

Deformation fields approximation by 3D Zernike polynomials for the analysis of continuous heterogeneity

David Herreros Calero

Programa de Doctorado en Ingeniería Informática y de
Telecomunicación

Centro Nacional de Biotecnología - CSIC

Departamento de Ingeniería Informática

Madrid, 2023



UNIVERSIDAD AUTÓNOMA DE MADRID

**DEPARTAMENTO DE INGENIERÍA
INFORMÁTICA**

**Deformation fields approximation by 3D
Zernike polynomials for the analysis of
macromolecular continuous heterogeneity**

David Herreros Calero

Graduado en Ingeniería Biomédica

Para optar al grado de Doctor en Ingeniería Informática y de
Telecomunicación por la Universidad Autónoma de Madrid

Director:

Prof. José María
Carazo García

Co-director:

Dr. Carlos Óscar
Sánchez Sorzano

Tutor:

Dr. Roberto Marabini
Ruiz

CENTRO NACIONAL DE BIOTECNOLOGÍA – CSIC

Madrid, 2023

El trabajo presentado en esta memoria ha sido realizado en el Departamento de Estructura de Macromoléculas del Centro Nacional de Biotecnología bajo la dirección del Prof. José María Carazo García y el Dr. Carlos Óscar Sánchez Sorzano.

Declaration

Hereby I declare that this thesis is my original work, which I have developed on my own during my PhD studies. All resources, references, and literature needed for the elaboration of this work are properly cited and listed according to the due source.

Acknowledgements

I would like to thank everyone who has participated, helped and accompanied me in the development of this thesis, specially to my directors José María Carazo and Carlos Óscar Sánchez Sorzano, my colleagues at the Biocomputing Unit in the National Centre of Biotechnology, and my family. Thank you so much for giving me the opportunity to enjoy such an amazing journey all together.

Agradecimientos

Quiero agradecer a todos los que han participado, ayudado y acompañado en el desarrollo de esta tesis, especialmente a mis directores José María Carazo y Carlos Óscar Sánchez Sorzano, a mis compañeros en la Unidad de Bioinformática del Centro Nacional de Biotecnología y a mi familia. Muchas gracias por darme la oportunidad de disfrutar de un viaje tan increíble todos juntos.

Abstract

Investigating the close relationship between the three-dimensional structure and biological function of macromolecules is challenging because of their dynamic nature. This intrinsic dynamism allows them to adopt different three-dimensional structures, giving rise to conformational variability that gives them the ability to alter their biological function depending on the needs and conditions of their environment. To respond to the processes and characteristics responsible for this variability, structural biology focuses its efforts on understanding the molecular phenomena responsible for changes in the three-dimensional structure of a macromolecule.

Among the structural biology techniques, CryoEM Single Particle Analysis (SPA) has become one of the preferred choices for the analysis of the structural characteristics of macromolecular complexes. By imaging a sample at different poses, it is possible to recover the 3D density map of a macromolecule directly from the CryoEM images, thus allowing to understand the structure adopted by the sample, as well as the biological implications in the macromolecule function derived from its conformational variability.

However, the understanding on how macromolecule structure and biological function meet from just a single conformational state is limited. Macromolecules are dynamic entities able to adopt and modify their structure in response to their biological environment, allowing them to modify their function to accomplish a given task. Therefore, reconstructing a single CryoEM map from the macromolecule images discards a large amount of key structural information about the dynamics of the sample, information that would be needed to have a more complete biological picture of the sample under study.

Instead of restricting the reconstruction of a macromolecule to a limited number of discrete states, it should be possible to take advantage of the fact that the CryoEM microscope can capture many conformational states when a sample is imaged, providing a large amount of structural data with the potential of unveiling the experimental

conformational landscape a macromolecule is most likely exploring under the CryoEM conditions. Thus, restricting protein reconstruction to just one (or a small number) of structural states is not a limitation of the CryoEM technique, but an intrinsic limitation of the assumptions currently taken by image processing algorithms when faced against the conformational variability challenge.

To provide more insights on the structural variability of a macromolecule in CryoEM, we present in this dissertation (structured as a compendium of articles) a new advanced image processing method called Zernike3D specifically designed to tackle the heterogeneity problem from CryoEM data. The Zernike3D algorithm has proven to be a versatile tool when analyzing continuous flexibility, even when applied to different CryoEM datatypes. The Zernike3D method enables a more accurate identification of the motions describing a given conformational change, allowing to extract as many conformational states as CryoEM maps, particles, or structural models involved in the analysis, information that can be posteriorly gathered to describe an approximation of a macromolecule conformational landscape.

In addition to the Zernike3D method, we propose a new reconstruction algorithm able to correct for the elastic deformation (previously estimated with the Zernike3D method) of CryoEM particles called ZART. Thanks to the heterogeneity correction and the characteristics of ZART implementation, it is possible to generate CryoEM maps with reduced motion blur artefacts induced by the conformational variability of the sample, thus enhancing hidden structural features and increasing the resolution of the CryoEM reconstructions.

Keywords

Cryogenic electron microscopy (CryoEM), Single Particle Analysis (SPA), structural biology, image processing, conformational variability, macromolecular flexibility, Zernike3D, Zernike3D-based Algebraic Reconstruction Technique (ZART).

Resumen

Investigar la estrecha relación presente entre la estructura tridimensional y la función biológica de las macromoléculas supone un desafío debido a la naturaleza dinámica de las mismas. Este dinamismo intrínseco les permite adoptar diferentes estructuras tridimensionales, dando lugar a una variabilidad conformacional que les brinda la capacidad de alterar su función biológica dependiendo de las necesidades y condiciones de su entorno. Para dar respuesta a los procesos y características responsables de esta variabilidad, la biología estructural centra sus esfuerzos en el entendimiento de los fenómenos moleculares responsables de los cambios en la estructura tridimensional de una macromolécula.

Entre las técnicas de biología estructural, el Análisis de Partículas Individuales (SPA) en CryoEM se ha convertido en una de las opciones preferidas para el estudio de las características estructurales de los complejos macromoleculares. Al obtener imágenes de una muestra en diferentes poses, es posible recuperar el mapa de densidad 3D de una macromolécula directamente a partir de dichas imágenes, permitiendo comprender la estructura adoptada por la muestra, así como las implicaciones biológicas en la función de la macromolécula derivadas de su variabilidad conformacional.

