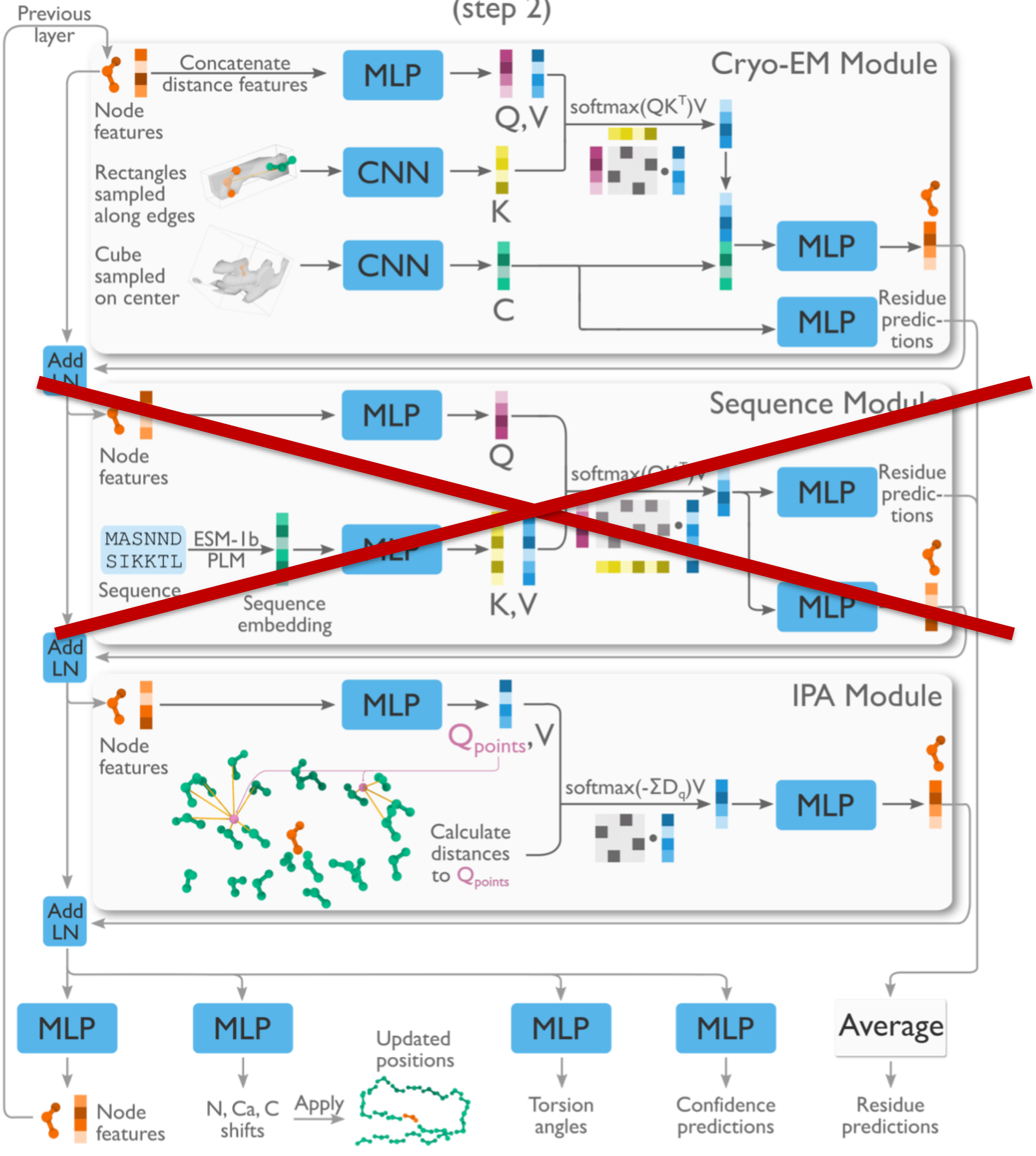
Results (nucleotides)



