

# Continuous heterogeneity analysis of CryoEM images through Zernike Polynomials and Spherical Harmonics

D. Herreros, R.R Lederman, J. Krieger, A. Jiménez-Moreno, M. Martínez, D. Myska, D. Strelak, J. Filipovic, I. Bahar, J.M. Carazo, C.O.S. Sorzano

## Introduction

Molecular flexibility has raised great interest amongst the CryoEM community, as it supposes a step forward in the study of molecular structure and function. Typically, 3D classification has been preferred to study molecular flexibility, but it has proven insufficient when the goal is to describe a molecule's conformational landscape. Luckily, it is possible to go a step before and study heterogeneity at the level of particles, where a larger number of states can be identified.

Following the aforementioned idea, we introduce a new formulation based on a mathematical basis to model continuous flexibility at the level of CryoEM particles. The information computed by this new method can be represented in a low-dimensional mapping that can be used to analyze and retrieve the conformational states represented by the different images.

## Conclusions

- Zernike3D basis allows studying molecular motions effectively thanks to the decomposition on varying frequency movements
- The application of this new tool to particle images results in a high-dimensional coefficient space where conformation and continuous motions can be identified
- Conformations can be retrieved at the level of particle refinement (high number of images) or volume deformation (low number of images)

## Methods

- Displacement finding problem:

$$\min_{g_L} \int |I(r) - P(V(r + g_L(r)))| dr + \lambda_1 \int \|g_L(r)\|^2 + \lambda_2 \frac{|\int (V_1(r) - V_1(r + g_L(r))) dr|}{\int V_1(r) dr}$$

- Definition of  $g_L$  based on the new basis:

$$g_L(r) = \sum_{l=0}^L \sum_{n=0}^N \sum_{m=-l}^l \begin{pmatrix} \alpha_{l,n,m}^x \\ \alpha_{l,n,m}^y \\ \alpha_{l,n,m}^z \end{pmatrix} Z_{l,n,m}(r)$$

- Definition of the basis

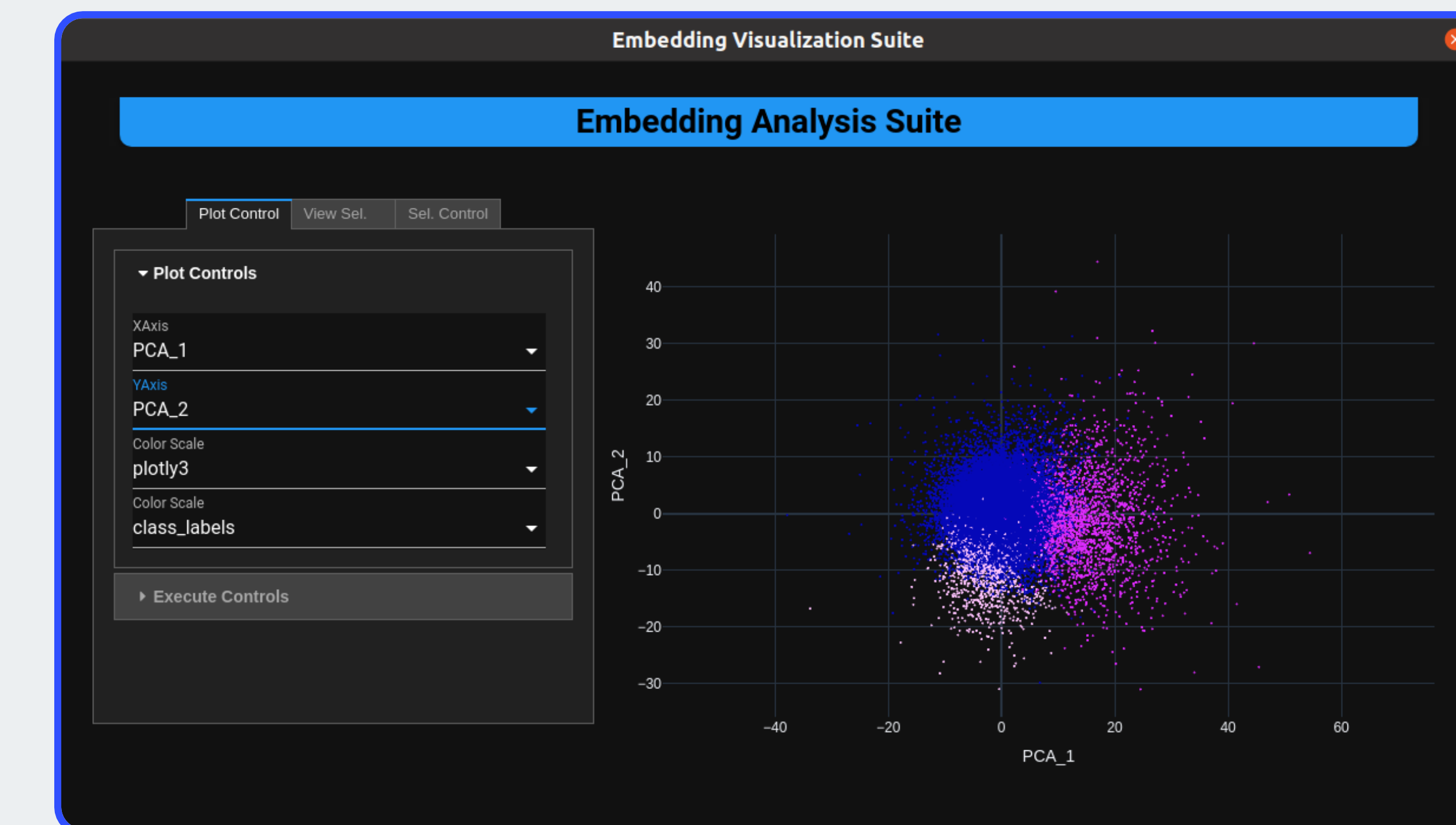
$$Z_{l,n,m}(r) = \bar{R}_{l,n}^1(r) y_l^m(\theta, \phi)$$