Sin embargo, la comprensión de cómo la relación entre la estructura de la macromolécula y su función biológica a partir de un único estado conformacional es limitada. Las macromoléculas son entidades dinámicas capaces de adoptar y modificar su estructura en respuesta a su entorno biológico, lo que les permite modificar su función para realizar una tarea determinada. Por lo tanto, la reconstrucción de un único mapa en CryoEM a partir de las imágenes de macromoléculas descarta una gran cantidad de información estructural clave para comprender la dinámica de la macromolécula, información que sería necesaria para tener una imagen biológica más completa de la muestra en estudio.

En lugar de restringir la reconstrucción de una macromolécula a un número limitado de estados discretos, es posible aprovechar la capacidad del microscopio de capturar una

cantidad considerable de estados conformacionales en cada imagen de una muestra, proporcionando una gran cantidad de datos estructurales con el potencial de revelar el paisaje conformacional experimental que una macromolécula explora bajo las condiciones impuestas durante la preparación de la muestra. Por lo tanto, restringir la reconstrucción de proteínas a solo uno (o un pequeño número) de estados estructurales no es una limitación de la técnica, sino una limitación intrínseca de las asunciones actualmente tomadas por los algoritmos de procesamiento de imágenes cuando se enfrentan una muestra con un alto grado de variabilidad conformacional.

Para proporcionar más información sobre la variabilidad estructural de una macromolécula en CryoEM, presentamos en esta tesis doctoral (estructurada como un compendio de artículos) un nuevo método avanzado de procesamiento de imágenes llamado Zernike3D, diseñado específicamente para abordar el problema de la heterogeneidad presente en los datos de CryoEM. El algoritmo Zernike3D ha demostrado ser una herramienta versátil al analizar la flexibilidad continua, incluso cuando se aplica a diferentes tipos de datos. El método permite, además, una identificación más precisa de los movimientos que describen un cambio conformacional determinado, permitiendo extraer tantos estados conformacionales como mapas, partículas o modelos estructurales involucrados en el análisis, información que se puede recopilar posteriormente para describir una aproximación de un paisaje conformacional para la macromolécula.

Además del método Zernike3D, proponemos un nuevo algoritmo de reconstrucción capaz de corregir la deformación elástica (previamente estimada con el método Zernike3D) de las partículas en CryoEM llamado ZART. Gracias a la corrección de heterogeneidad y las características de la implementación del método, es posible generar mapas con una reducción de los artefactos de desenfoque de derivados de variabilidad conformacional de la muestra, mejorando así las características estructurales y aumentando la resolución de las reconstrucciones.

Palabras clave

Criomicroscopía electrónica (CryoEM), Análisis de partículas individuales (SPA), biología

estructural, procesamiento de imágenes, variabilidad conformacional, flexibilidad macromolecular, Zernike3D, Zernike3D-based Algebraic Reconstruction Technique (ZART).

Table of contents

Introduction	1
1.1. Objectives	3
1.2. Preliminaries	4
1.2.1. CryoEM processing workflow	4
1.2.2. CryoEM 3D classification.....	7
1.3. State of the art	8
1.3.1. Overview	8
1.3.2. Density-based heterogeneity approach	9
1.3.3. Deformation field-based approach.....	10
1.3.4. ManifoldEM	11
1.3.5. CryoDRGN	11
1.3.6. Gaussian Mixture Models	12
1.3.7. Normal Mode Analysis.....	12
1.3.8. 3DFlex	13
Methodology and results	14
2.1. Mathematical description.....	14
2.1.1. The Zernike3D basis	14
2.1.2. Closure under rotations	16
2.1.3. Numerical stability of the Zernike3D basis	16
2.1.4. Deformation field scaling.....	17
2.1.5. ZART	18
2.2. Zernike3D basis applications.....	21
2.2.1. Application to CryoEM density maps.....	21
2.2.2. Application to structural models	23
2.2.3. Application to CryoEM particles	25
2.3. Results and discusión.....	28
2.3.1. Zernike3D application to density maps	28
2.3.2. Zernike3D application to particles	32
2.3.3. ZART reconstruction	39
Conclusions and future work	45
3.1. Future work.....	46
Conclusiones y futuro	48
3.2. Trabajo futuro	49
References.....	51
Appendices.....	55
Approximating deformation fields for the analysis of continuous heterogeneity of biological macromolecules by 3D Zernike polynomials.....	56

Estimating conformational landscapes from Cryo-EM particles by 3D Zernike polynomials
..... 71

*ZART: A Novel Multiresolution Reconstruction Algorithm with Motion-blur Correction for
Single Particle Analysis..... 83*

Lists

List of Figures

<i>Figure 1</i>	4
<i>Figure 2</i>	5
<i>Figure 3</i>	6
<i>Figure 4</i>	8
<i>Figure 5</i>	15
<i>Figure 6</i>	26
<i>Figure 7</i>	27
<i>Figure 8</i>	30
<i>Figure 9</i>	31
<i>Figure 10</i>	32
<i>Figure 11</i>	33
<i>Figure 12</i>	34
<i>Figure 13</i>	36
<i>Figure 14</i>	37
<i>Figure 15</i>	38
<i>Figure 16</i>	38
<i>Figure 17</i>	39
<i>Figure 18</i>	40
<i>Figure 19</i>	41
<i>Figure 20</i>	42
<i>Figure 21</i>	43

Introduction

Cryogenic Electron Microscopy (CryoEM) has become over the years one of the main fields in structural biology to understand the structural characteristics of macromolecular complexes. Since it originated in 1960s [1], CryoEM has undergone a series of breakthroughs before becoming the versatile technique we know nowadays.

One of the main advances made in CryoEM is the set of technological changes that were captured under the words *resolution revolution* which supposed the end of the low-resolution revolution era around 2013. Before the resolution revolution, the CryoEM density maps resolved from the acquired particles were limited to a spatial resolution of around 10 Å. At this resolution levels, the biological knowledge that could be drawn from the CryoEM maps supposed a major challenge, as most of the protein structure was represented by a set of *“blobs”* without internal structural features. The publication of the first near-atomic resolution structure [2] from non-crystalline data at 3 Å put an end to the *“blob-era”* with the culmination of a Nobel Prize granted to Jacques Dubochet, Joachim Frank, and Richard Henderson in 2017.