Building nucleotides and proteins

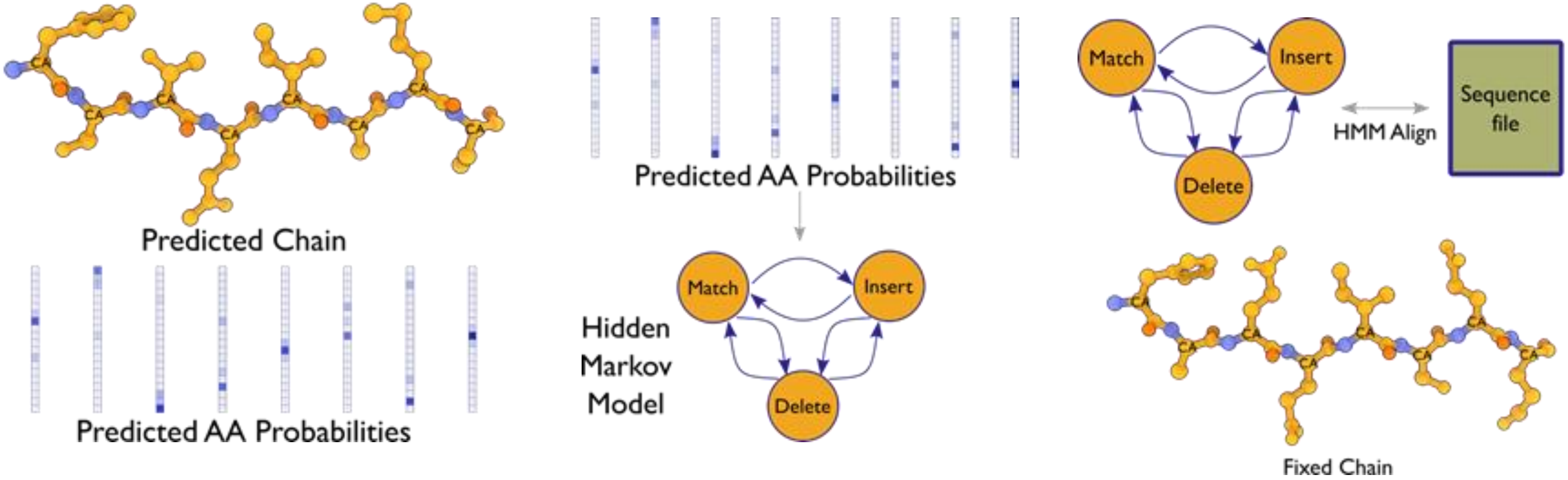


PDB:8EA4


(step 2)
Previous layer
Cryo-EM Module
Node features
Concatenate distance features
MLP
Q,V
softmax(QK^T)V
Rectangles sampled along edges
CNN
K
Cube sampled on center
CNN
C
MLP
MLP
Residue predic-tions
Add LN
Sequence Module
Node features
MLP
Q
softmax(QK^T)V
MLP
Residue predic-tions
MASNND SIKKTL
Sequence
ESM-1b PLM
Sequence embedding
MLP
K,V
MLP
Add LN
IPA Module
Node features
MLP
Q_points, V
softmax(-ΣD_q)V
MLP
Calculate distances to Q_points
Add LN
MLP
MLP
MLP
MLP
Average
Node features
N, Ca, C shifts
Apply
Updated positions
Torsion angles
Confidence predictions
Residue predictions

