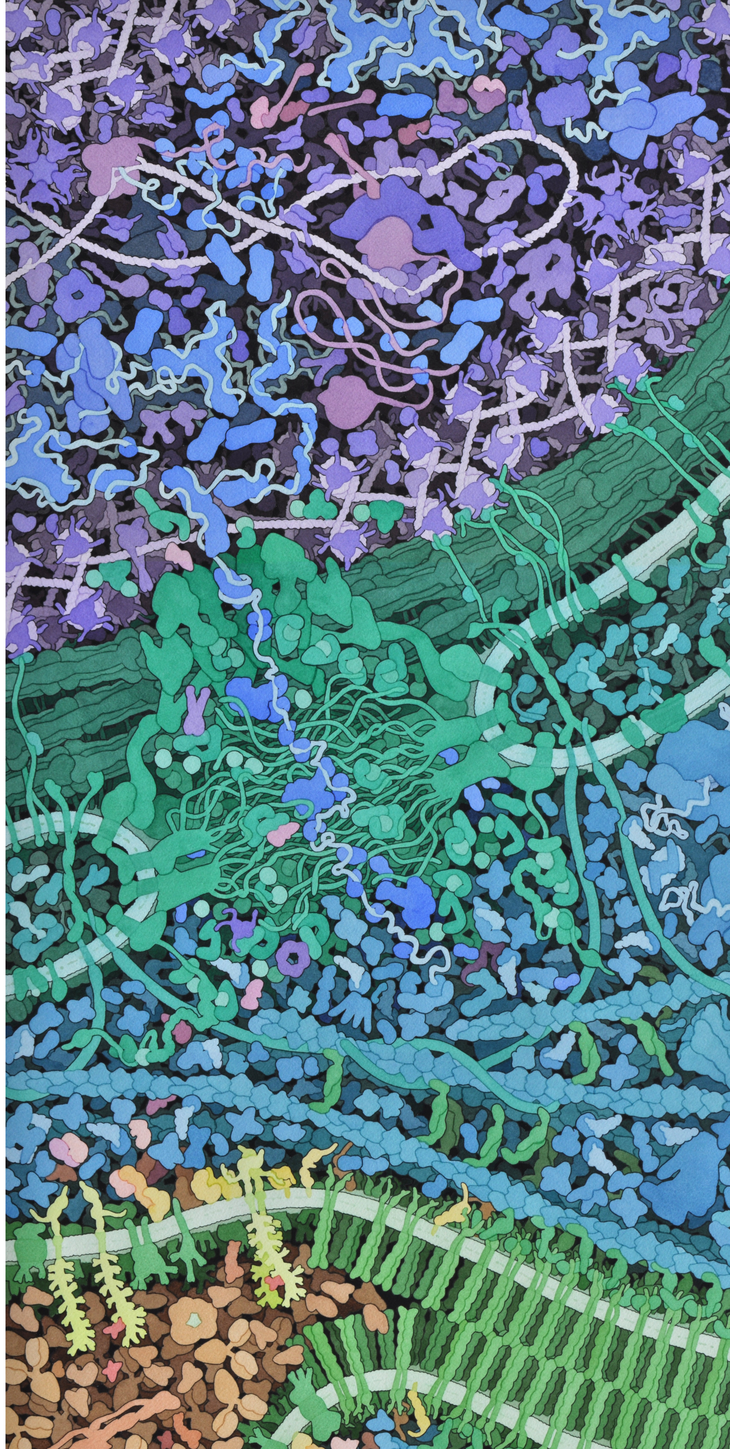
Pose estimation for heterogeneous ab initio reconstruction from cryo-EM

CryoDRGN results, cryo-EM metrics, CryoDRGN-AI, and future directions

Robert Heeter
COS 598L

24 February 2026
38 slides



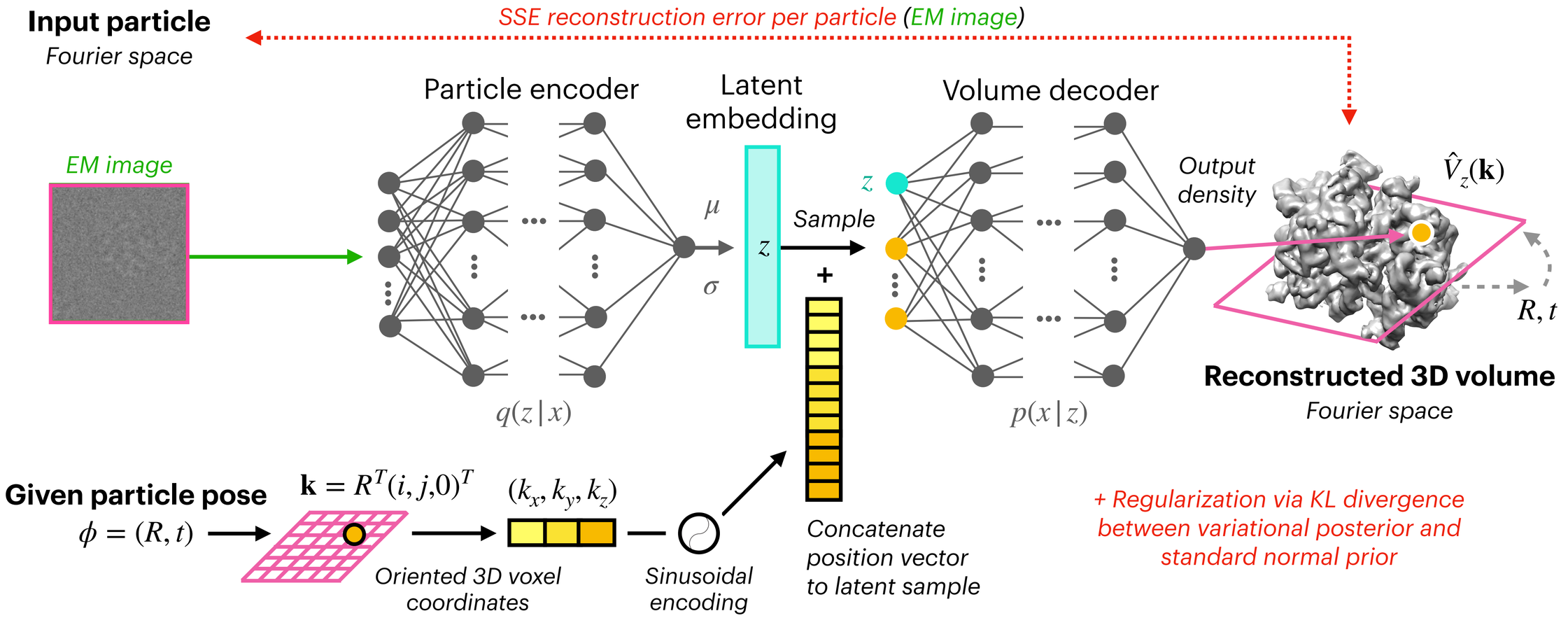
▸1. CryoDRGN Results

2. Pose Estimation
3. CryoDRGN-AI Method
4. CryoDRGN-AI Results
5. Future Directions

Roadmap

CryoDRGN: Variational autoencoder for heterogeneity analysis

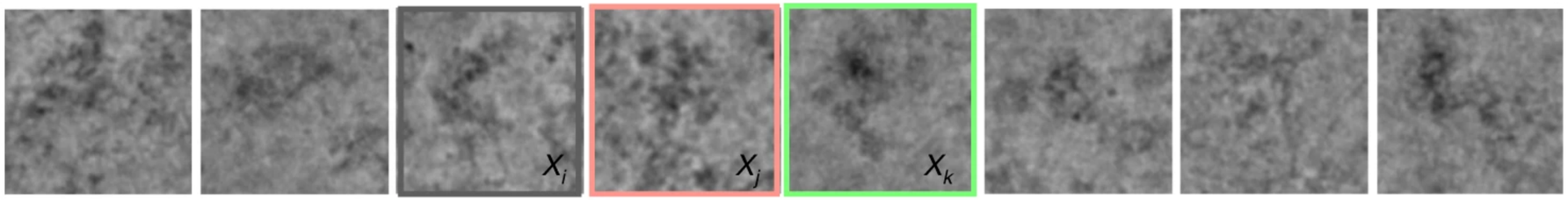
Latent manifold encodes conformation of particle, which conditions 3D reconstruction

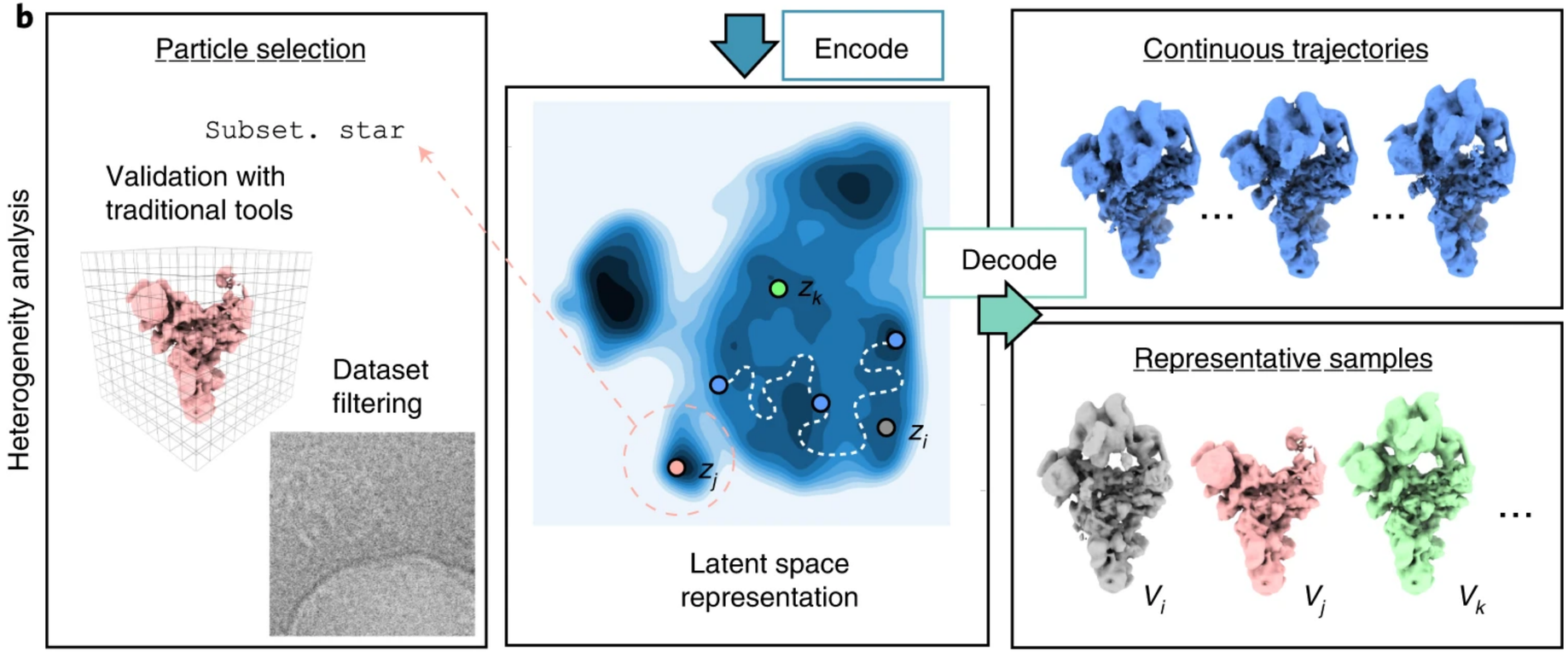


Zhong et al, *Nat Methods*, 2021

Analyzing the generative model

- Use the encoder network to evaluate the latent embedding **z** for each image
- Use the decoder network to generate **V** at different values of **z**





Zhong *et al*. ICLR 2020. Nature Methods 2021

Aside: Reconstruction of simulated datasets with GT structures

Point spread function (CTF)