From this point on, the interest in CryoEM increased, due to its capacity to estimate high resolution macromolecular structures while overcoming the main limitations of other popular techniques such as X-ray crystallography [3] and Nuclear Magnetic Resonance (NMR) [4]. Still, there were challenges to overcome at the image processing level, due to the complex conditions the acquired CryoEM particles are subjected to: CTF, low signal-to-noise ratio, conformational variability...

One of the main advantages of CryoEM is the ability to capture macromolecular complexes at different conformational states. In other techniques, it is only possible to represent a macromolecule by a single 3D structure, suggesting that the macromolecule under study is static. The previous assumption is, however, not true in a biological

environment. For example, proteins exhibit an intrinsic dynamic nature, allowing them to restructure under different conditions to accomplish a given biological task. Thus, without explicitly addressing the dynamic nature of proteins it would not be possible to drive the cellular processes that make life possible. Therefore, the structural characteristics of proteins are strongly related to their biological function.

Therefore, the main objective of structural biology is not only to resolve and understand the 3D structure a given macromolecular complex adopts, but to understand how the resolved structure and biological function meet to drive a relevant biological process, and how dynamics play a role in the modification of this function.

Classically, the extraction of the different experimental conformations captured by CryoEM has relied on a process called 3D classification [5, 6, 7]. During a 3D classification, images corresponding to a different conformation are identified, isolated, and used to reconstruct several conformational states. However, 3D classification has two main limitations:

1. It is only possible to obtain a limited number of states (that, furthermore, are predefined by the user) reliably out of the whole conformational landscape captured by the CryoEM particles.
2. The reconstruction process of a given state supposes an averaging of the macromolecule signal, yielding a map which may still contain a mixture of conformational states.

During this dissertation, structure as a compendium of articles, a new method able to overcome the limitations of classical 3D classification will be presented and discussed. The dissertation is organized as follows: Chapter 1 focuses on the description of the objectives of the dissertation (Chapter 1.1), followed by a presentation of a typical CryoEM pipeline and a brief introduction to the state of the art in CryoEM conformational heterogeneity methods (Chapters 1.2 and 1.3 respectively). Chapter 2 presents the proposed method and main results. Chapter 3 includes a list of all author publications. Chapter 4 summarizes the conclusions and future steps of this work.

1.1. Objectives

This dissertation, structured as a compendium of articles, introduces a new CryoEM algorithm specifically developed to tackle the conformational heterogeneity problem in CryoEM Single Particle Analysis (SPA) [8]. The new method, called Zernike3D, has been specifically designed to estimate per-particle conformational states, allowing to estimate richer conformational landscapes compared to classical methods like 3D classification. The dissertation will be mostly focused on the mathematical description of the new algorithm, as well as its application to the main CryoEM datatypes (particles, volumes, and structural models).

During the development of the Zernike3D algorithm, we identified the following research topics:

- Definition of the Zernike3D basis, focusing on its ability to approximate molecular motions in the form of a deformation field expressed in term of the basis components.
- Application of the Zernike3D basis to estimate conformational states starting from diverse types of CryoEM data (particles, volumes, and structural models), as well as overcoming the challenges each different datatype introduces in the identification of a meaningful motion.
- Analysis of the estimated Zernike3D deformation fields to extract the main forces driving a molecular transition in the form of a strain and rotation field.
- Application of the estimated deformation fields to perform a finer correction of the conformational heterogeneity during the reconstruction process in a new algorithm called ZART (Zernike3D-based Algebraic Reconstruction Technique). The new reconstruction algorithm has proven to be able to further improve the resolution of highly flexible molecular regions compared to standard CryoEM reconstruction algorithms.

The Zernike3D algorithm, as well as their specific implementations needed to deal with the main CryoEM datatypes, have been implemented in the software Xmipp [9] and Scipion [10], allowing them to be easily accessed and used by the CryoEM community.

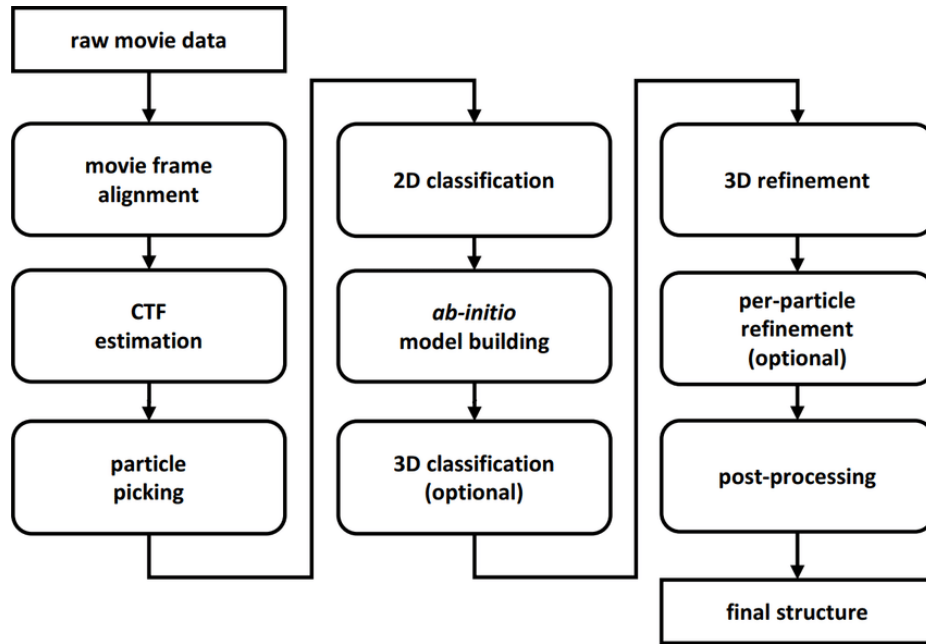


Figure 1: Scheme summarizing the main steps followed in a CryoEM image processing workflow [76].

1.2. Preliminaries

1.2.1. CryoEM processing workflow

The following sections provides an overview of the processing pipeline commonly followed in CryoEM Single Particle Analysis (SPA). To complete the description of the SPA workflow, a graphical representation summarizing all the step to be described is provided in Figure 1.

Note that the section will present only on the image acquisition and image processing steps and skip the sample preparation. Regarding the preparation of the CryoEM samples, it suffices to say that it involves the preparation of a solution containing the samples, which is posteriorly deposited on a grid made of carbon and gold. Once the sample has been placed in the grid, it is vitrified prior to imaging.