Protein identification

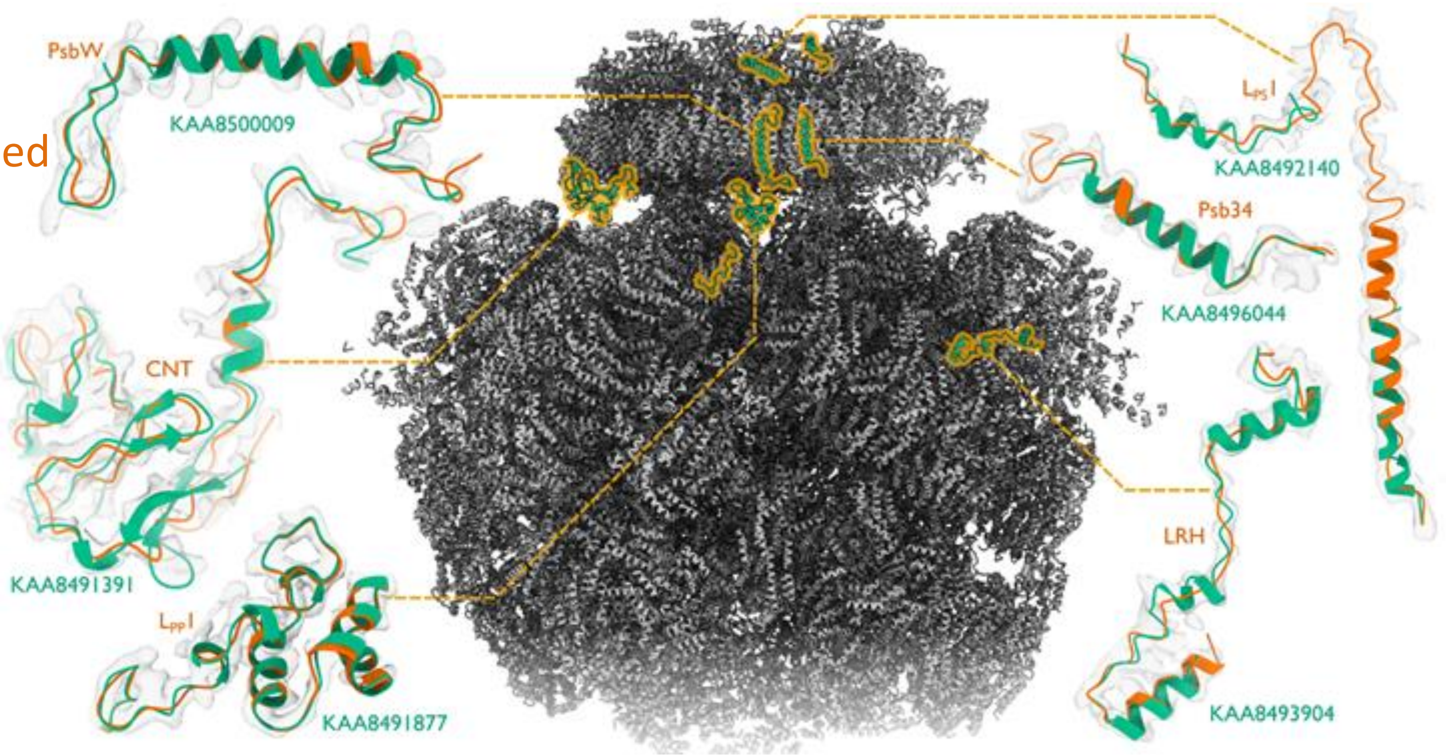
What if we use the proteome as our sequence file?