Projection about 5D pose

3D volume

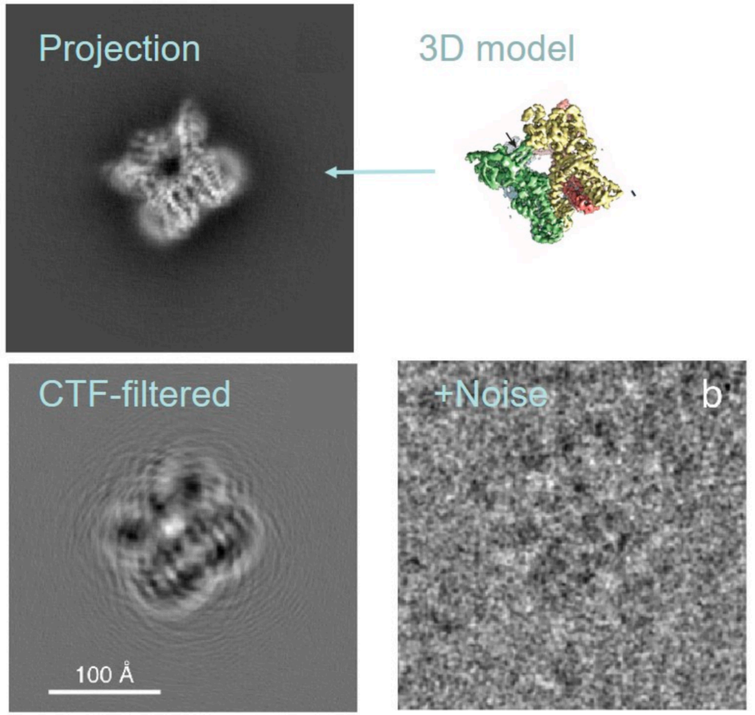
2D image of particle i

$$Y_i = C_i * P_{\phi_i} V_i + \eta_i$$ ····· *Gaussian noise*

Complex aberrations

Orientation bias
Anisotropic pose distribution

Multiple sources of noise
Shot, detector, impurities, radiation damage, ice



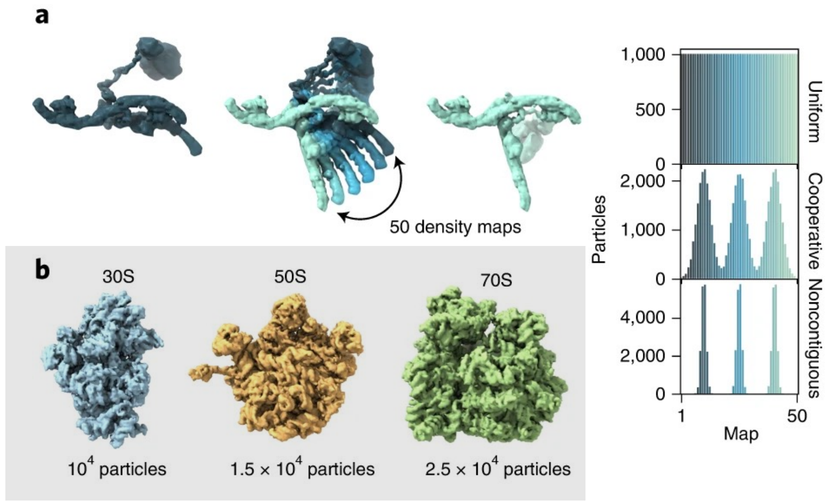
Amit Singer

Simulated

Reconstruction of simulated datasets with GT structures

#1: Sample along 1D transition for conformational heterogeneity

#2: Sample different ribosome complexes for compositional heterogeneity

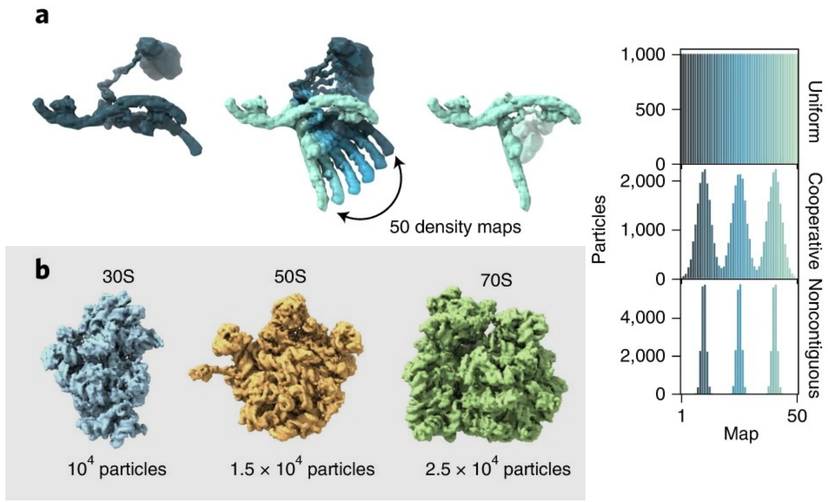


Simulated

Reconstruction of simulated datasets with GT structures

#1: Sample along 1D transition for conformational heterogeneity

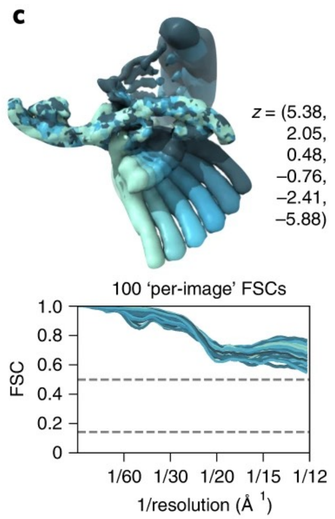
#2: Sample different ribosome complexes for compositional heterogeneity

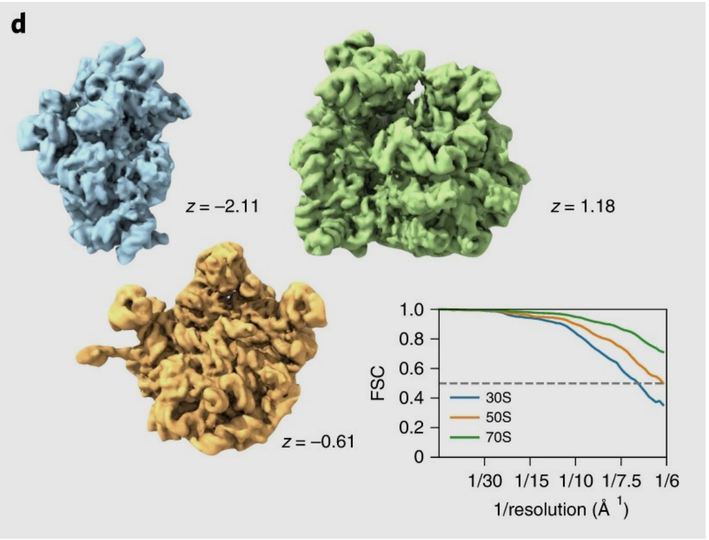


CryoDRGN reconstructions

#1: Conformational

#2: Compositional





Simulated

Reconstruction of simulated datasets with GT structures

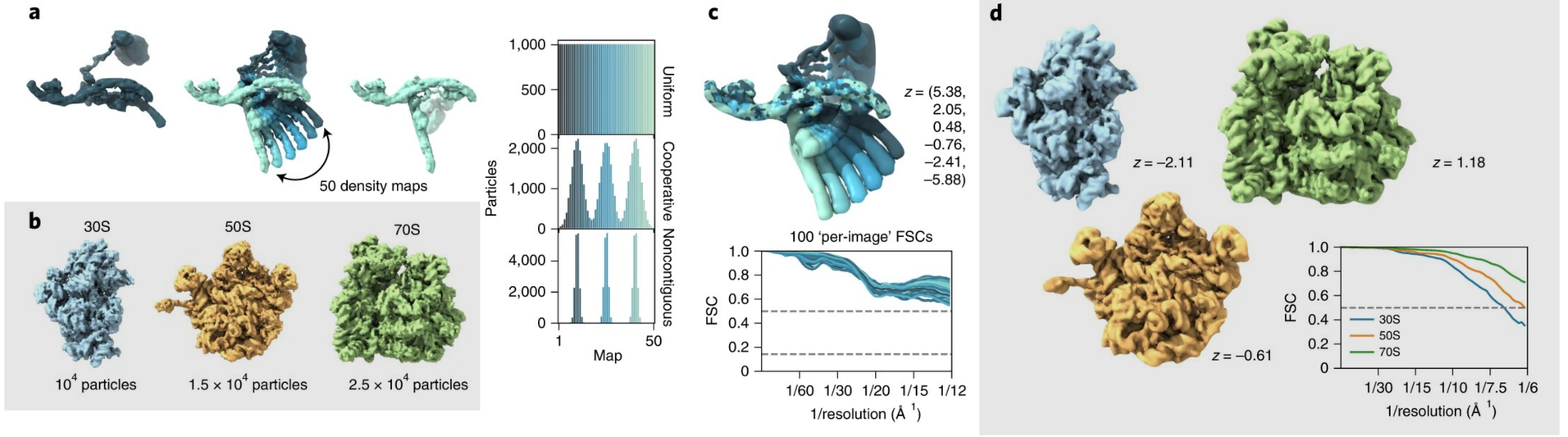
#1: Sample along 1D transition for conformational heterogeneity

CryoDRGN reconstructions

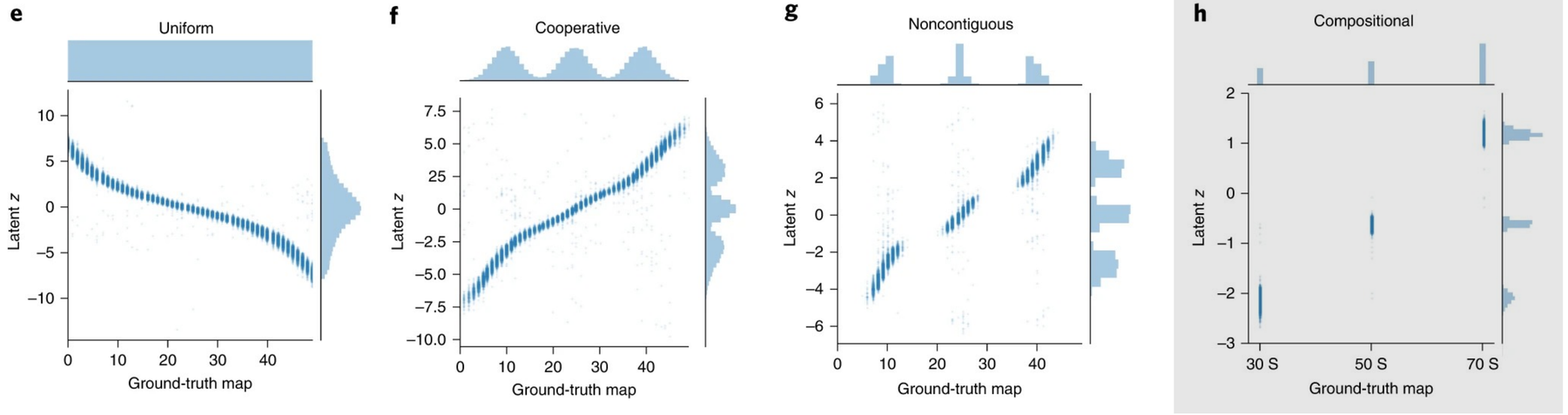
#1: Conformational

#2: Compositional

#2: Sample different ribosome complexes for compositional heterogeneity

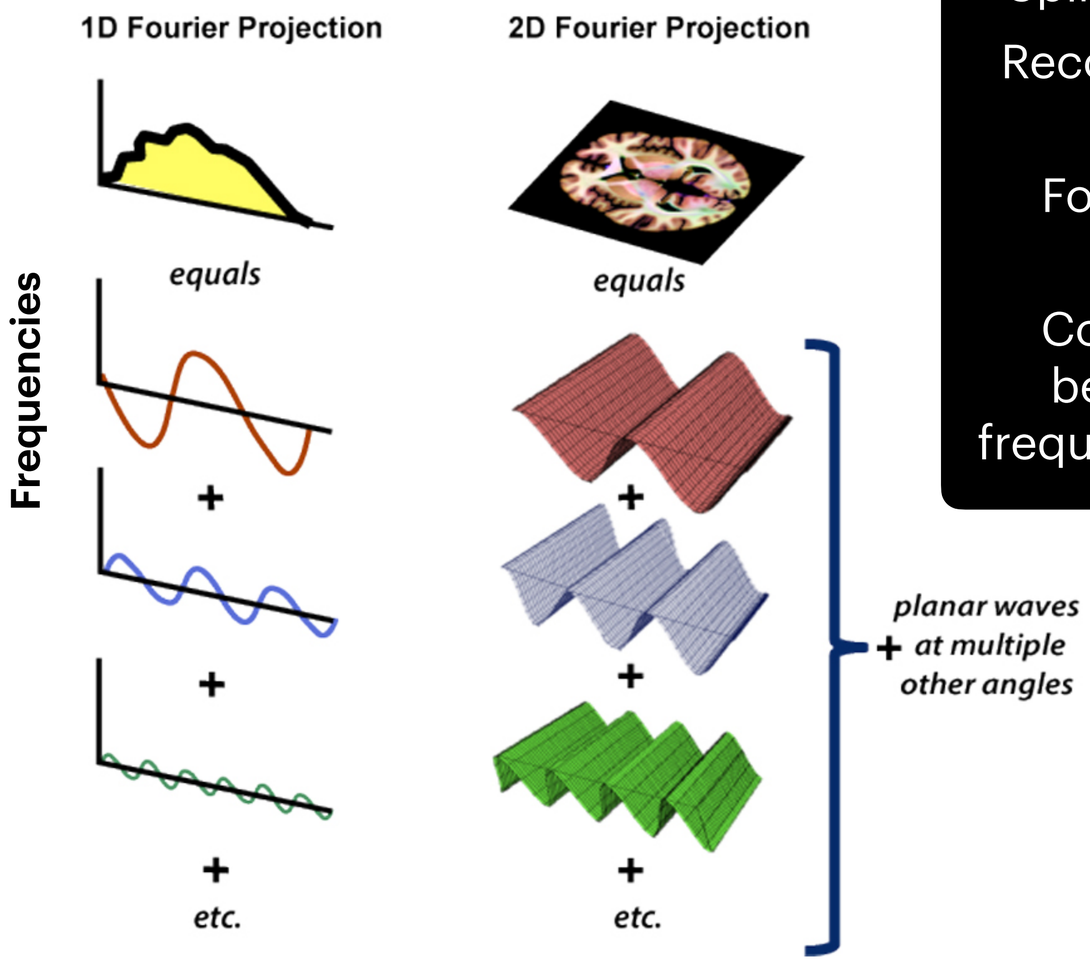


Predicted latent vs. ground truth reaction coordinate



Simulated

Aside: Structure validation + resolution determination with Fourier shell correlation

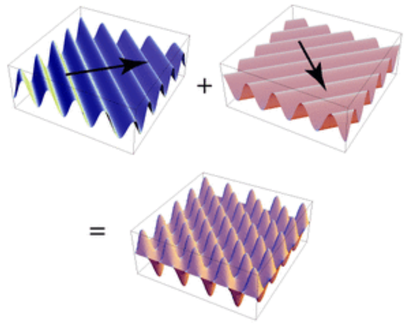


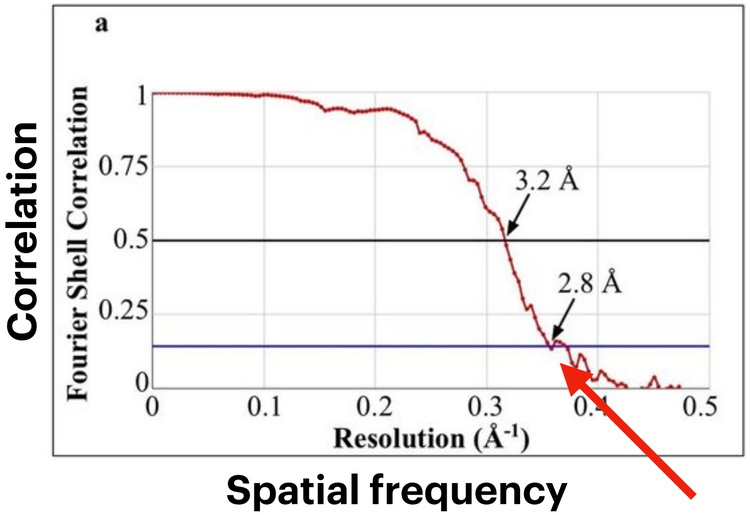
Split data in half (2 sets)

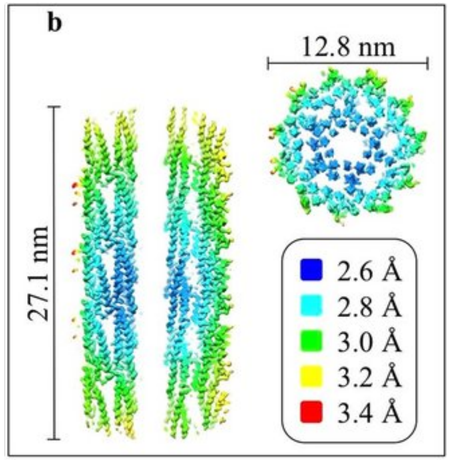
Reconstruct 3D volumes independently

Fourier-transform 3D volumes

Compute correlation between sets within frequency bands/3D shells







Local resolution estimation via masking

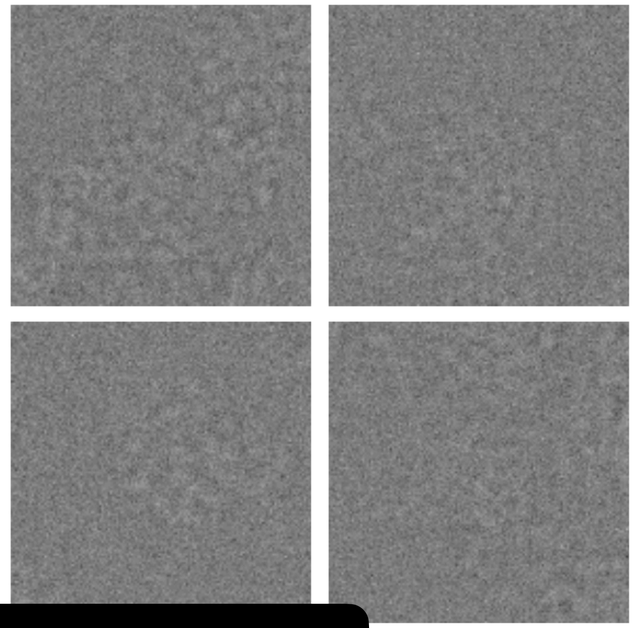
Learning ribosome assembly landscapes

Modular assembly of the bacterial large ribosomal subunit (LSU)