CryoEM grids are composed by many patches containing the vitrified sample. Vitrification of the sample in an aqueous solution is essential to avoid sample denaturalization in vacuum, which is a condition needed by the CryoEM microscope to operate properly. Ideally, the vitrification should lead to a thin layer of vitreous ice, thus preventing the overlapping of particles on top of each other. A scheme of a typical CryoEM grid is shown in Figure 2.

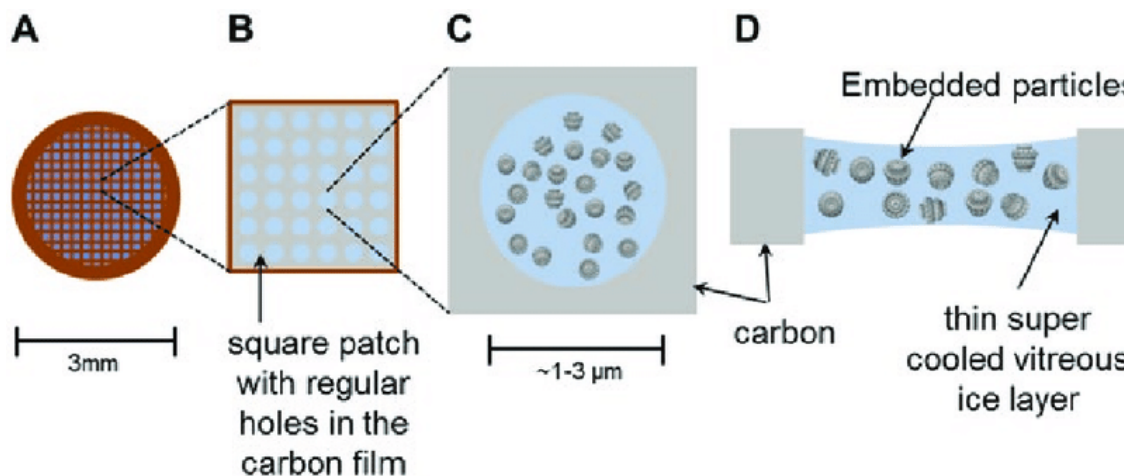


Figure 2: Representation of a CryoEM grid after vitrification [77].

Within a grid patch, the orientational arrangement of the sample will be random. The previous condition is beneficial for SPA, as it allows capturing many projection images at different projection orientations, information that can be posteriorly used to reconstruct the CryoEM density map and analyze the sample structure. However, in some cases the distribution of orientations might exhibit some preferred orientations that might impact negatively in the quality of the CryoEM reconstruction.

From all the different patches in grid, only those showing a good distribution of particles and quality are imaged. The quality of the patches can be analyzed in an automatic or manual way, allowing to define an acquisition planning.

The image formation process in Transmission Electron Microscopy (TEM) relies on a beam of electrons flowing through the microscope column and reaching the path to be image. At this point, the interaction of the electrons with the sample in the patch will determine whether they are able to reach the electron detector or not. The number of electrons reaching the detector is counted and arranged in an image called “*frame*”. The number of electrons that interact with the patch is represented by a number called “*dose*” measured in number of electrons per \AA^2 . The dose plays a key role in the imaging process, as it controls the tradeoff between radiation damage and the signal to noise ratio of the frames. In general, dose is kept low to reduce radiation image as much as possible. The CryoEM image formation process is exemplified in Figure 3.

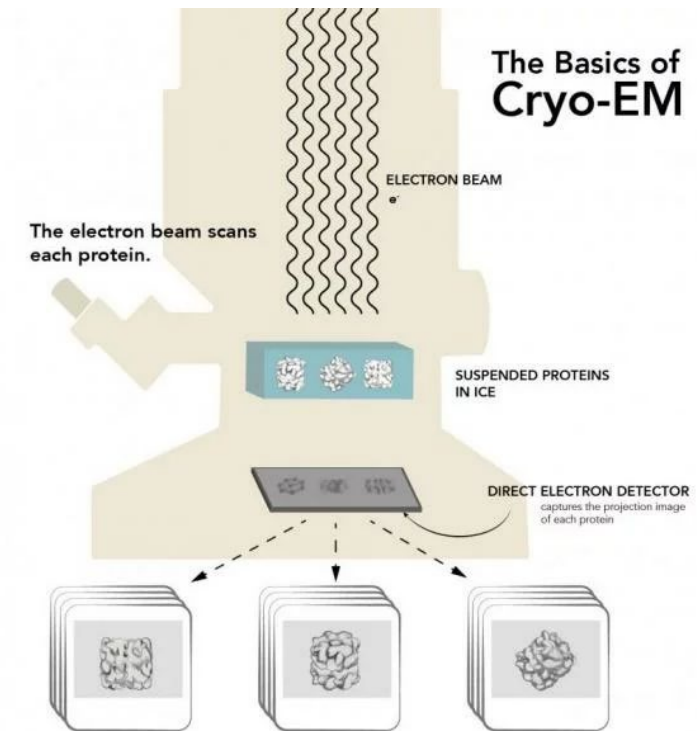


Figure 3: Summary of the image formation process in CryoEM (modified from [80]).

In addition, exposing the sample to the electrons for a long time might also lead to undesired effects. Therefore, it is common to acquire not one but several frames from the same patch, which are posteriorly grouped to create a “movie”.

The next step in the image processing workflow consists of aligning the different frames [11, 12, 13] composing a movie to posteriorly compose them in a single image with improved signal to noise ratio called “micrograph”. Once the micrograph is obtained, it is possible to further analyze it to estimate the Contrast Transfer Function (CTF) [14, 15, 16], which is used later in the workflow to correct for the microscope defocus. CTF correction is mandatory to push the resolution of a map to near atomic levels.

Apart from the CTF estimation, the micrograph can also be used to identify and isolate the sample into smaller images called “particles” in a process called picking [17, 18, 19]. Therefore, ideally a particle should contain only a single projection of the sample at a given orientation. In practice, the picking process will yield a mixture of ideal and unwanted particles, which need to be cleaned before feeding them into the reconstruction process.

The cleaning step of picked particles is referred as 2D classification [20, 21, 22]. During this process, particles coming from a similar projection orientation are identified, aligned, and averaged to generate a cleaner particle image from the sample. If a 2D class is mostly

composed by garbage particles, it will be averaged into a blurred 2D class, which can be used to filter and clean the particle set.