Phycobilisome

Backbone deposited

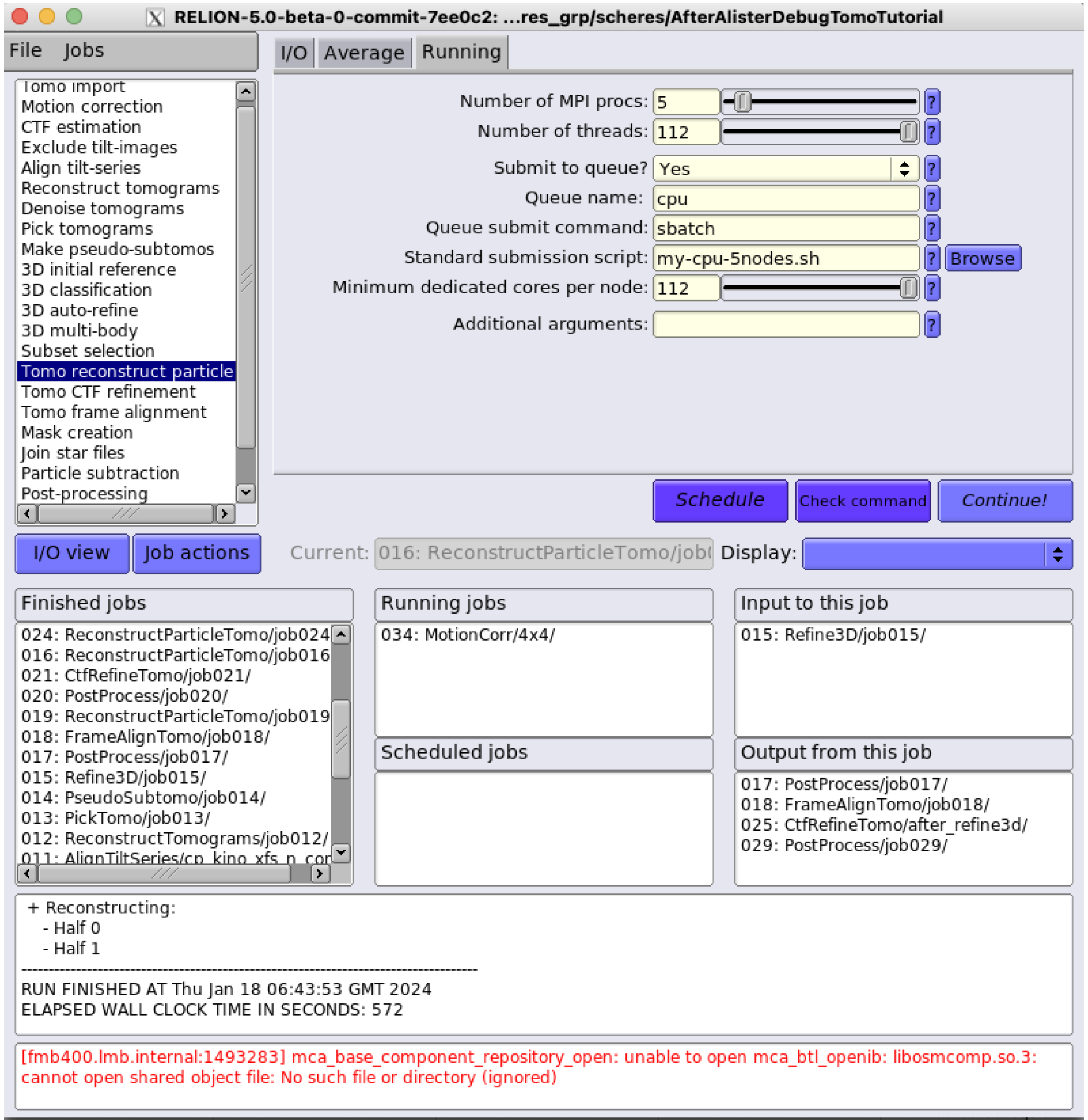
ModelAngelo



You et al. *Nature* 2023

A complete tomography pipeline

relion --tomo



RELION-5.0-beta-0-commit-7ee0c2: ...res_grp/scheres/AfterAlisterDebugTomoTutorial

Number of MPI procs: 5

Number of threads: 112

Submit to queue? Yes

Queue name: cpu

Queue submit command: sbatch

Standard submission script: my-cpu-5nodes.sh

Minimum dedicated cores per node: 112

Additional arguments:

Current: 016: ReconstructParticleTomo/job

Display:

Finished jobs
024: ReconstructParticleTomo/job024
016: ReconstructParticleTomo/job016
021: CtfRefineTomo/job021/
020: PostProcess/job020/
019: ReconstructParticleTomo/job019
018: FrameAlignTomo/job018/
017: PostProcess/job017/
015: Refine3D/job015/
014: PseudoSubtomo/job014/
013: PickTomo/job013/
012: ReconstructTomograms/job012/
011: AlignTiltSeries/cp_kino_xfs_n_cor