Being  $\bar{R}_{l,n}^1(r)$  the Zernike Polynomials and  $y_l^m(\theta, \phi)$  the real Spherical Harmonics

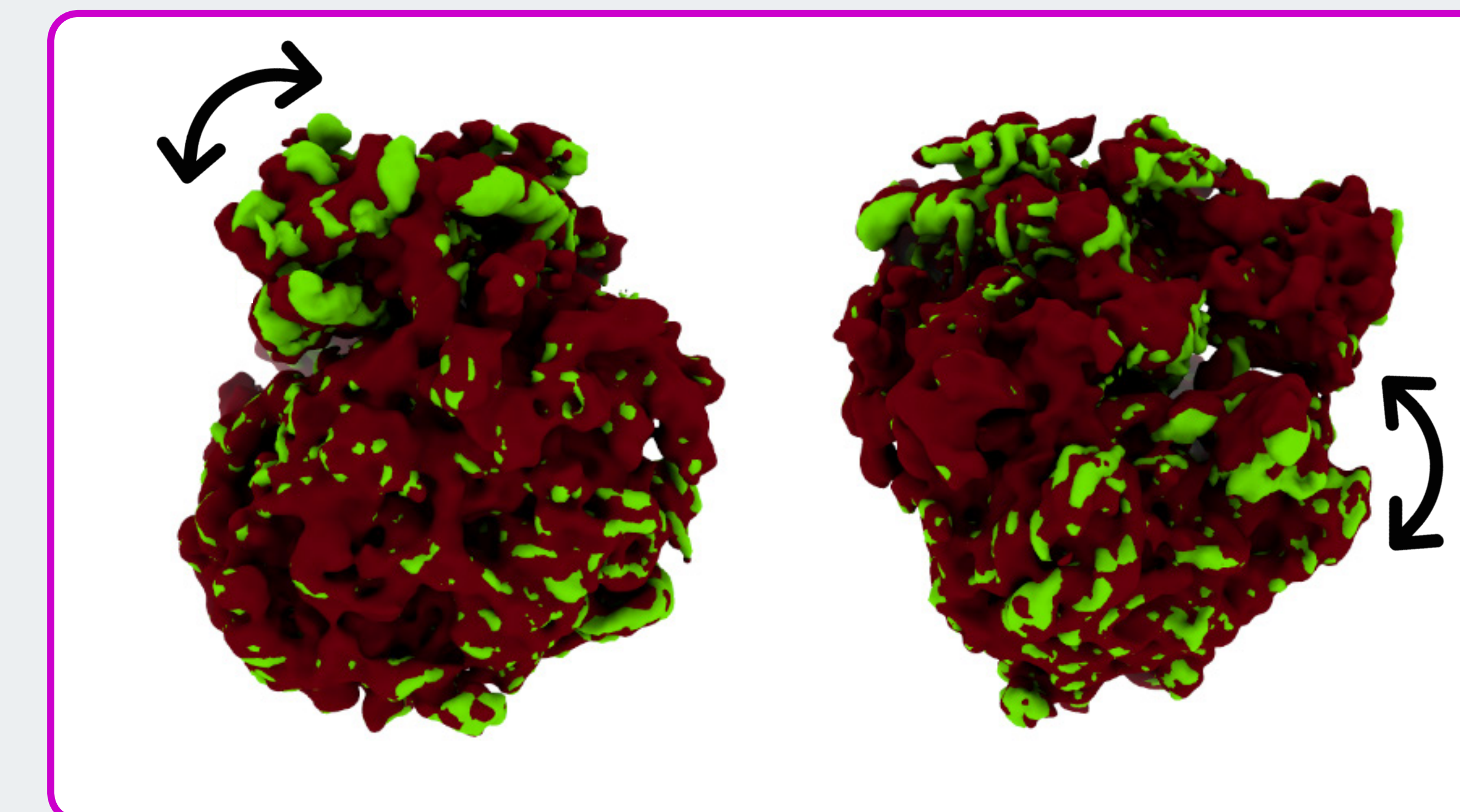
## References

1. C.O.S. Sorzano, et a. Survey of the analysis of continuous conformational variability of biological macromolecules by electron microscopy. *Acta Crystallogr.*, 75:19–32, jan 2019.
2. J. Frank, A. Ourmazd. Continuous changes in structure mapped by manifold embedding of single-particle data in cryo-EM. *Methods*, 100:61-67, may 2016.
3. M.Chen, M.L., Baker and S.J. Ludtke. Application of Deep Learning to CryoEM Heterogeneity Analysis. *Biophys. Journal*, 110:158a, feb 2016.