Good 2D averages can also be used to generate an initial reconstruction of the sample structure called initial volume [23, 24, 25]. In general, initial volumes end up having very low resolution, but they represent a good starting point to enter the 3D refinement step [26, 27]. During the 3D refinement, particle orientations are finely tuned, progressively improving the reconstruction resolution.

Once the refinement processed has finished, it is possible to further analyze the density map to either improve it (e.g., apply a sharpening step [28, 29, 30], denoising [31, 32, 33], resolution estimation [34, 35, 36]...) or trace the atomic structure if the resolution allows for it [37, 38, 39]. In general, if we assume that the sample is almost static, the previous workflow should lead to a good resolution map. However, the conformational variability captured in the particle images plays a significant role in the maximum quality achievable. For that reason, it is also common to analyze the conformational variability through a 3D classification.

1.2.2. CryoEM 3D classification

The most established method in CryoEM to disentangle the conformational variability information captured in a set of particles is 3D classification. The aim of the classification approach is to identify those particles whose projections match the structure of K different 3D objects. Therefore, 3D classification leads to a coarse approximation of the conformational landscape of a macromolecule called discrete heterogeneity.

Many classification methods have been proposed. For example, Relion [40] 3D classification takes advantage of maximum likelihood estimators to determine the probability that a given projection comes from one of the K possible underlying 3D structures that the algorithm will try to generate. In 3D classification, the number of underlying objects K is a parameter that needs to be set by the user. Although it is possible to ask a classification algorithm to look for many classes, most of the time only a small fraction of the classes found will attract enough particles to yield meaningful conformations. Therefore, 3D classification is limited to reconstruct the most stable structural states a given macromolecule adopts.

It should be noted that the discrete heterogeneity approximation 3D classification

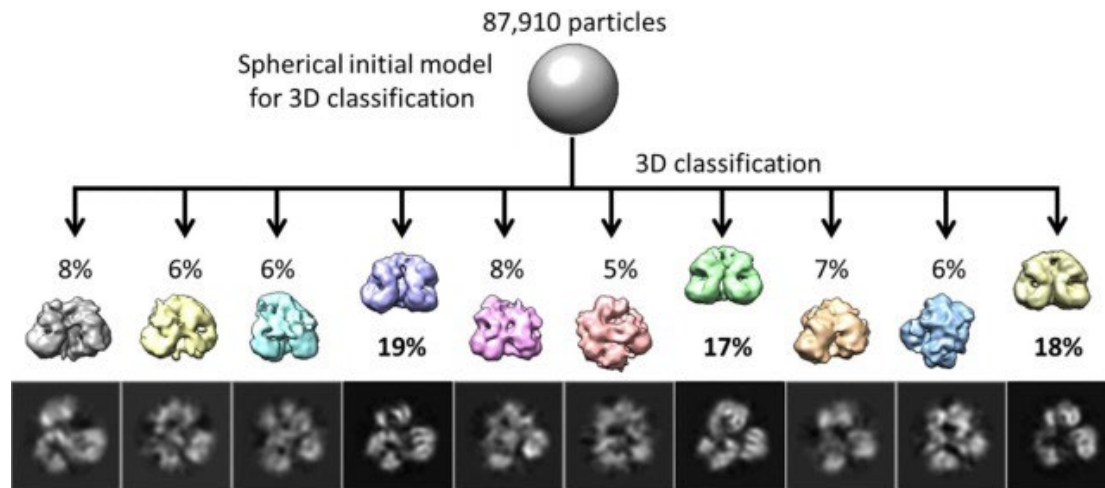


Figure 4: Example of a 3D classification of the Contorsbody analog 2 protein (modified from [78]).

proposes has the capability to identify two types of variability: compositional variability and conformational variability. The main difference between the two is that compositional variability is mostly related to changes in the mass of the density map (e.g., a region of the macromolecule that is lost). In the case of conformational variability, structural differences are due to macromolecular motions.

An example of a 3D classification process is provided in Figure 4.

1.3. State of the art

This chapter describes the current state of the art regarding the estimation of conformational landscapes in CryoEM. The chapter will start with an overview of the different approaches currently available to approximate conformational states from CryoEM data, followed by a more in detail description of the different algorithms being currently developed to accomplish the previous task.

1.3.1. Overview

The exploration of macromolecular conformational states is not new in CryoEM [41], although it has drawn a large attention lately due to the development of advanced algorithms capable of overcoming the main limitations of classical processing methods in a wide range of cases.

Therefore, we could consider the conformational heterogeneity field is suffering a revolution, allowing researchers to extract for the first-time experimental approximations of

conformational landscapes, in contrast with the discrete approaches that limit the analysis to a small number of stable/highly represented states.

The new conformational heterogeneity software could be classified based on the approach they follow to estimate conformational states from CryoEM data (usually, from CryoEM particles, as they capture a considerable number of states enough to approximate experimental conformational landscapes:

- **Density-based approaches:** most software fall into this group, including the classical approach of 3D classification as well. Currently, CryoDRGN [42], Gaussian Mixture Models [43], and ManifoldEM [44] constitute the main representatives of this group.
- **Deformation field-based approaches:** this approach is mostly followed when working with structural model, as it provides a very convenient way to deal with atomic positions. However, efforts are also being made to adapt their usage to the discretize world defined by voxels and pixels. 3DFlex [45] or Normal Mode Analysis [46] belong to this group of methods. In addition, the Zernike3D algorithm to be introduced in this dissertation relies also on deformation field principles.

1.3.2. Density-based heterogeneity approach

The density-based heterogeneity approach to estimate conformational heterogeneity could be considered as an evolution of the reconstruction process. Due to the low signal to noise ratios CryoEM particles exhibit, reconstruction requires many particles to average out the noise while giving more strength to the protein signal. However, the reconstruction principles under noisy conditions only hold under the assumption of having a static sample.

If the number of particles is large enough, the 3D classification process will be able to identify the most stable states and isolate the particles that more likely correspond to that structure. However, in order not to sacrifice the protein signal, averaging of a substantial number of particles is still mandatory, leading to a mixture of more consistent conformations that help decreasing the effect of motion blur artefacts.