Running jobs
034: MotionCorr/4x4/

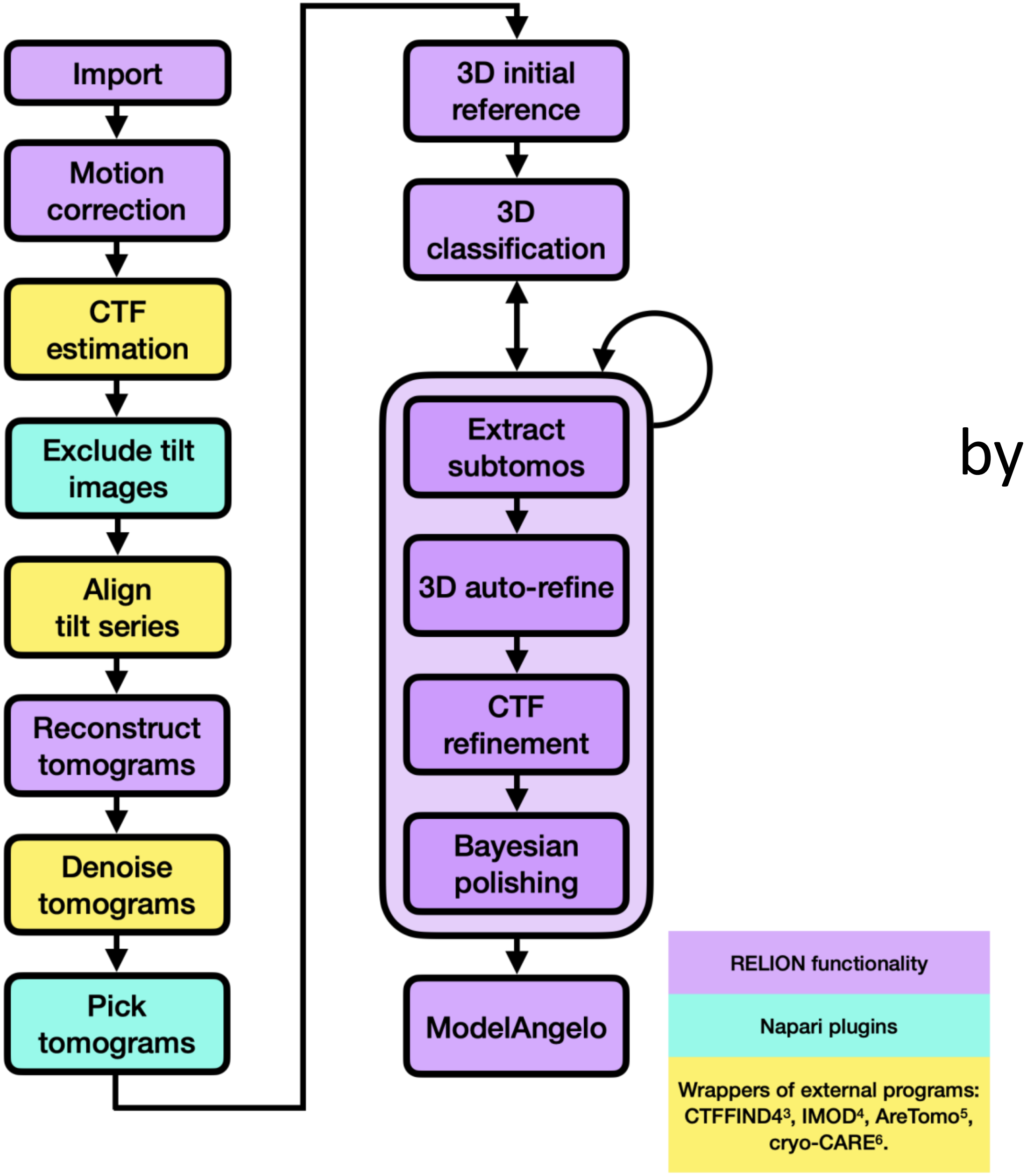
Scheduled jobs

Input to this job
015: Refine3D/job015/

Output from this job
017: PostProcess/job017/
018: FrameAlignTomo/job018/
025: CtfRefineTomo/after_refine3d/
029: PostProcess/job029/

```
 + Reconstructing:
   - Half 0
   - Half 1
--------------------------------------------------------------------------------
RUN FINISHED AT Thu Jan 18 06:43:53 GMT 2024
ELAPSED WALL CLOCK TIME IN SECONDS: 572
```

```
[fmb400.lmb.internal:1493283] mca_base_component_repository_open: unable to open mca_btl_openib: libosmcomp.so.3:
cannot open shared object file: No such file or directory (ignored)
```