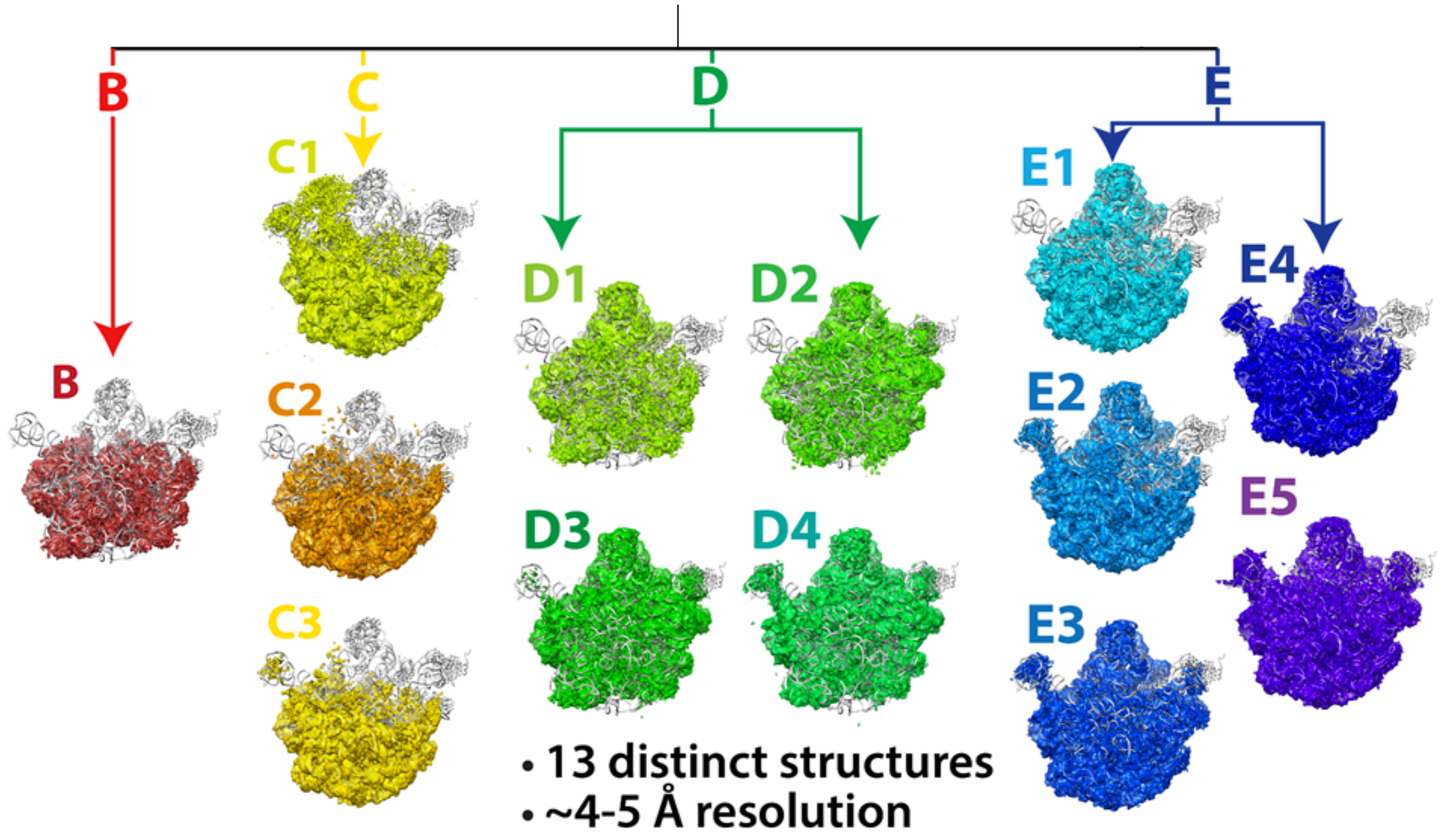
Dataset: 131k cryo-EM images of a mixture of LSU assembly intermediates

4 major and 13 minor states of the LSU identified from hierarchical multiclass reconstruction

Example images



Experimental [EMPIAR-10076]



Davis et al. Cell 2016

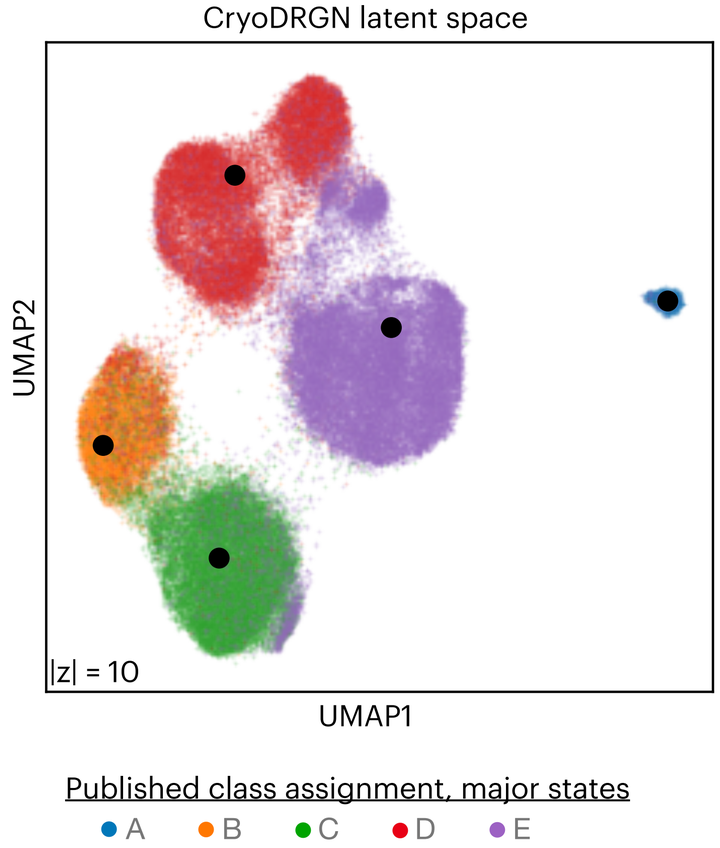
Learning ribosome assembly landscapes

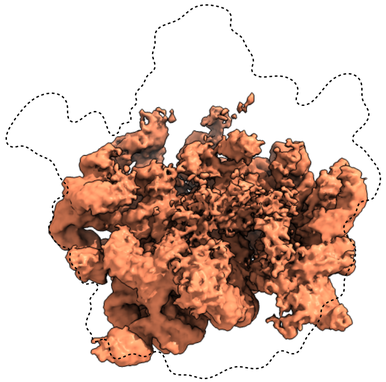
Modular assembly of the bacterial large ribosomal subunit (LSU)

Dataset: 131k cryo-EM images of a mixture of LSU assembly intermediates

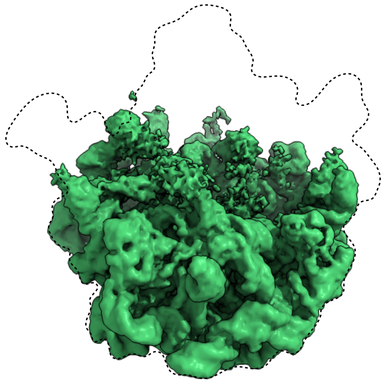
4 major and 13 minor states of the LSU identified from hierarchical multiclass reconstruction

