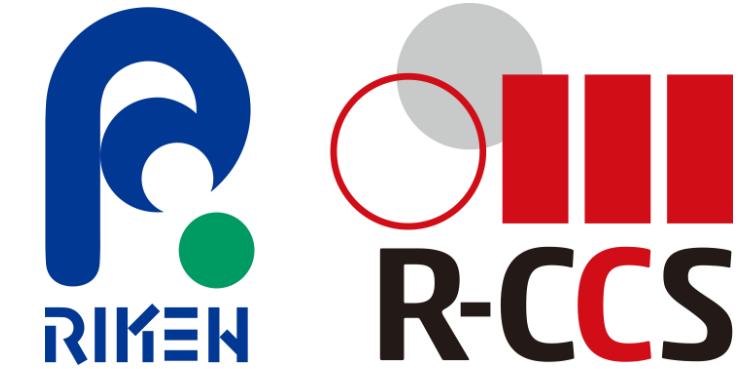


Scalable HPC Workflow on “Fugaku” for Uncovering Multiple Protein Conformational States from Cryo-EM Single-Particle Images



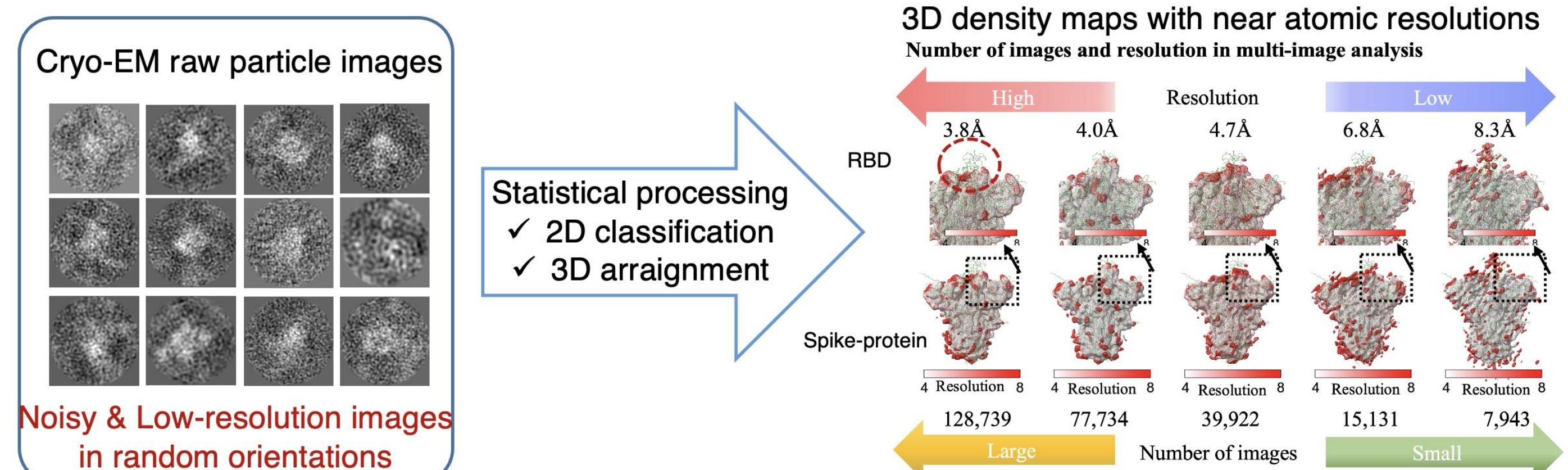
Atsushi Tokuhisa

RIKEN Center for Computational Science, Japan

BACKGROUND / MOTIVATION

- Transient protein conformations are essential for biological function but are often short-lived and heterogeneous.
- In cryo-electron microscopy (cryo-EM), ensemble averaging and image noise obscure rare and intermediate states.
- Bridging experimental cryo-EM images with atomistic molecular simulations is therefore required to resolve such hidden conformations.
- We introduce **cryoMDM**, a molecular-simulation–driven structural matching framework that directly estimates 3D atomic models from individual cryo-EM images, bypassing density-map reconstruction.
- To enable large-scale conformational searching, we implemented a **scalable HPC workflow** on the Fugaku supercomputer.