Import

Motion correction

CTF estimation

Exclude tilt images

Align tilt series

Reconstruct tomograms

Denoise tomograms

Pick tomograms

3D initial reference

3D classification

Extract subtomos

3D auto-refine

CTF refinement

Bayesian polishing

ModelAngelo

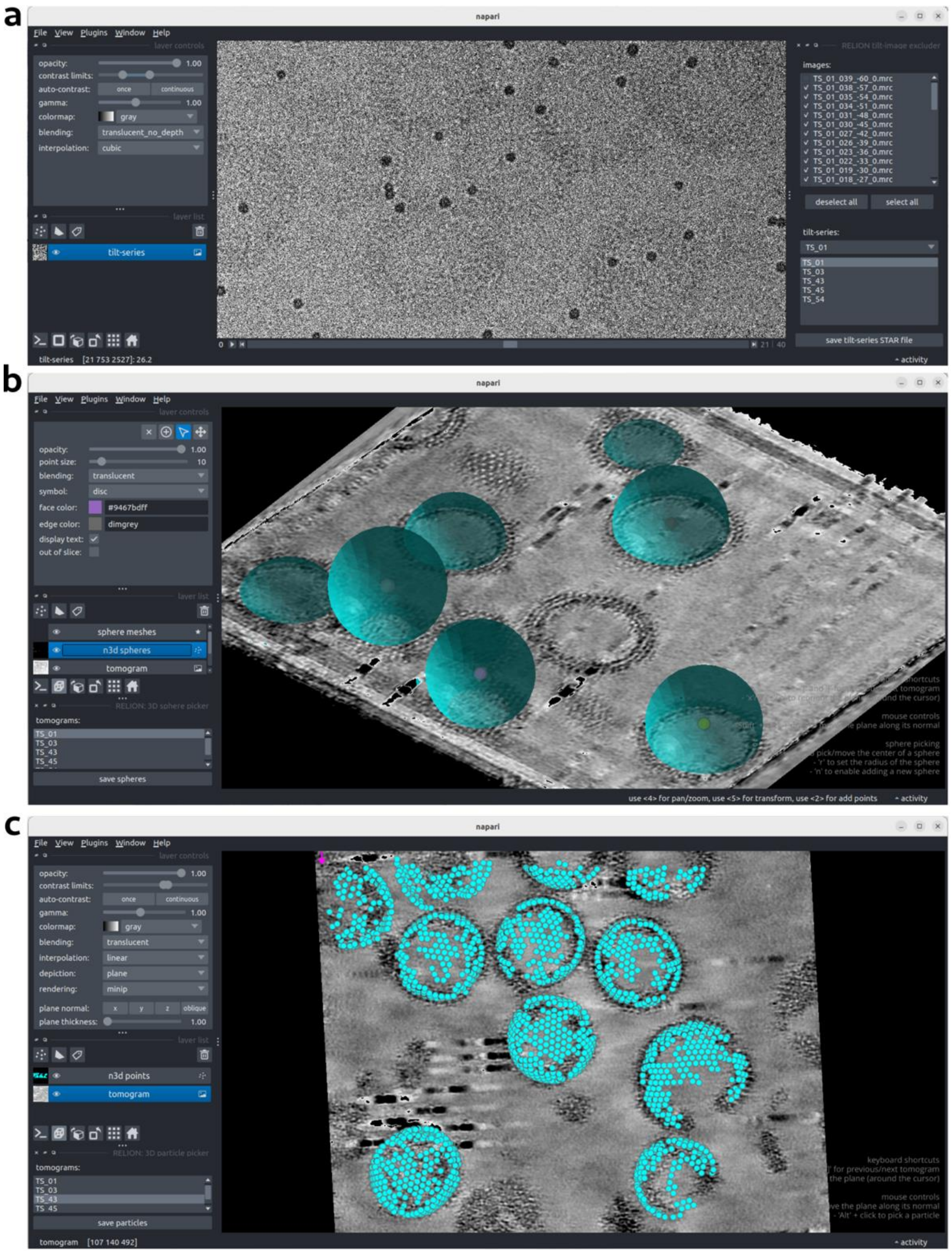
RELION functionality

Napari plugins

Wrappers of external programs: CTFFIND4[3], IMOD[4], AreTomo[5], cryo-CARE[6].