Experimental

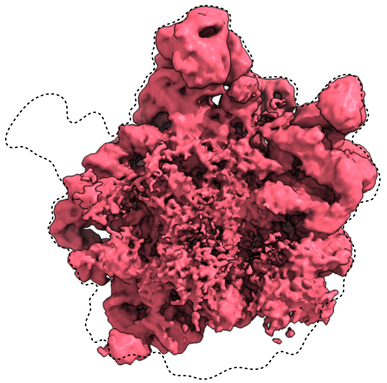




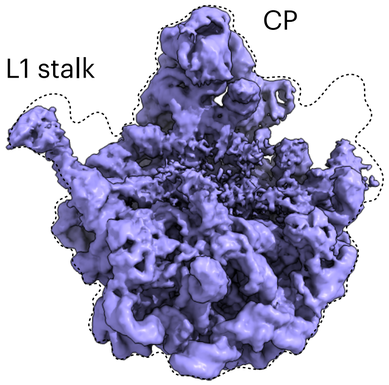
LSU assembly class B



LSU assembly class C

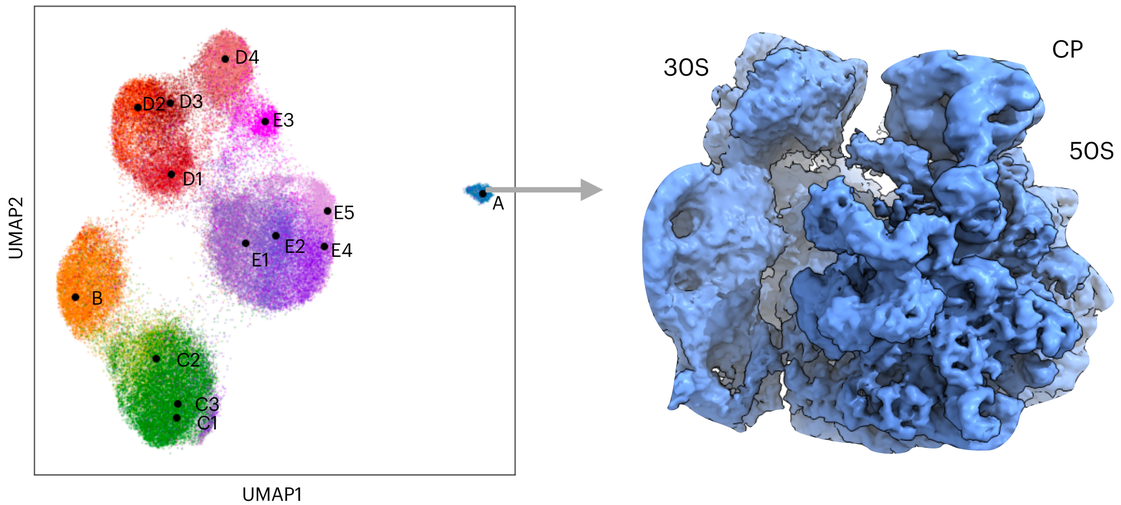


LSU assembly class D



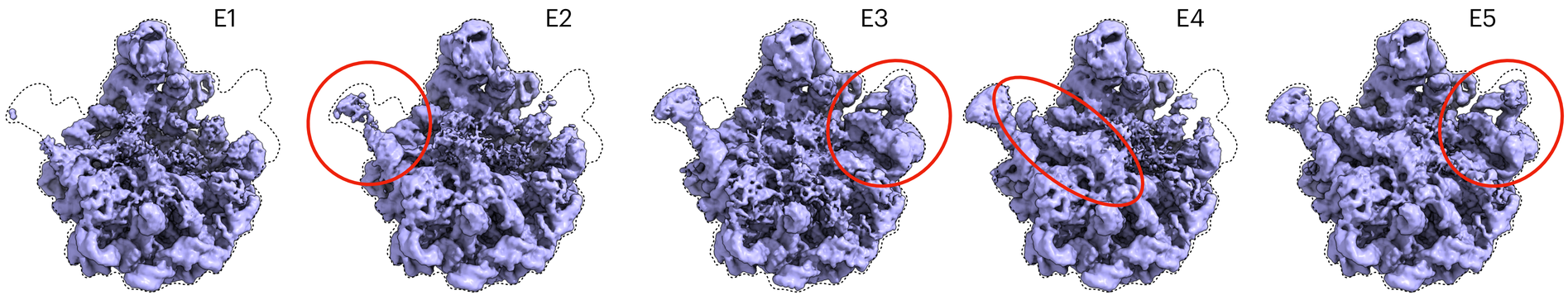
LSU assembly class E

Additional samples from the latent space



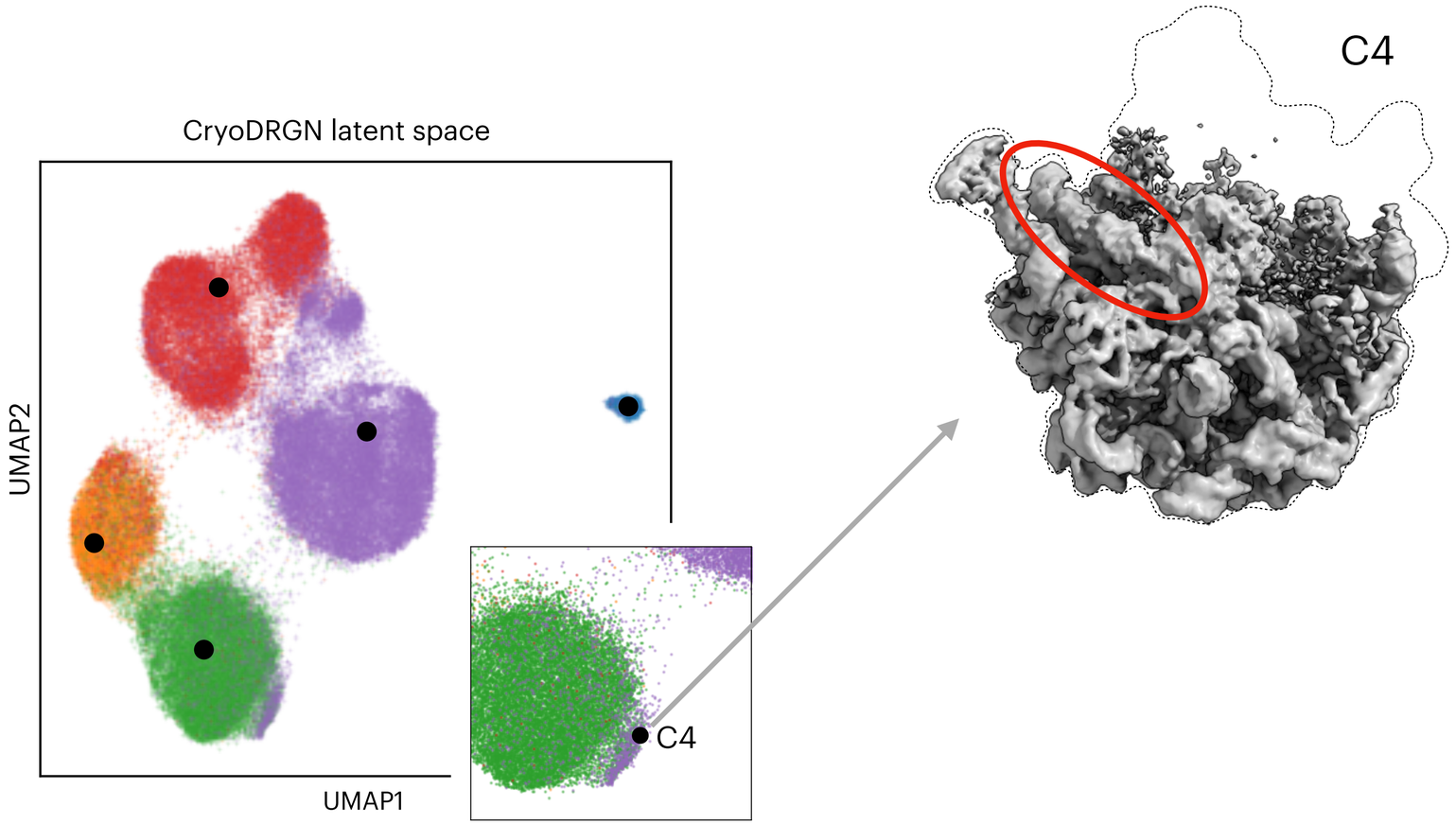
the 70S ribosome

E class minor assembly states



Experimental

Discovery of a new assembly state, C4

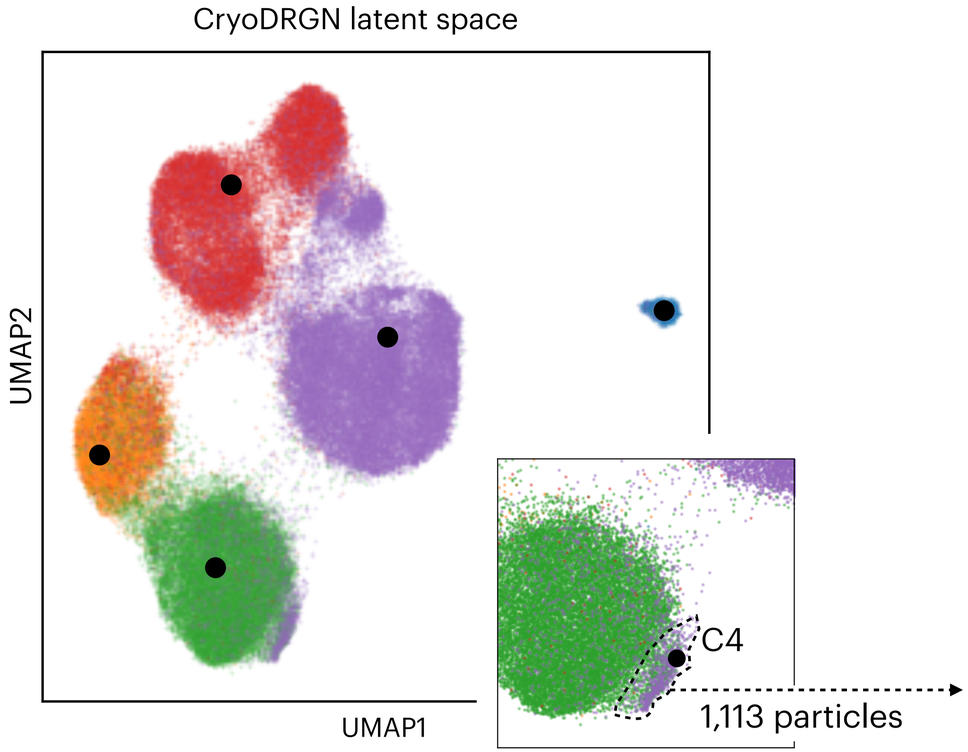


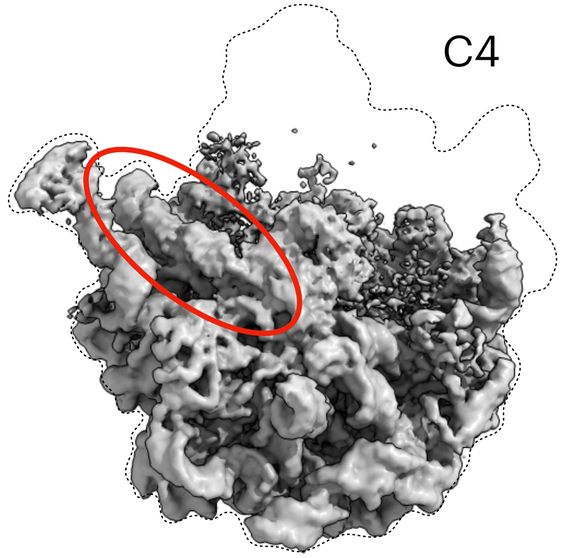
Published 3D classification, major states

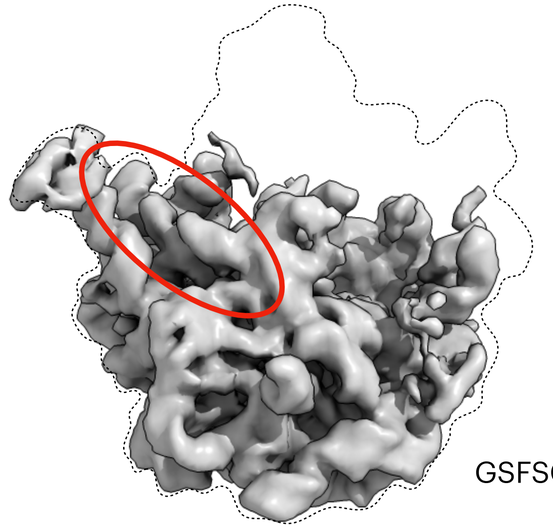
A B C D E

Experimental

Discovery of a new assembly state, C4







Homogeneous refinement

GSFSC 7.6 Å

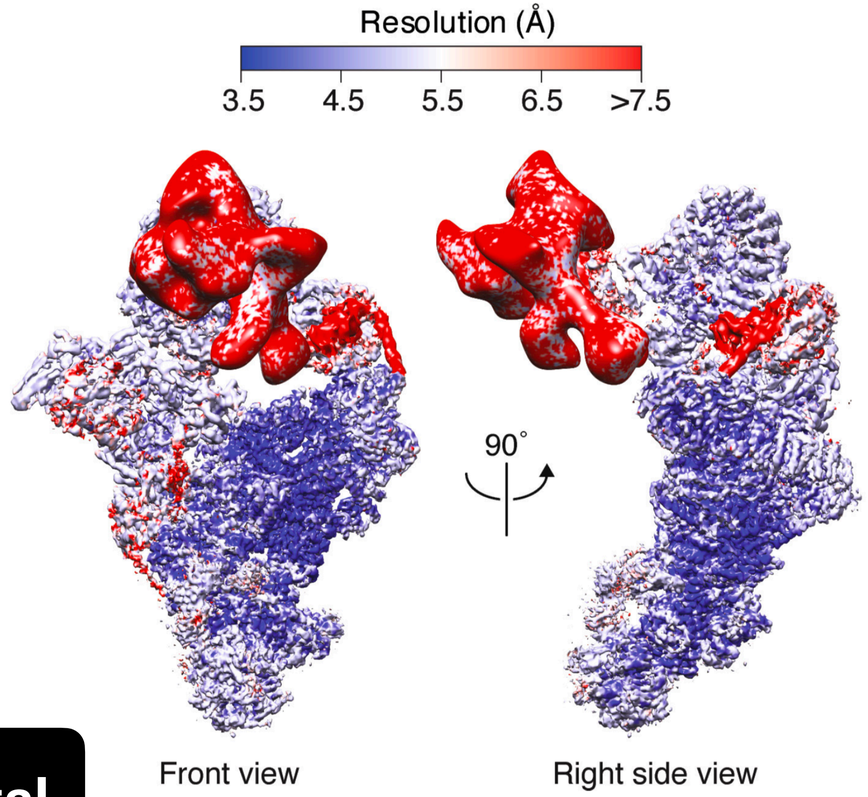
Published 3D classification, major states

A B C D E

Experimental

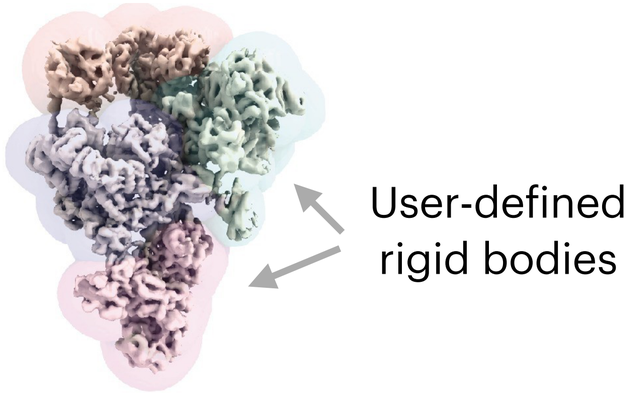
Structure of the pre-catalytic spliceosome

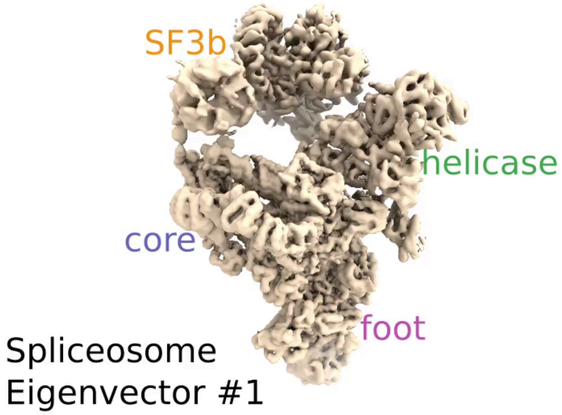
Sub-complexes resolved separately through many rounds of focused classification



Plaschka, Lin, & Nagai. Nature 2017

Multibody refinement



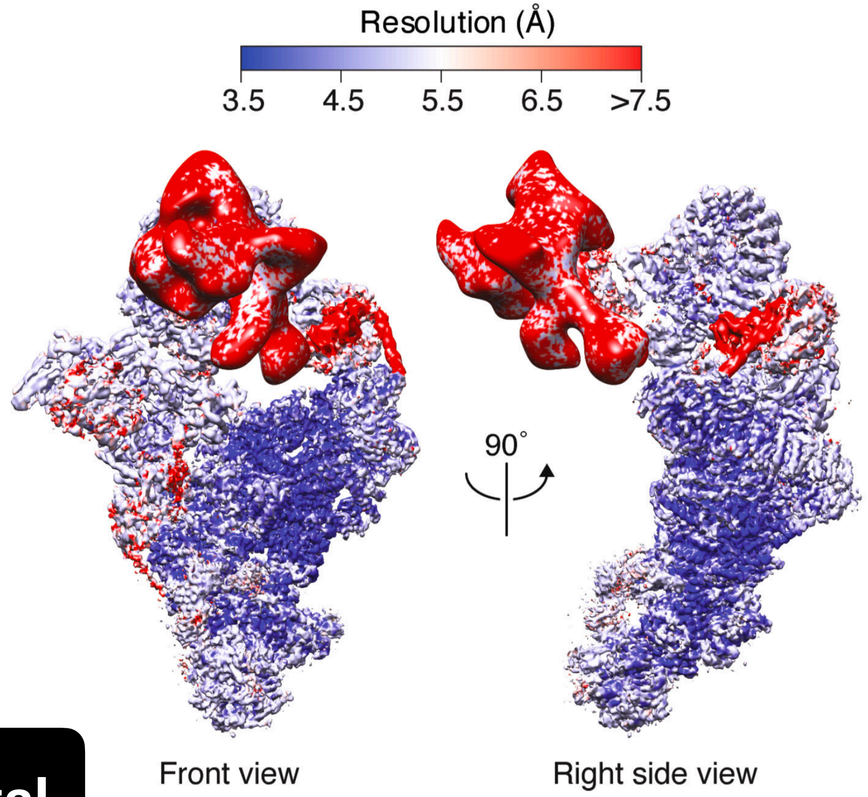


Nakane et al. eLife 2018

Experimental

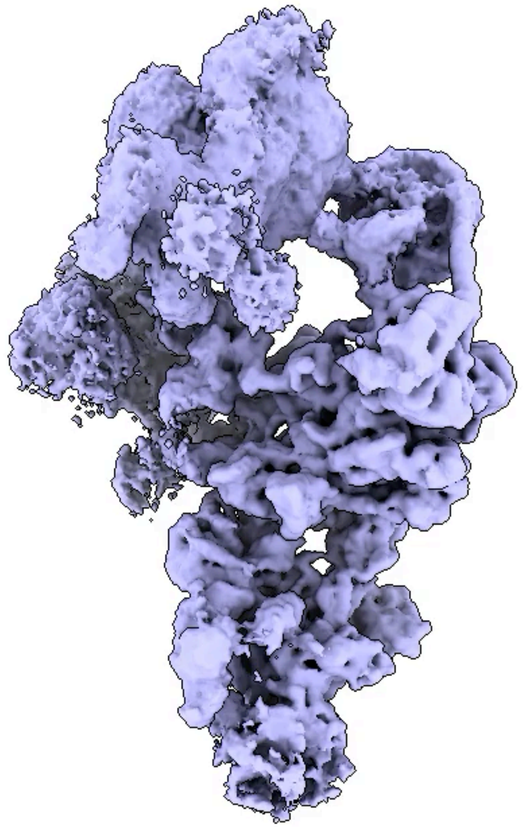
Structure of the pre-catalytic spliceosome