Therefore, if we want more accurate conformational states, we need to define a reconstruction process able to denoise enough the reconstructed maps from a fewer number

of particles to reduce the conformational averaging. Thus, density-based approaches have arisen as an improved reconstruction that tries to minimize the number of particles needed to recover a meaningful structure, usually referred as “heterogeneous reconstruction”.

One of the main advantages of the heterogeneous reconstruction method is its ability to deal with both, continuous and compositional heterogeneity. Continuous heterogeneity refers to the structural variability directly originated from the dynamic nature of macromolecules, allowing them to adopt different structural states. In contrast, compositional variability arises when the specimen under study suffers changes in its mass composition due to different causes (e.g., loss of regions due to denaturalization, ligand binding...).

1.3.3. Deformation field-based approach

The deformation field-based approaches are specifically designed to estimate the motion that a macromolecule needs to undergo to reach a different conformational state. Therefore, this approach does not estimate directly conformational states, but it defines a warping field that upon application leads to a particular structure.

The most common way to define the previous warping field is as a deformation field. In essence, a deformation field is a composition of vectors, each of which defines the displacement needed to move a given point to a different location. For our purposes, the definition of point is rather general: although it has been commonly associated with an atomic position, it could also be extended to a pixel or voxel location in space.

Deformation fields are the predilect choice when dealing with continuous heterogeneity, as they try to estimate the natural motion that connects two different conformational states. Compared to density based-approaches, deformation fields have proven to be more versatile, as they can be applied directly to different datatypes (particles, volumes, and atomic structures). In addition, they open the possibility to add conformational heterogeneity corrections in a reconstruction process to improve resolution, as it will be discussed in Section 2.1.5. Nevertheless, deformation fields cannot handle compositional heterogeneity unlike density-based approaches, so it needs to be handled beforehand to avoid misleading conformational estimations.

1.3.4. ManifoldEM

ManifoldEM [44] was the first continuous and compositional heterogeneity algorithm introduced in the field, which proposes a slightly different approach compared to the other methods to be presented in this chapter. In ManifoldEM, the conformational landscape represented by a set of images is directly estimated from the images themselves, without the need to generate a 3D representation of the different conformational states.

To that end, ManifoldEM disentangles the orientation and conformational variability by splitting the particle dataset into several projection orientation cones. Ideally, if each cone thickness becomes infinitesimally small, the orientation variability within the cone would be zero. In practice, due to the low signal to noise ratio particles exhibit, there is a tradeoff between cone thickness and orientation variability, although the choice of thickness usually makes this variability negligible.

Within a cone, all particle images capture the protein virtually in the same orientation. Therefore, particle variability within a cone is mostly arising due to conformational variability. The next step ManifoldEM performs is to identify the continuous variational changes occurring in each cone, leading to an embedding approximating the particles' conformational landscape. Lastly, the embedding of each cone is merged in a process called stitching.

1.3.5. CryoDRGN

CryoDRGN [42] is a deep learning-based algorithm designed to tackle the conformational heterogeneity problem with a heterogeneous reconstruction. To that end, CryoDRGN design involves a deep neural network able to learn how to produce heterogeneous reconstructions directly from a set of CryoEM particles with CTF and angular information stored.

CryoDRGN has proven to be a powerful choice to deal with both, compositional and continuous heterogeneity through the estimation of a latent space that summarizes the conformational landscape of the macromolecule under study. The variational autoencoder architecture of CryoDRGN allows to transform any point in the latent space representation into a CryoEM density map by forwarding it through the decoder network. To reconstruct a whole volume, CryoDRGN relies on a positional encoding of the voxel coordinates, which are

combined with the latent space representations to decode the appropriate voxel value needed to recover the desired density map.

1.3.6. Gaussian Mixture Models

Gaussian Mixture Models (GMM) [43] is a density-based compositional and flexibility heterogeneity algorithm included as part of EMAN package [47]. Unlike CryoDRGN, GMM proposes a deep neural network that decomposes the representation of a CryoEM density map as the sum of different Gaussian functions in 3D with variable widths and heights. The resolution and potential overfitting of the generated map will therefore depend on the number of Gaussian involved in the generation process.

GMM proposes a training based on two different steps: Initially, the network learns how to produce a homogeneous (standard) reconstruction from a set of CryoEM particles. Once the first training step is over, the network is fine-tuned with the particle dataset to learn the heterogeneous reconstruction. After the previous two step process, the network will learn to produce the five main parameters needed to describe how to produce the gaussian composition yielding the density map: the gaussian spatial position (X, Y, Z) , and the gaussian widths and heights.

1.3.7. Normal Mode Analysis

Normal Mode Analysis (NMA) [46] is one of the first methods developed to analyze continuous conformational variability around an equilibrium state. The method is based on the description of the oscillations a macromolecule undergoes when it is perturbed from a stable state representing a minimum of energy. The vibration modes associated with the previous oscillation can be combined to describe complex motions around the initial equilibrium state in the form of a deformation field, that can be expressed in terms of the vibration modes and the Normal Mode basis.

Normal Mode Analysis requires a structural model representing the starting equilibrium position previously described. In many cases, the structural model can be traced directly on the CryoEM density map reconstructed from the particles. However, it is also possible to create a pseudoatomic model directly from the map [48] to feed into the analysis.

NMA has also proven to be able to approximate conformational states from particles, as done by HEMNMA-3D [49] (integrated in ContinuousFlex software [50]). However, NMA has an intrinsic limitation on its usage, making the conformational landscape estimation semiautomatic: the vibration modes needed to describe the expected motions in a particle dataset needs to be preselected by the user.

In the next Chapter, we will show how the Zernike3D approach is able to overcome the previous limitation of NMA, making the conformational estimation process fully automatic.

1.3.8. 3DFlex

3DFlex [45] is a deformation field-based software integrated as part of CryoSPARC [51] software. 3DFlex proposes an autoencoder neural network able to learn both, a homogeneous reconstruction obtained from the input particles, and a latent space that approximates the conformational landscape captured by the particles. By forwarding a latent space point into the decoder, 3DFlex network can translate the latent space representation into a deformation field used to warp the learn volume to generate a different conformation.

In addition, the combined learning of the map and conformational latent space coordinates to yield a homogeneous “consensus” reconstruction with reduced motion blur, as the deformation fields allow the network to learn how to correct for the conformational heterogeneity present in the particle dataset.