Poster 83
by Bogdan Toader

Napari-based picker



opacity: 1.00
contrast limits:
gamma: 1.00
colormap: gray
blending: translucent_no_depth
interpolation: cubic

tilt-series

tilt-series [21 753 2527]: 26.2

0 21 | 40

images:

☐ TS_01_039_-60_0.mrc
✓ TS_01_038_-57_0.mrc
✓ TS_01_035_-54_0.mrc
✓ TS_01_034_-51_0.mrc
✓ TS_01_031_-48_0.mrc
✓ TS_01_030_-45_0.mrc
✓ TS_01_027_-42_0.mrc
✓ TS_01_026_-39_0.mrc
✓ TS_01_023_-36_0.mrc
✓ TS_01_022_-33_0.mrc
✓ TS_01_019_-30_0.mrc
✓ TS_01_018_-27_0.mrc

tilt-series: TS_01

TS_01
TS_03
TS_43
TS_45
TS_54

opacity: 1.00
point size: 10
blending: translucent
symbol: disc
face color: #9467bdff
edge color: dimgrey
display text: ☑
out of slice: ☐

sphere meshes
n3d spheres
tomogram

tomograms:

TS_01
TS_03
TS_43
TS_45

shortcuts
tomogram
the cursor)

mouse controls
plane along its normal

sphere picking
pick/move the center of a sphere
- 'r' to set the radius of the sphere
- 'n' to enable adding a new sphere

use <4> for pan/zoom, use <5> for transform, use <2> for add points

opacity: 1.00
contrast limits:
gamma: 1.00
colormap: gray
blending: translucent
interpolation: linear
depiction: plane
rendering: minip
plane normal:
plane thickness: 1.00

n3d points
tomogram

tomograms:

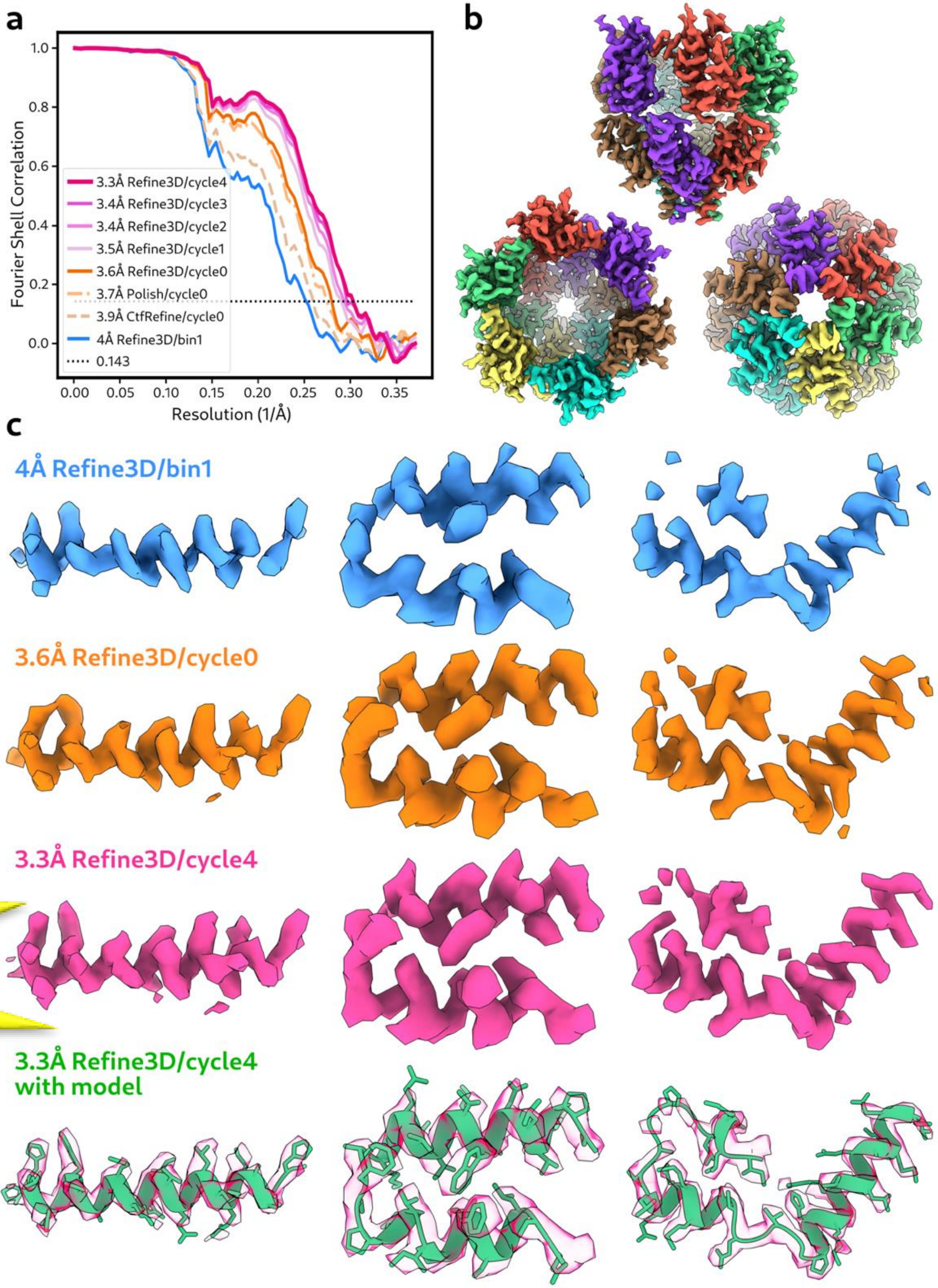
TS_01
TS_03
TS_43
TS_45

keyboard shortcuts
]' for previous/next tomogram
the plane (around the cursor)

mouse controls
ve the plane along its normal
- 'Alt' + click to pick a particle

tomogram [107 140 492]

HIV immature capsids (5tomos)



~48 hours from import to atomic model

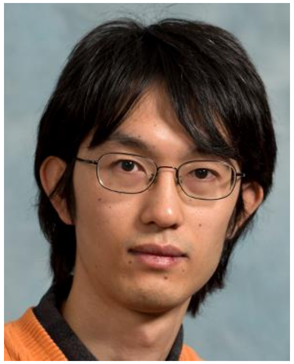
Cryo-ET pipeline documentation

- https://relion.readthedocs.io/release-5.0
  - NOW UPDATED!
  - Pre-calculated results from Zenodo

Conclusions

- Blush makes difficult refinements better
- DynaMight improves maps from data sets with continuous flexibility
  - But be careful with atomic models as input!
- ModelAngelo builds models automatically
  - as good as humans for protein building
  - better than humans at identifying unknown proteins
- A start-end tomo pipeline will make sub-tomogram averaging a lot easier
  - Data model also to be adopted by CCP-EM
- #OpenSoftwareAcceleratesScience
  - CCP-EM pipeliner & Doppio will be important!

Thanks!


Takanori
Nakane


Dari
Kimanius


Johannes
Schwab


Kiarash
Jamali

#teamtomo


Bogdan
Toader


Jasenko
Zivanov


Alister
Burt


Rangana
Warshamanage

- Rafa Leiro (CNIO)
- CCP-EM
- Ozan Öktem
- Carola Schönlieb
- LMB EM facility
- Computing @LMB

- ○ Global Phasing
- ○ Turing Institute

UKRI MRC Laboratory of Molecular Biology

@SjorsScheres

#OpenSoftwareAcceleratesScience