Sub-complexes resolved separately through many rounds of focused classification



Plaschka, Lin, & Nagai. Nature 2017

cryoSPARC 3DVA

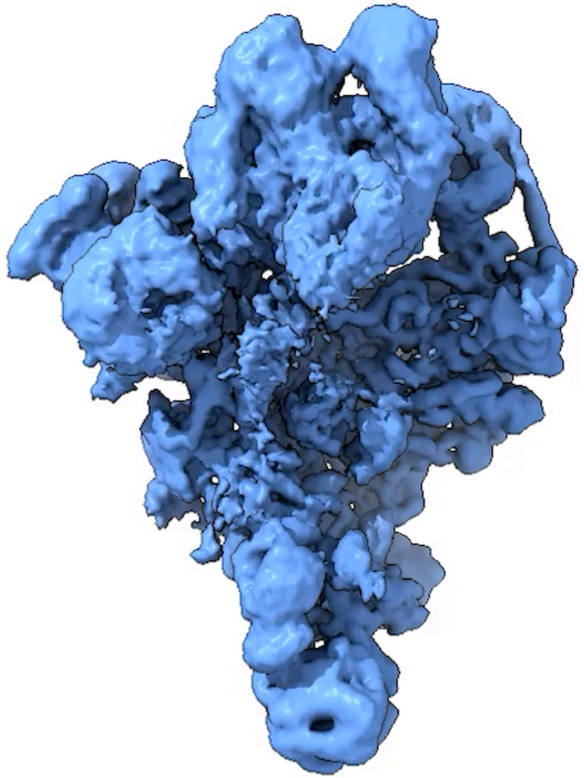


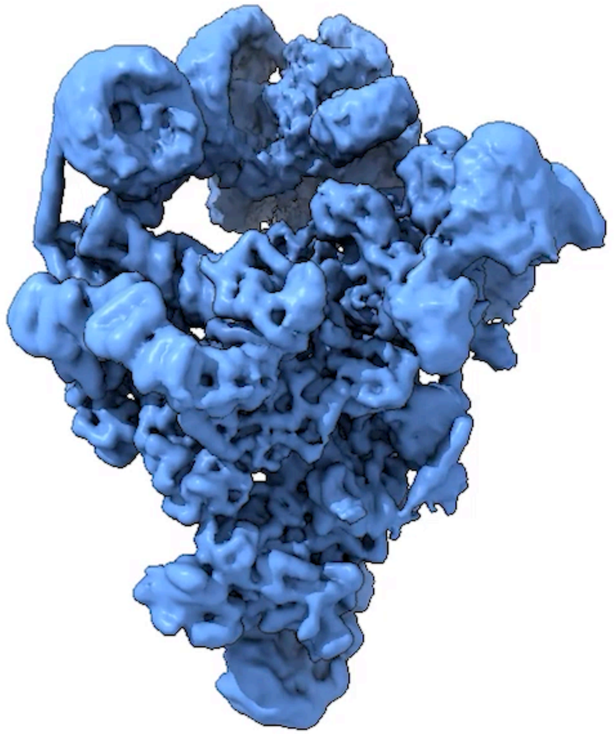
Punjani & Fleet. JSB 2021

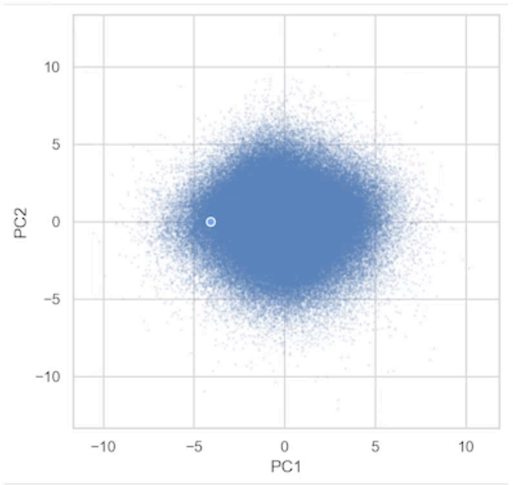
Experimental

Reconstructing continuous variability of the spliceosome

Trajectories along principle component axis of the latent space show variability within dataset







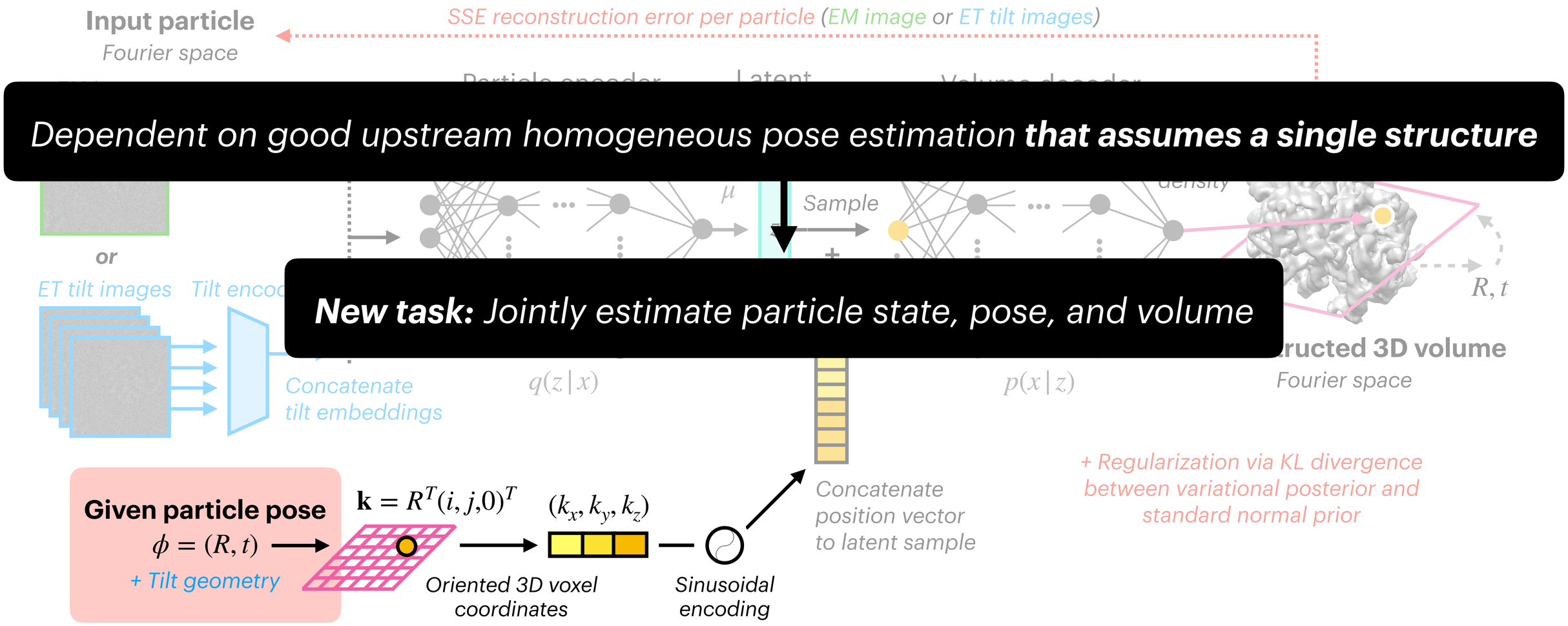
[EMPIAR-10180]

Experimental

Caveat: Interpolation along PCs can produce nonphysical motions e.g. under compositional heterogeneity and in general when the data distribution is not supported along the interpolation path

CryoDRGN: Variational autoencoder for heterogeneity analysis

Latent manifold encodes conformation of particle, which conditions 3D reconstruction



Zhong et al, *Nat Methods*, 2021; Rangan, Feathers et al, *Nat Methods*, 2024

1. CryoDRGN Results

▸2. Pose Estimation

3. CryoDRGN-AI Method
4. CryoDRGN-AI Results
5. Future Directions

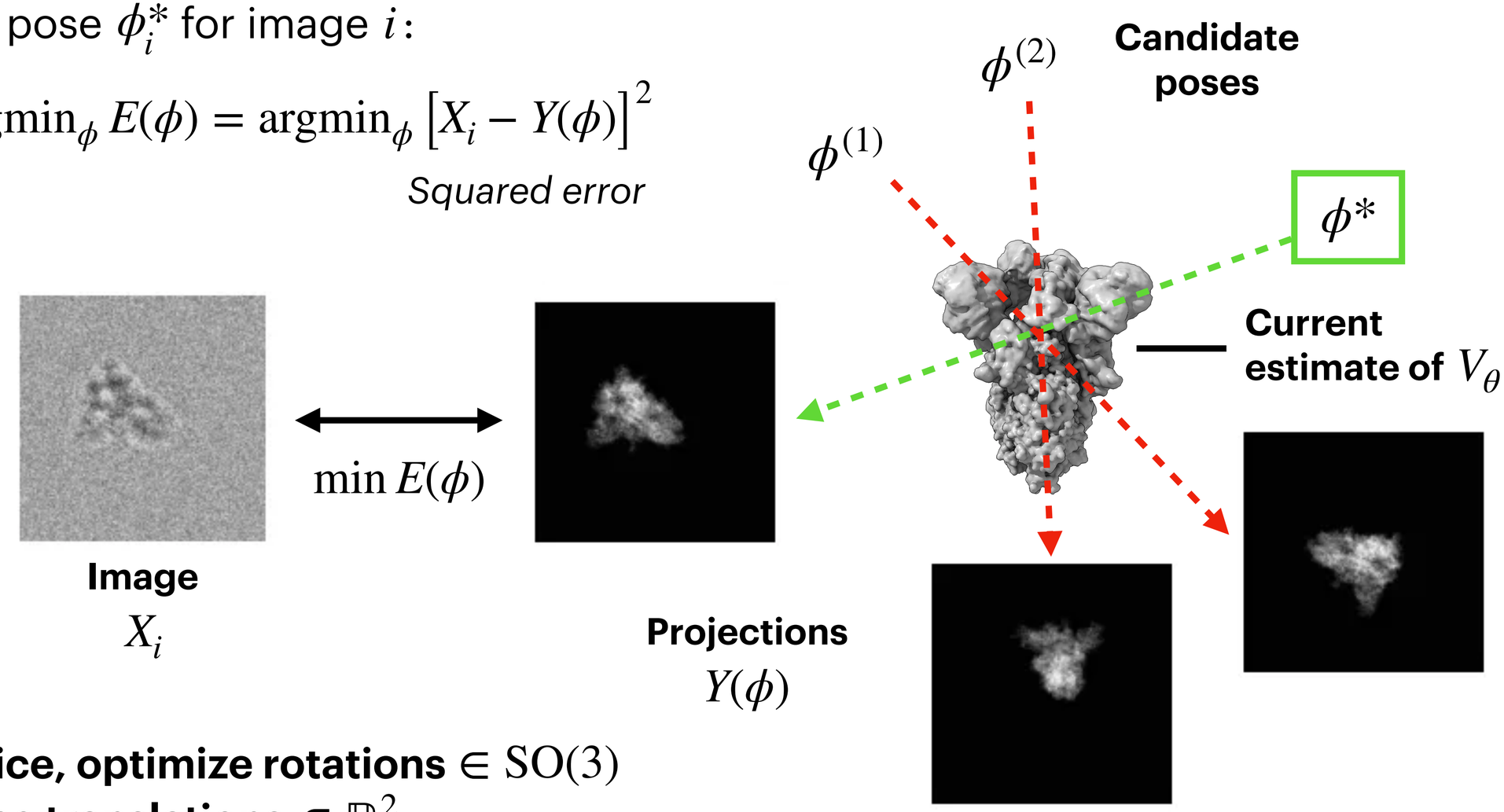
Roadmap

Image pose estimate evaluation via projection matching

Optimal pose ϕ_i^* for image i:

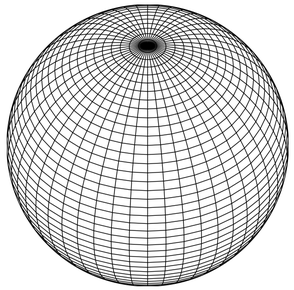
$$\phi_i^* = \text{argmin}_\phi \, E(\phi) = \text{argmin}_\phi \left[X_i - Y(\phi)\right]^2$$

Squared error



In practice, optimize rotations $\in \text{SO}(3)$
+ in-plane translations $\in \mathbb{R}^2$

Brute force alignment is intractable for cryo-EM

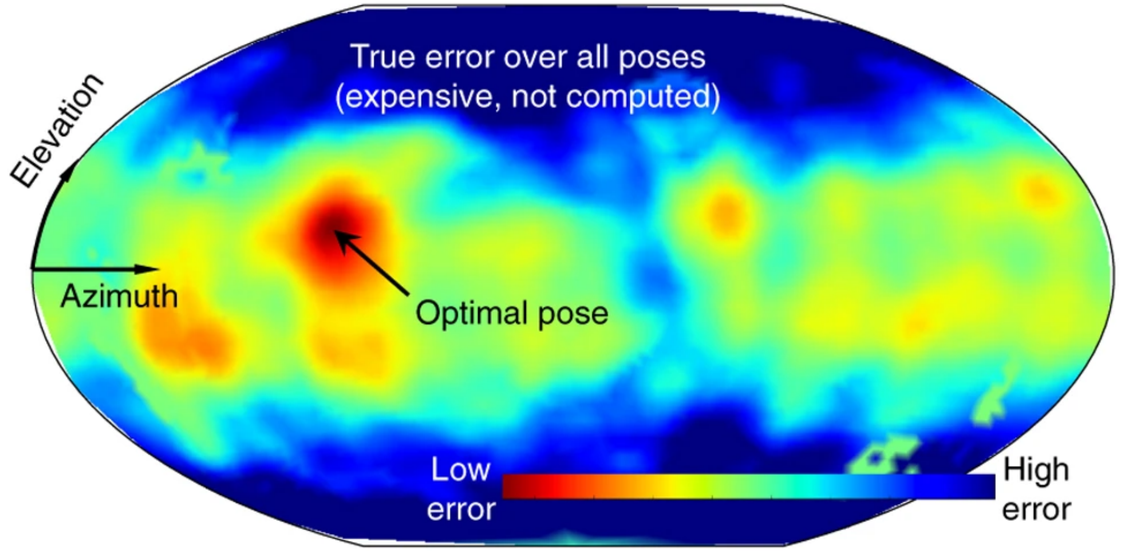


Evaluate error at all SO(3) candidates at desired resolution

Resolution	# candidate poses
15°	4,608
7.5°	36,864
3.75°	294,912
1.88°	2,359,296
0.94°	18,874,368

Product of S^1 grid of in-plane rotations and Healpix S^2 grid of out-of-plane rotations

Observation: Pose error is smooth



Idea: Cheaply reject large areas of pose space that certainly don't contain optimum

Punjani et al, *Nat Methods*, 2017

Single $D = 128$ pix image :

$128 \times 128 \times 18{,}874{,}368$ $\quad N = 100{,}000$ images?

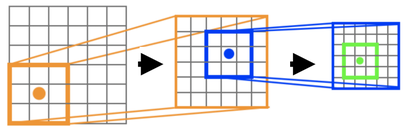
≈ 310 billion operations $\quad D = 256$ pix? Multiple epochs?