Advances and Challenges in cryo-EM

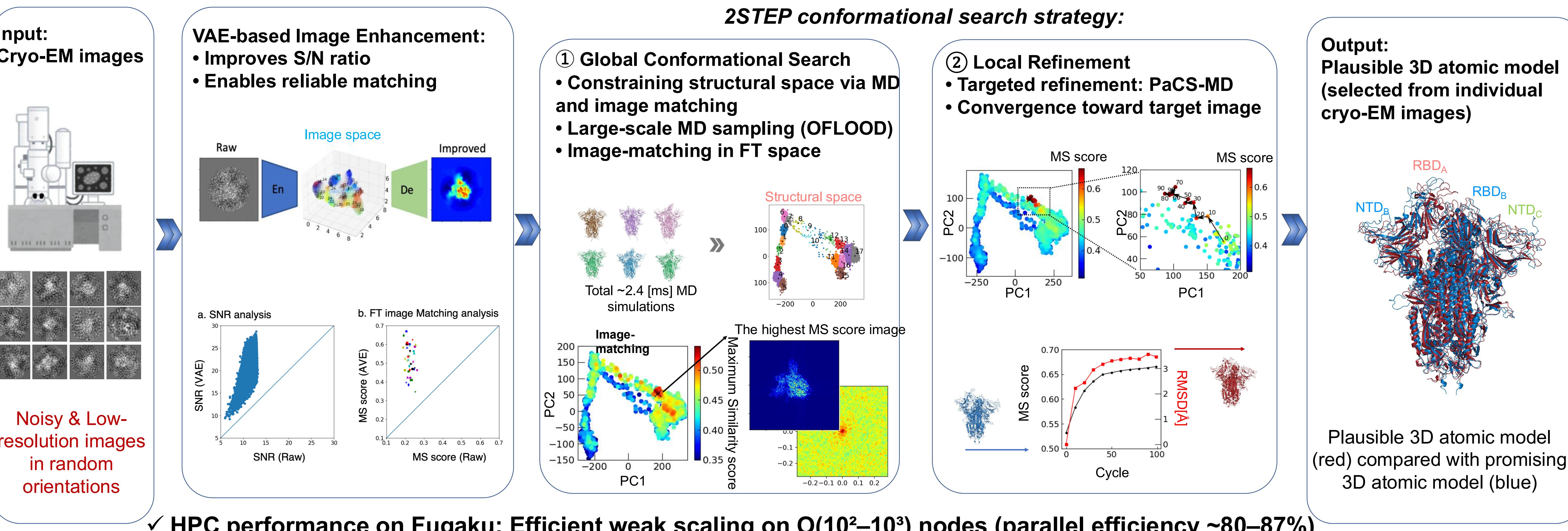


✓ Statistical processing improve signal, but at same time it blurs flexible regions, and then masks rare conformations.

METHODS

- CryoMDM matches MD-generated atomic conformations directly with individual cryo-EM images.
- The workflow consists of VAE-based image enhancement, large-scale global search with OFLOOD-MD, and local refinement using PaCS-MD. Post-analysis is performed using Fragment Molecular Orbital (FMO) calculations.
- The workflow is optimized for the Fugaku supercomputer and scales efficiently to several hundred nodes.
- This enables high-throughput structural matching without full-system execution.

Fig.1 | Scalable cryoMDM Workflow on Fugaku (with VAE-based image enhancement)

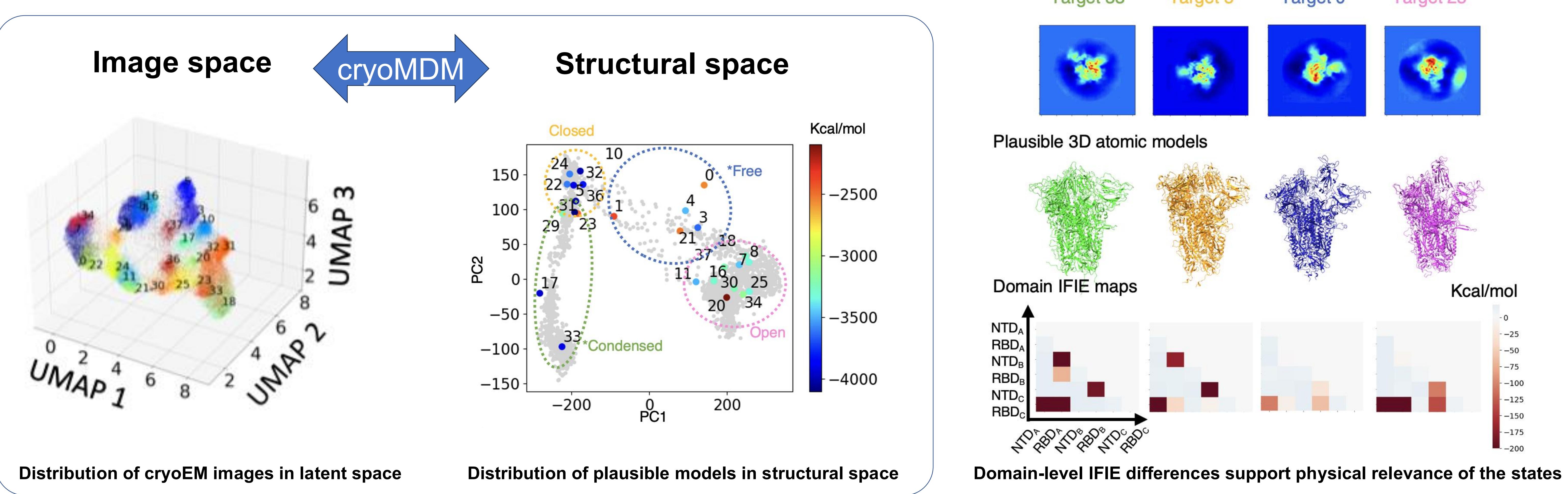


✓ HPC performance on Fugaku: Efficient weak scaling on $O(10^2\text{--}10^3)$ nodes (parallel efficiency ~80–87%)

RESULTS

- CryoMDM identified multiple plausible 3D atomic models of the SARS-CoV-2 spike protein directly from individual cryo-EM images.
- Analysis of 26 target images revealed four characteristic conformational states: Condensed, Closed, Free, and Open.
- The Condensed and Free states represent previously uncharacterized intermediate conformations.
- FMO analysis confirmed distinct interaction patterns stabilizing each state.
- These results demonstrate that large-scale searching on Fugaku enables direct identification of transient protein conformations.

Fig.2 | Four characteristic conformational states revealed by cryoMDM



CONCLUSIONS / TAKE-HOME MESSAGE

- Hundreds-of-node HPC workflows on Fugaku enable direct, atomic-resolution discovery of transient protein conformations from noisy cryo-EM images.