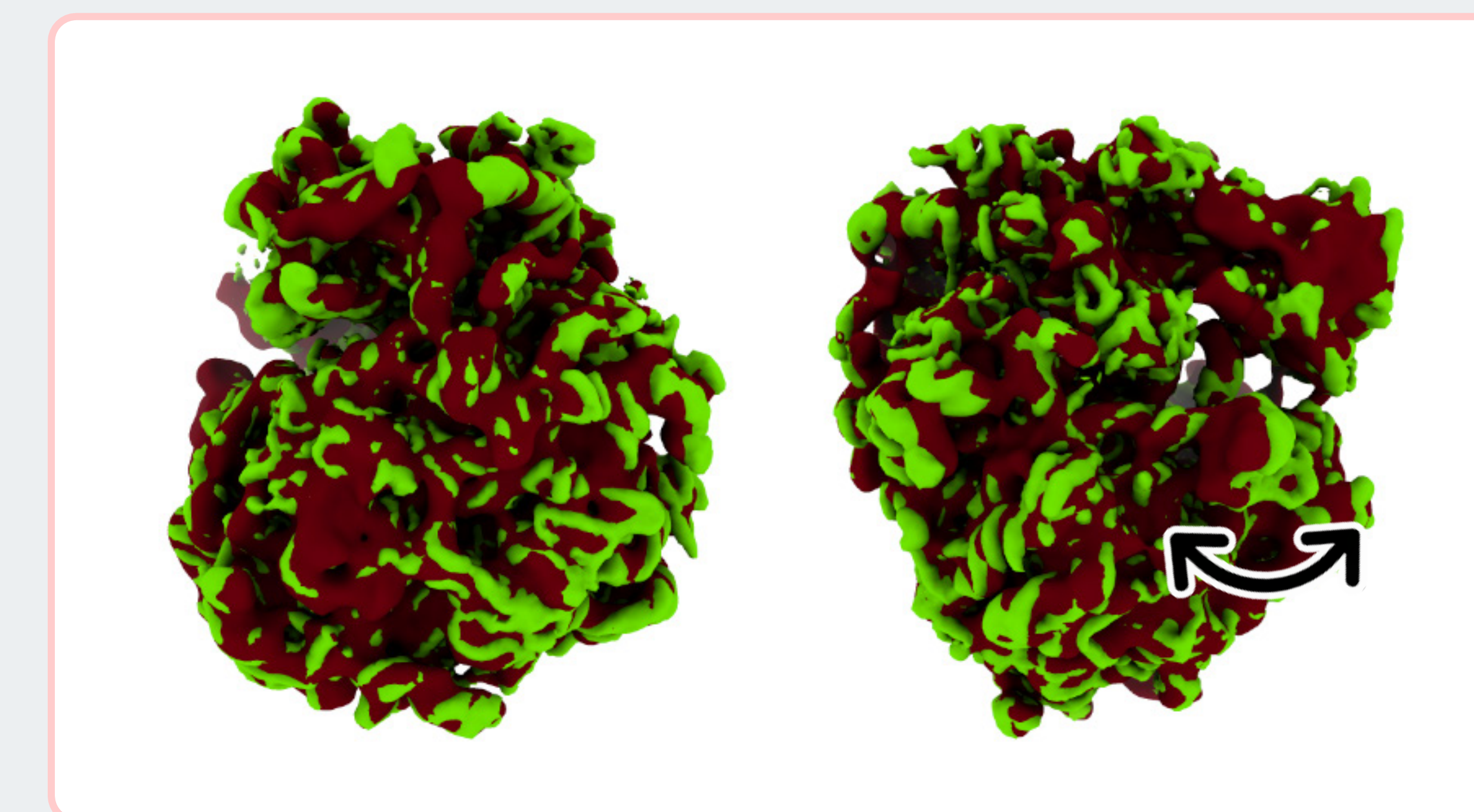
## Results



Zernike3D coefficient space (PCA) - dataset 10028 from EMPIAR  
Here we also introduce our new tool for embedding analysis



Conformation retrieve from the particles corresponding to the purple cluster in the Zernike3D embedding (Cryosparc refinement)



Conformation retrieve from the particles corresponding to the pink cluster in the Zernike3D embedding (Cryosparc refinement)