Górski et al, *Astrophysical Journal*, 2005

Pose estimation approaches for ab initio reconstruction

Exhaustive search

Rotation



Translation

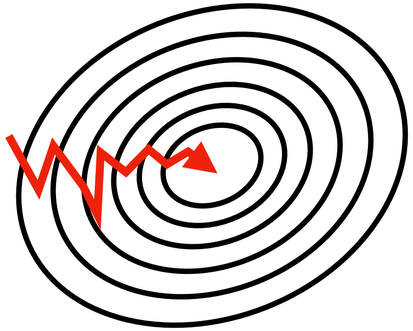
$$\phi_i^{(t+1)} = \text{argmin}_{\phi \in \text{grid}} \mathcal{L}(\phi)$$

Robust to high noise

Slow optimization

RELION, CryoSPARC, CryoDRGN2, ***CryoDRGN-AI****, etc.*

Differentiable point estimates



(Stochastic) Gradient descent

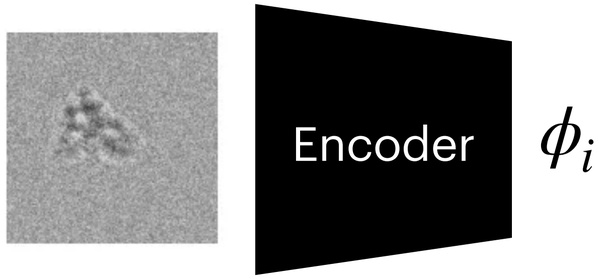
$$\phi_i^{(t+1)} = \phi_i^{(t)} - \lambda \nabla_\phi \mathcal{L}$$

Fast optimization

Needs initialization

CryoDRGN-AI

Amortized inference



Predict pose directly

$$\phi_i = f_\xi(Y_i)$$

Fast evaluation

Needs learning

CryoAI, CryoFIRE

1. CryoDRGN Results
2. Pose Estimation

▸ 3. CryoDRGN-AI Method

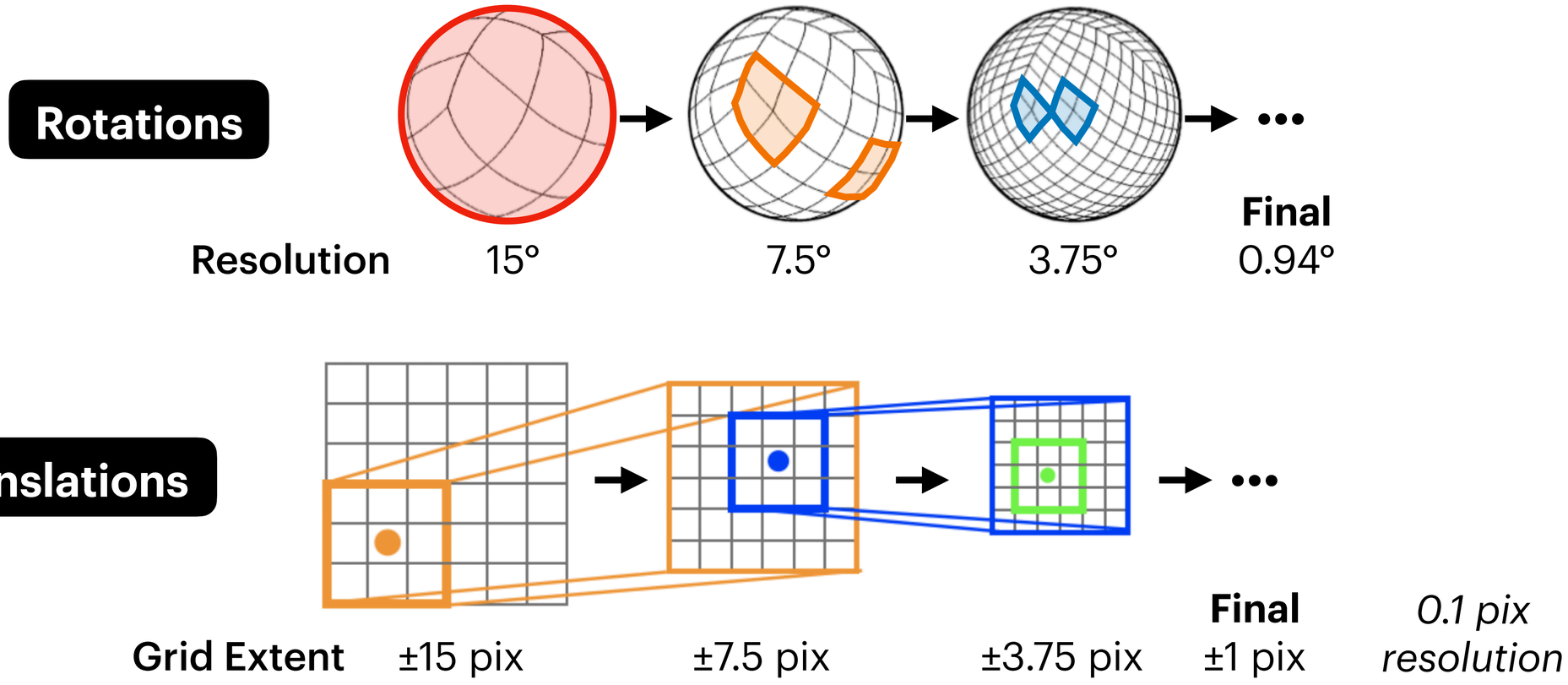
4. CryoDRGN-AI Results
5. Future Directions

Roadmap

CryoDRGN-AI: Coarse- to fine-grained hierarchical search grid

Subdivision

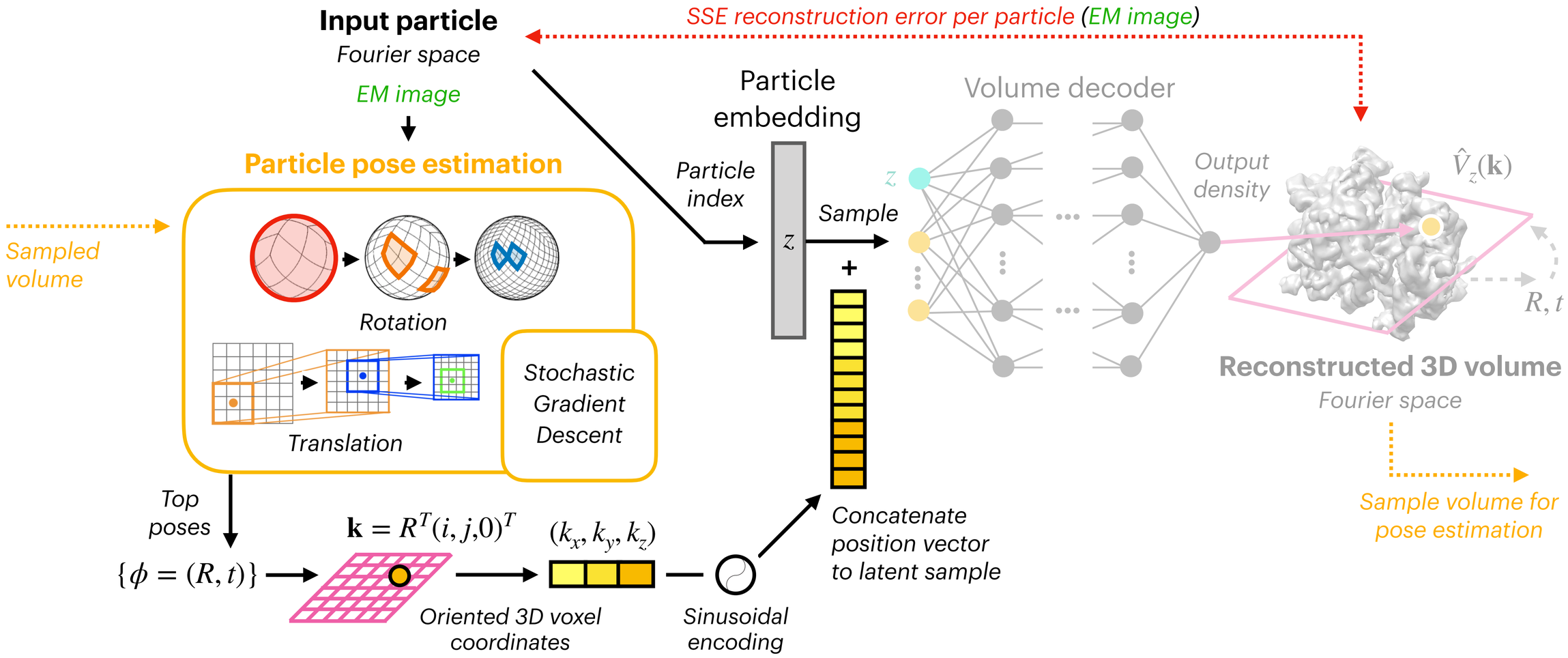
Sample poses coarsely to start, then more finely in promising areas of pose space



Levy et al, *Nat Methods*, 2025

CryoDRGN-AI: Ab initio heterogeneous reconstruction

Addition of hierarchical and stochastic gradient descent (SGD) pose estimation, no autoencoder

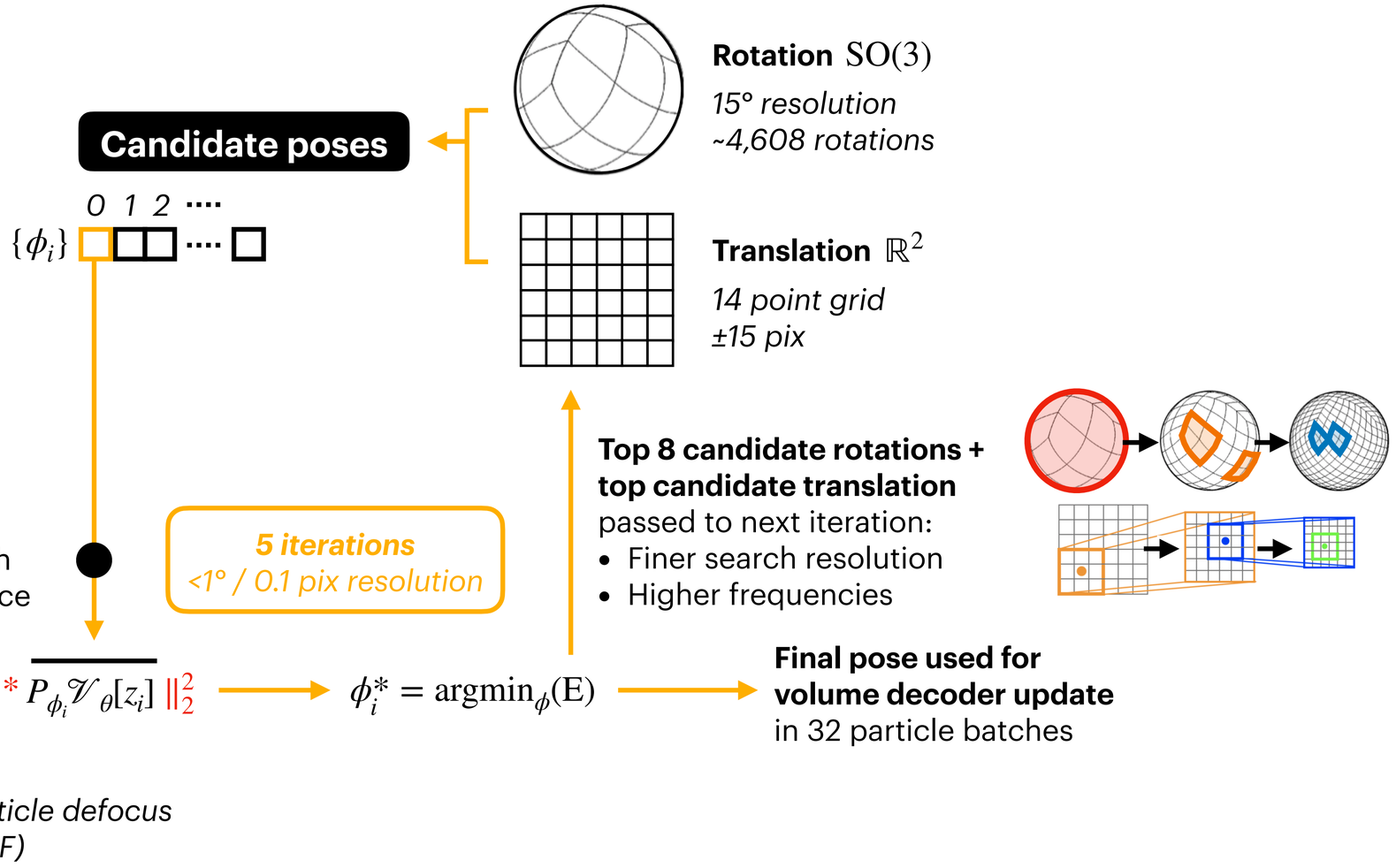


Levy et al, *Nat Methods*, 2025

Poses optimized per particle image during hierarchical search

Particle pose estimation

Hierarchical pose search for particle i :

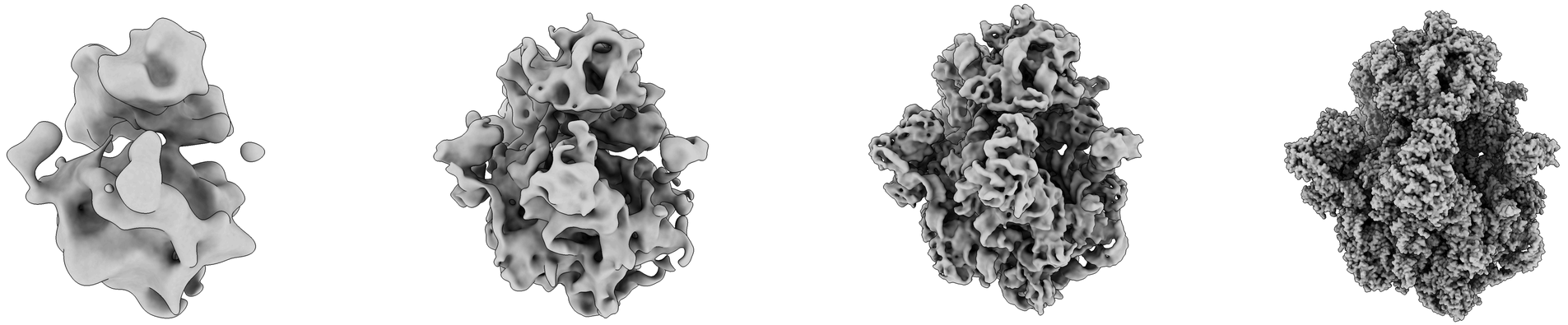


CryoDRGN-AI: Ramp spatial frequencies during search

Frequency marching

Frequency cutoff *L* is ramped up across pose search iterations to increasingly include higher-resolution information in alignment (Fourier space)

$$\hat{E}(\phi) = \sum_{||\ell|| \leq L} [X^{(\ell)} - Y(\phi)^{(\ell)}]^2$$



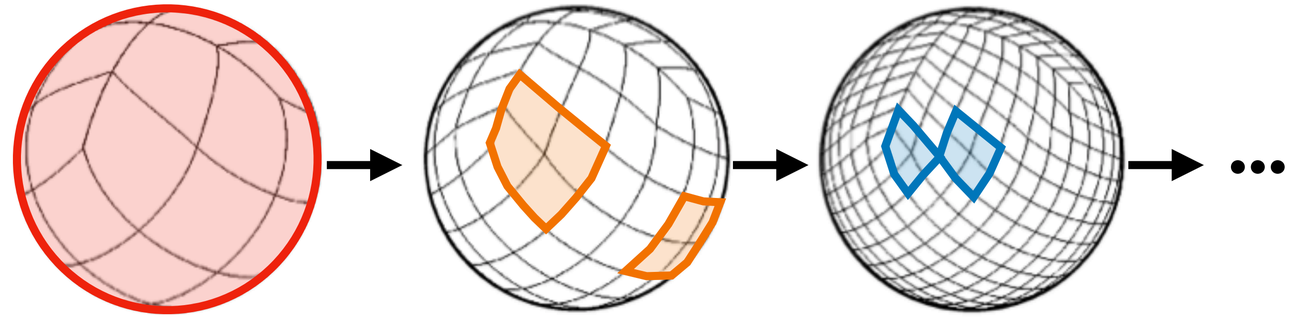
L

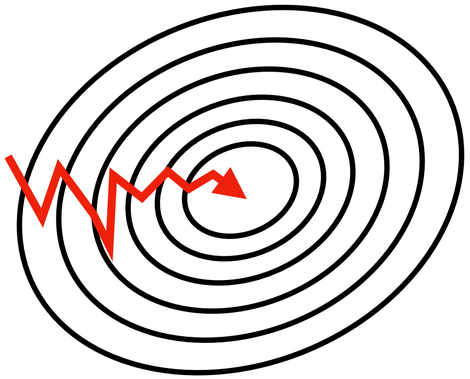
CryoSPARC: Double from $L_{min} = 12$ to L_{max} dependent on map resolution

CryoDRGN-AI: Linear from $L_{min} = 12$ to $L_{max} = 32$

Gradient descent following good pose initialization from hierarchical search

Particle pose estimation





Hierarchical Pose Search

150k or 2N particles processed

Robust, but slow and resolution-limited by search grid

Stochastic Gradient Descent

100 epochs

Fast and not resolution-limited, but requires good initialization

CryoDRGN-AI

Pose Search Hyperparameters

l_start: [int, default = 12] Number of frequencies used during the first pose search step.

l_end: [int, default = 32] Number of frequencies used during the last pose search step.

n_iter: [int, default = 4] Number of pose search iterations.

t_extent: [float, default = 20.0] Extent of the translation seach grid, in pixels.

t_n_grid: [int, default = 7] Number of point per dimension in the translation search grid.

t_x_shift: [float, default = 0.0] X-axis shift of the translation search grid.

t_y_shift: [float, default = 0.0] Y-axis shift of the translation search grid.

no_trans_search_at_pose_search: [bool, default = False] Flag for by-passing the translation search.

n_kept_poses: [int, default = 8] Number of poses kept per image.

base_healpy: [int, default = 2] Base healpy index.