Methodology and results

The next Chapter focuses on the description of the Zernike3D basis and its possibilities, followed by the results obtained during the application of this method to different scenarios.

The results included and discussed in this Chapter has been taken from the publications listed and included as part of this dissertation.

The Chapter is structured as follows: the first sections provide the mathematical details of the Zernike3D basis and its main properties. Sections following the mathematical description of the basis offer more insights on how to apply the Zernike3D method to estimate the conformational variability from different CryoEM data types (volumes, structural models, and particles), as well as its application in ZART to correct the conformational heterogeneity to improve CryoEM reconstructions. Lastly, the main results obtained during the development of the Zernike3D and ZART methods will be presented and discussed.

2.1. Mathematical description

2.1.1. The Zernike3D basis

To express the deformation field, we propose in this work a mathematical basis able to expand any function in the unit ball B . In general, the expansion could take the form:

$$\mathbf{g}(\mathbf{r}) = \sum_{l=0}^{\infty} \sum_{n=0}^{\infty} \sum_{m=-l}^l \alpha_{l,n,m} Z_{l,n,m}(\mathbf{r}) \quad (1)$$

Where $Z_{l,n,m}$ represents the Zernike3D basis and $\alpha_{l,n,m}$ is a Zernike3D coefficients determining the contribution of that component to the final expansion. It is important to note that the choice of the Zernike3D basis is not unique, as it is possible to find other

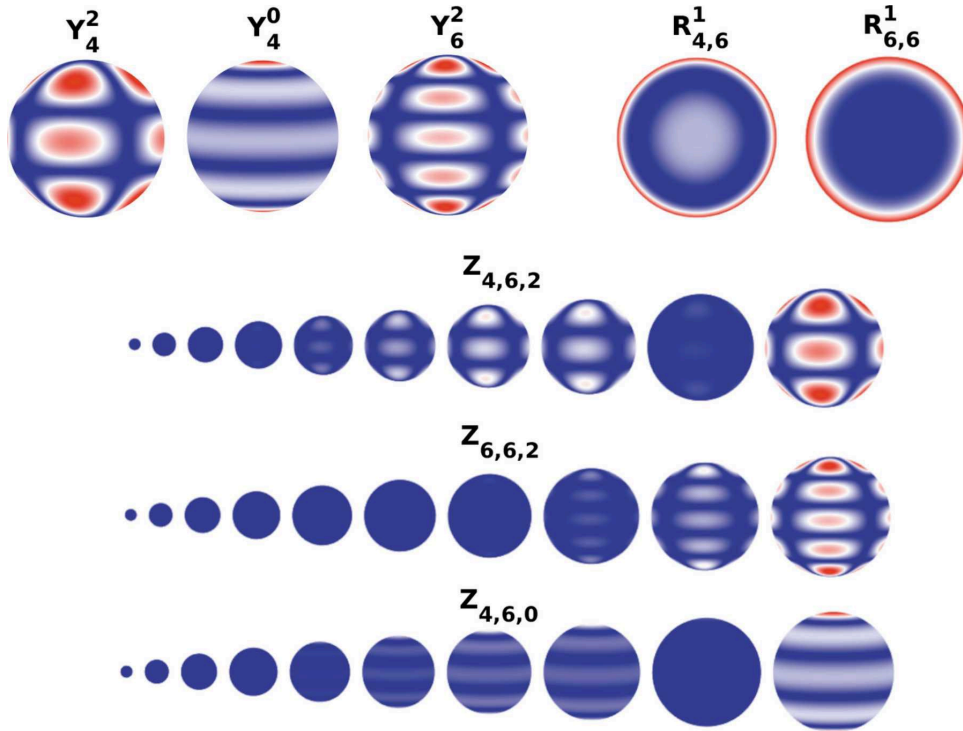


Figure 5: Example of some Zernike3D basis components. The functions represented at the top show the spherical harmonics Y_l^m and the Zernike polynomials $R_{l,n}^1$ used for the composition of the basis components. The Zernike3D components $Z_{l,n,m}$ shown are defined in the unit ball B . To simplify the visualization of the component variation along the ball, different shells have been represented [52].

sensible choices of function combinations that would give rise to valid basis. However, the Zernike3D basis properties become helpful when applied to the conformational heterogeneity problem, as it will be discussed during the next sections.

The Zernike3D basis $Z_{l,n,m}$ is defined in terms of two different set of functions, the 3D real-valued (normalized) generalized Zernike polynomials (radial component) and the real-valued spherical harmonics (angular component). References to the expressions of the functions composing the Zernike3D basis, as well as a table containing the expression of some of the basis components can be found here [52]. For this dissertation, it suffices to say the final expression of the Zernike3D basis can be obtained as:

$$Z_{l,n,m}(\mathbf{r}) = \bar{R}_{l,n}^1(r)y_l^m(\theta, \phi) \quad (2)$$

where $\bar{R}_{l,n}^1(r)$ represents the 3D real-valued and normalize Zernike polynomials and $y_l^m(\theta, \phi)$ refers to the real-valued spherical harmonics, r, θ, ϕ are the spherical coordinates of the point being considered. The values l and n are non-negative integers referring to the degrees of the spherical harmonics and Zernike polynomials, respectively, and m is an integer such that $-l \leq m \leq l$. An example of some Zernike3D basis components is provided in Figure

5.

Similarly to other common basis in image processing such as the Fourier Transform, the basis degrees l and n control the frequency of the basis components. In macromolecular motions, smaller degrees of the basis will be able to represent collective motions (low frequency), while higher degrees will be responsible for approximating more localized movements (high frequency).

2.1.2. Closure under rotations

One of the most interesting properties of the Zernike3D basis is that it is closed under rotations, meaning that changes in the rotation of the reference frame deformation fields can also be represented by the Zernike3D basis. The previous property holds as long as the origin of the reference frame does not change.

Since a deformation field is just a 3D vector associated with a given position in space \mathbf{r} , thanks to the closure under rotation property we can propagate any rotation applied to the field as follows:

$$\mathbf{A} \mathbf{g}_L(\mathbf{A}^{-1} \mathbf{r}) = \sum_{l=0}^{\infty} \sum_{n=0}^{\infty} \sum_{m=-l}^l \mathbf{A} \alpha_{l,n,m} Z_{l,n,m}(\mathbf{A}^{-1} \mathbf{r}) \quad (3)$$

Being \mathbf{A} a valid unitary rotation matrix.