1. CryoDRGN Results
2. Pose Estimation
3. CryoDRGN-AI Method

▸ 4. CryoDRGN-AI Results

5. Future Directions

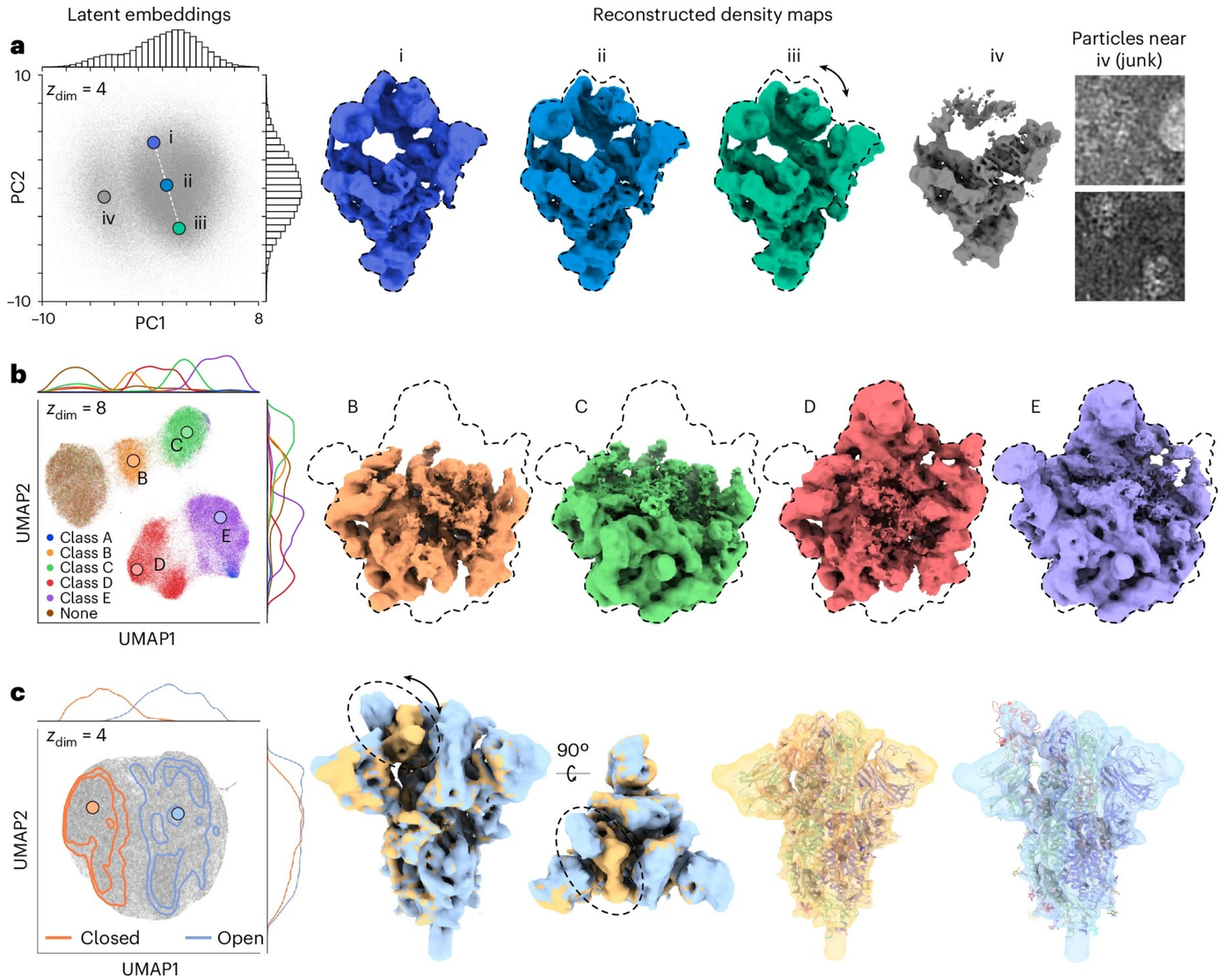
Roadmap

Single-shot reconstruction from unfiltered datasets

Pre-catalytic spliceosome
Extended vs. retracted bending motion

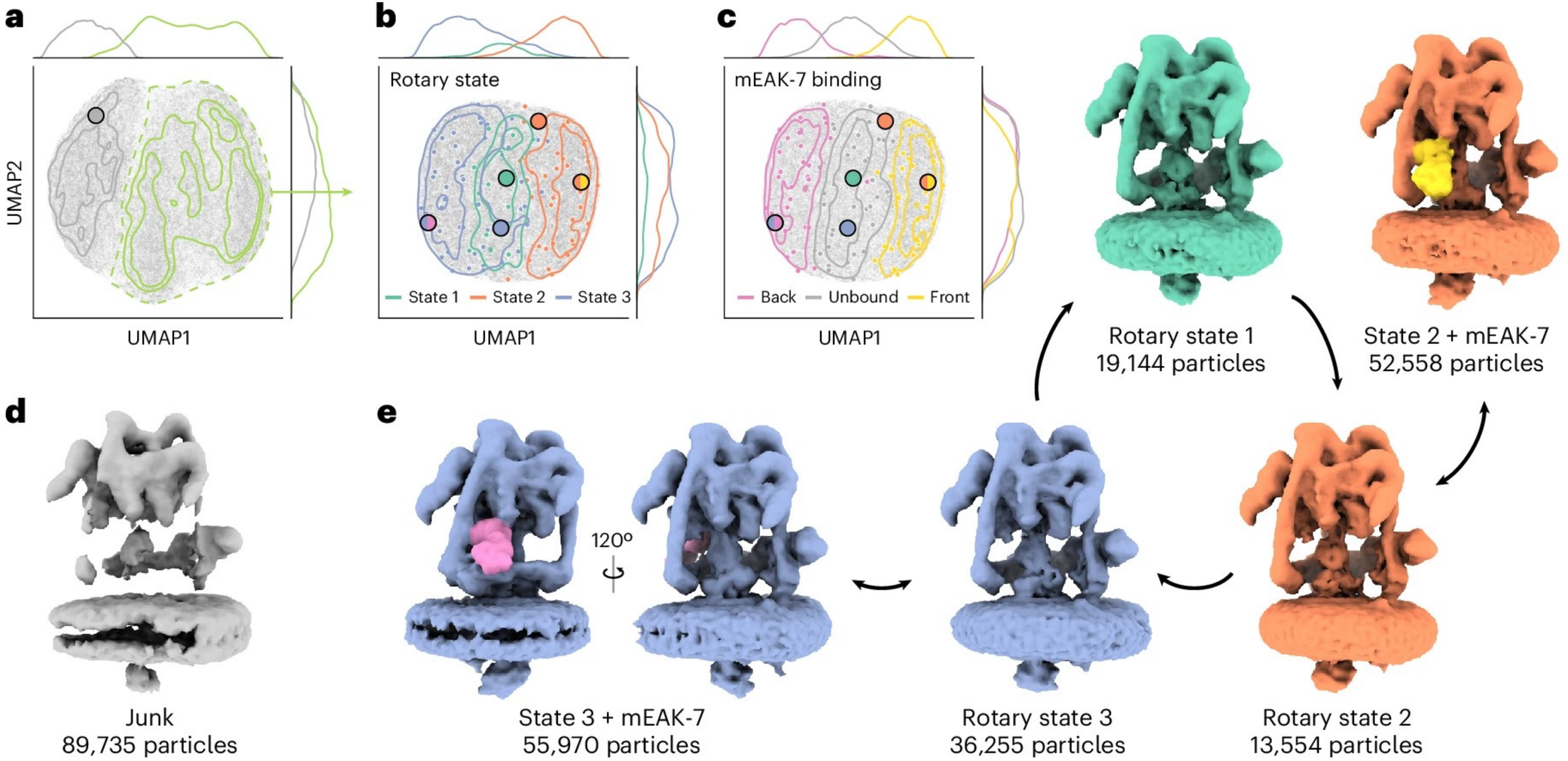
Bacterial large ribosomal subunit
Assembly states

SARS-CoV-2 spike protein
Conformational change of receptor-binding domain



V-ATPase complex (transmembrane proton pump)

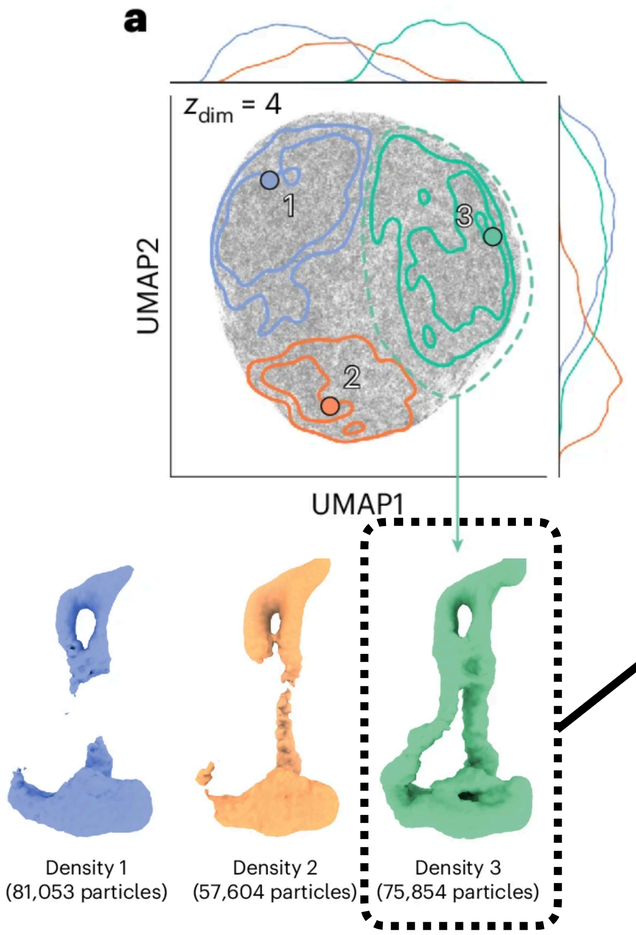
Junk identification

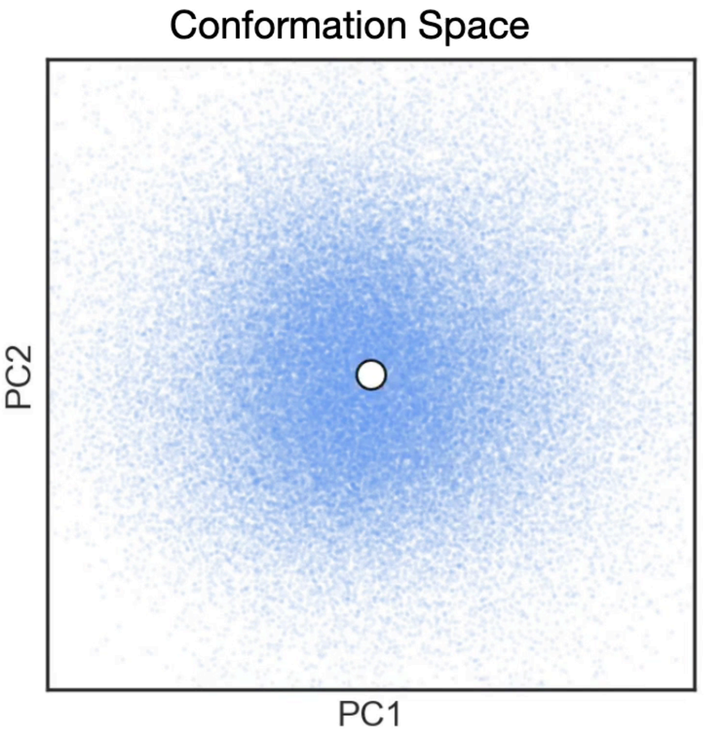


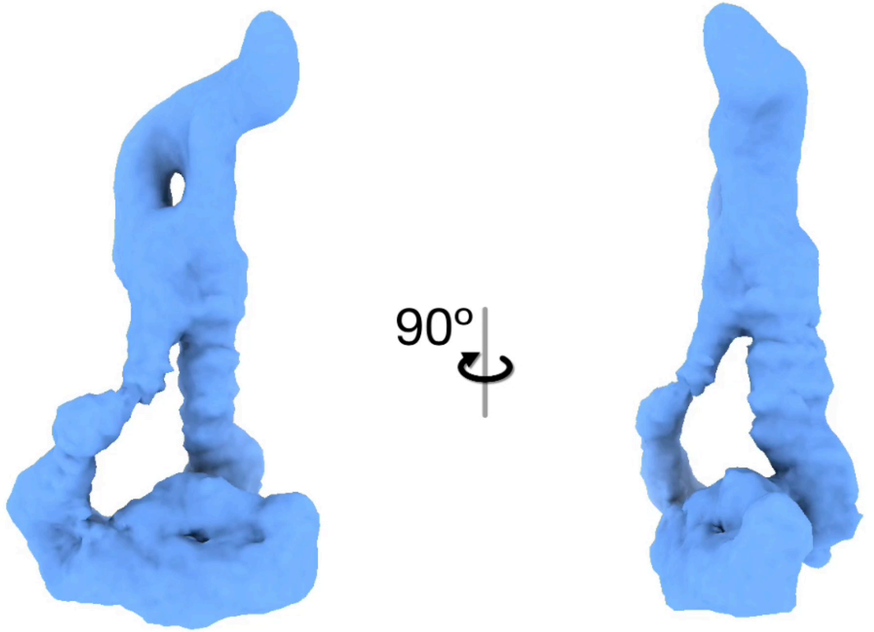
Rotary states

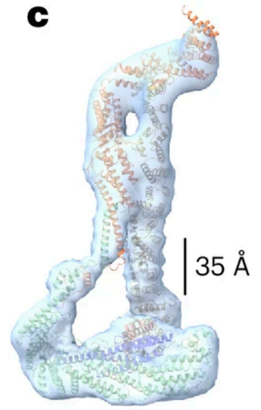
Reconstructing the flexible DSL1/SNARE complex

A highly flexible complex is reconstructed despite the presence of junk particles







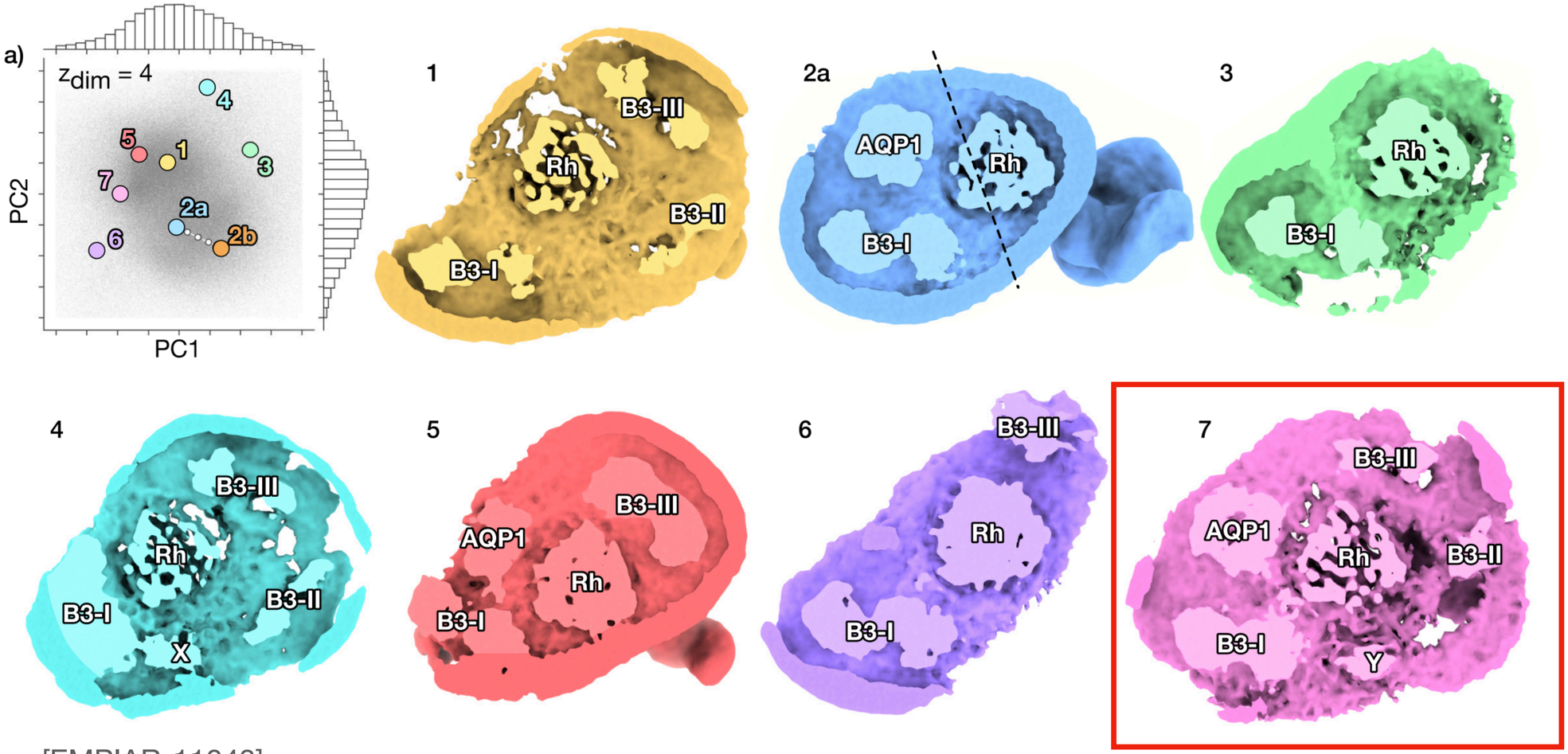


[EMPIAR-11846]

Levy *et al*. bioRxiv 2024. Nature Methods 2025

Revealing new states of the ankyrin-1 complex

Ab initio heterogeneity analysis reveals a rare “supercomplex” state



[EMPIAR-11043]

Levy and Raghu *et al*. Nature Methods 2025

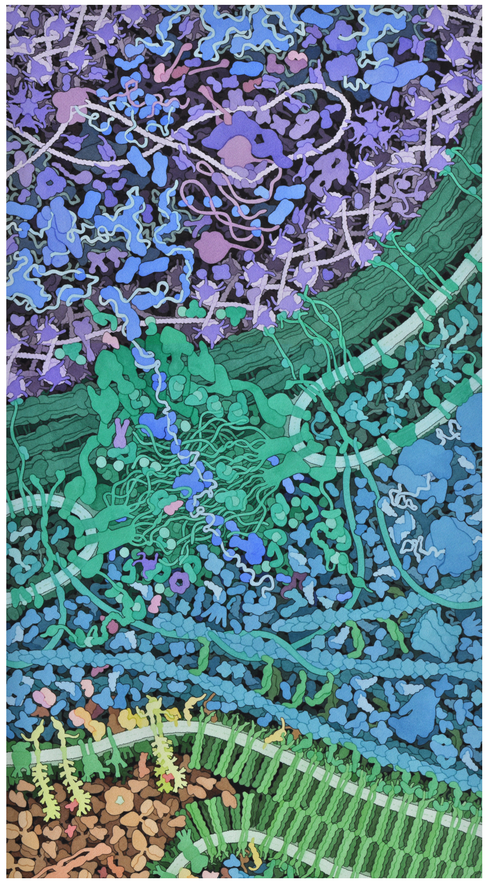
1. CryoDRGN Results
2. Pose Estimation
3. CryoDRGN-AI Method
4. CryoDRGN-AI Results

▸ 5. Future Directions

Roadmap

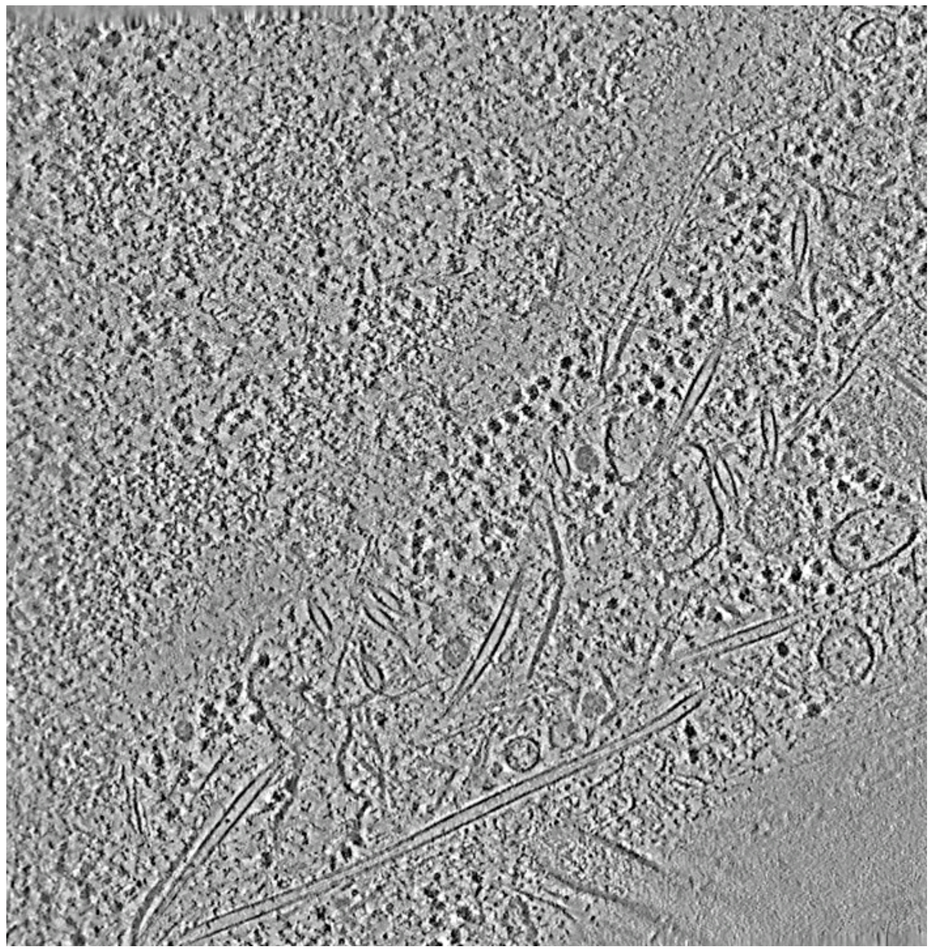
Direct cellular visualization with in situ cryo-electron tomography

Cellular biology with spatial context



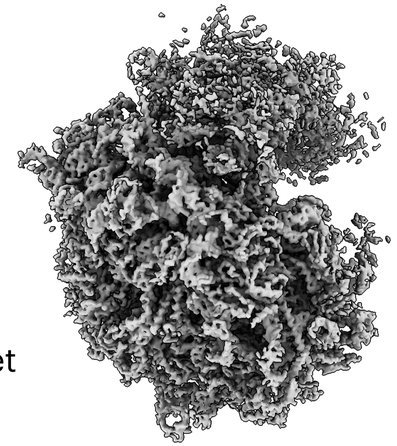
VEGF signaling, Goodsell, 2011

HeLa cell nuclear periphery,
Mahamid et al, *Science*, 2016



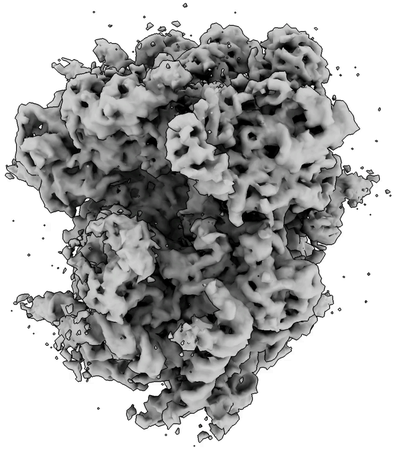
High resolution detail

80S ribosome,
Phantom Dataset



Structure dynamics

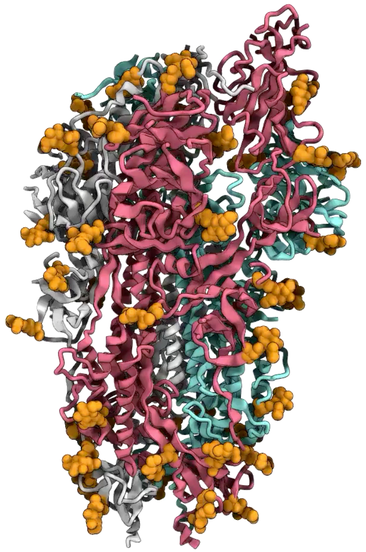
CryoDRGN-ET on
Phantom Dataset,
Ryan Feathers



Dynamic and diverse biomolecular systems exhibit extreme structural *heterogeneity*

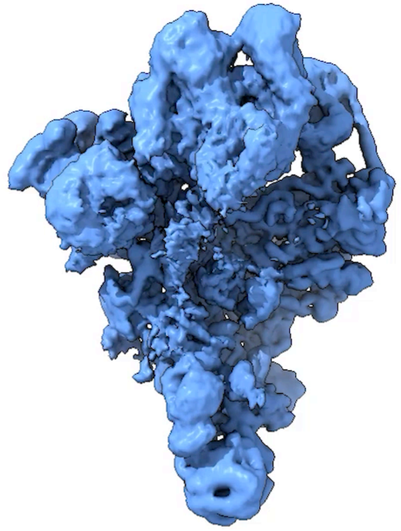
Conformational Variability

SARS-CoV-2 spike
MD simulation



D.E. Shaw Research

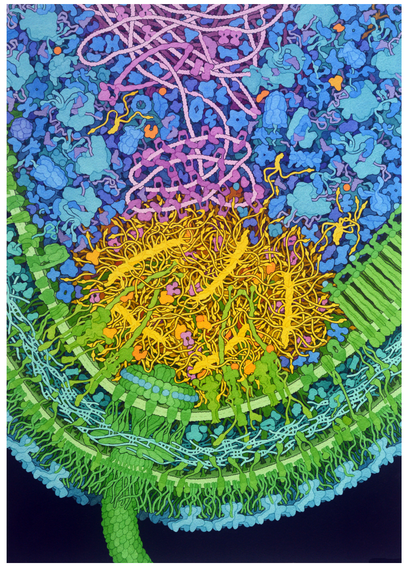
Spliceosome
Cryo-EM



Data: Plaschka et al, *Nature*, 2017
Model: Zhong et al, *Nat Methods*, 2021

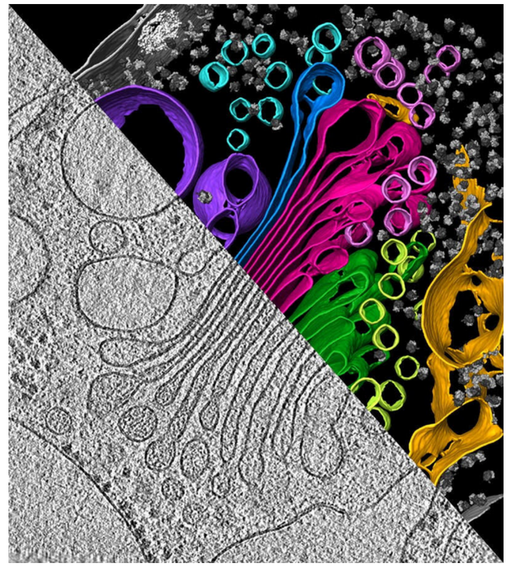
Compositional Variability

Flagellar motor
Illustration



Goodsell, 2022

Algae vesicles and golgi apparatus, *Cryo-ET*



Bykov et al, *eLife*, 2017

Directions for algorithms for cryo-EM structure determination

Open problems

Cryo-EM/ET

- Higher resolution
- Smaller targets
- Throughput
- Complex heterogeneity
  - Highly flexible structures
  - Cellular imaging with complex mixtures

Downstream tasks

- Spatial analysis
- Model building
- Conformational landscapes

Areas of exploration

Cryo-EM/ET

- More expressive and efficient neural models
- Agentic processing workflows
- Structural priors for reconstruction

Downstream tasks

- Spatial analysis adapted from spatial transcriptomics
- Structural priors for model building
- Statistical mechanics