It is important to mention that the closure under rotation property is only true for continuous spaces. In case any discretization is applied (as happens with arrays of pixels and voxels), closure under rotation no longer holds.

2.1.3. Numerical stability of the Zernike3D basis

The series expansion representation of a complex function in any given basis requires an infinite number of components to yield a perfect approximation. As this is unattainable in practice, series expansions need to be truncated to a fixed number of components. Therefore, there exists a tradeoff between the computational complexity and the accuracy of the series expansion that depends on the truncation applied.

Thus, the Zernike3D basis needs to be truncated in practice to a limited number of

components that will determine the accuracy and freedom the deformation field must represent a given motion. However, macromolecular motions tend to be collective [53], meaning that they can be approximated accurately by smooth deformation fields. Therefore, the number of Zernike3D components needed to expand macromolecular motions does not need to be large to get an accurate estimation of a given conformational transition.

2.1.4. Deformation field scaling

The application of deformation fields to different CryoEM data type is possible thanks to the possibility of scaling the field to match grids at different resolutions. For example, a deformation field referred to a CryoEM density map can only be applied to its corresponding structural model if the scaling associated with the grid sampling rate is considered. Another example is the possibility to compute the deformation field in a low-resolution version of an image for performance purposes and apply it later to warp the original resolution images.

Although scaling a deformation field is usually fast, as the number of deformation fields that need to be scaled becomes larger it can have a significant negative impact on the computational complexity. Luckily, the expansion of a deformation field in a series of Zernike3D coefficients can reduce considerably the number of operations needed to perform to successfully scale a field.

The scaling of an arbitrary deformation field vector \mathbf{r}_o by a factor k can be expressed as:

$$\mathbf{r}_d = k\mathbf{r}_o \quad (4)$$

If we expand the vectors \mathbf{r}_o and \mathbf{r}_d in term of the Zernike3D basis we reach:

$$\begin{aligned} \sum_{l=0}^{\infty} \sum_{n=0}^{\infty} \sum_{m=-l}^l \alpha_{l,n,m}^d Z_{l,n,m}(\mathbf{r}_d) &= k \sum_{l=0}^{\infty} \sum_{n=0}^{\infty} \sum_{m=-l}^l \alpha_{l,n,m}^o Z_{l,n,m}(\mathbf{r}_o) \\ &= \sum_{l=0}^{\infty} \sum_{n=0}^{\infty} \sum_{m=-l}^l (k\alpha_{l,n,m}^o) Z_{l,n,m}(k^{-1}\mathbf{r}_d) \end{aligned} \quad (5)$$

Therefore, from (5) we can conclude that scaling the Zernike3D deformation coefficients $\alpha_{l,n,m}$ is enough to a new expansion able to reproduce the scaled field (as long as the basis is evaluated in the correct positions).

$$\alpha_{l,n,m}^d = k\alpha_{l,n,m}^o \quad (6)$$

2.1.5. ZART

Although the most logical application of a conformational variability analysis is the exploration of the different conformational states a macromolecule may adopt, it is also possible to take advantage of it to perform some corrections.

Being able to reach high resolutions in CryoEM density maps is of great importance, as it simplifies tracing the chain of atoms that define the real underlying atomic structure represented by the CryoEM map. However, resolution is usually not homogeneously distributed along the map, as it is common to get regions able to reach better resolutions than others.

If we assume that the parameters needed to reconstruct a good CryoEM density map have been identified correctly (such as the CTF, particle alignment...), areas of the molecule that tend to be more static are usually recovered at a larger resolution. In contrast, highly flexible regions are characterized by a motion blur artefact that makes the interpretation of the map more difficult.

Therefore, if we manage to resolve the conformational variability associated with every particle in a dataset, it should be also possible to apply those estimations to “undo” the molecular motions. In this way, all particles in the dataset will virtually represent the same conformational state as if they were completely static, leading to a reconstruction with improved resolution and reduce motion blurring.

Following the previous hypothesis, we developed a new reconstruction algorithm called ZART (Zernike3D-based Algebraic Reconstruction Technique). The main contribution of ZART is its ability to include conformational variability estimations obtained from a Zernike3D analysis to correct for the non-rigid alignment of particles induced by molecular motions.

In addition to the previous characteristic, ZART has also been designed so that it can include a sharpening step during the reconstruction, allowing reaching more meaningful local minima during the reconstruction process. Thanks to the previous sharpening step, ZART has proven to be able to reconstruct maps with improved local resolution compared to other standard software and methods such as CryoSPARC [51], Relion [40], or Fourier-based

reconstruction [54].

ZART is based on the Algebraic Reconstruction Technique [55], which provides a method to solve linear systems of equations. Luckily, the reconstruction process of a map out of a series of its projections can be formulated as finding the solution of a system of linear equations.

Consider a set of experimental CryoEM images arranged in a vector \mathbf{I} , and a matrix A containing the projection matrices needed to recover every image stored in \mathbf{I} from the underlying 3D object they represent \mathbf{x} . Then, we can express the reconstruction problem as a linear system of equation of the form:

$$A\mathbf{x} = \mathbf{I} \quad (7)$$

To solve (7), ZART uses Block-ART [56] algorithm, an iterative algorithm that starts from an initial guess $\mathbf{x}^{(0)}$. A sensible choice could be, for example, a volume filled with zeros, although other choices are also possible (e.g., an initial volume obtained from the particle dataset). By iterating over all the images in the dataset, it is possible to find \mathbf{x} with:

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \lambda_k \sum_i \frac{I_i - \mathbf{a}_i \mathbf{x}^{(k)}}{\|\mathbf{a}_i\|^2} \mathbf{a}_i \quad (8)$$

Where \mathbf{a}_i is a vector with the contribution of each voxel in $\mathbf{x}^{(k)}$ onto pixel i , and i is the corresponding index of the currently selected image (different for every k). The parameter λ_k is a relaxation factor controlling the convergence speed of the ART iterative process (which might take values between 0 and 2). For noisy images, choosing a λ_k that converges too fast (values close to 1) might lead to divergence.

The reason for choosing ART instead of other more standard reconstruction methods in CryoEM such as those based on the Fourier central slice theorem is the simplicity of including geometrical deformations in ART. Thus, iterative process described in (8) can be modified to include the heterogeneity correction coming from the Zernike3D per-particle deformation fields $\mathbf{g}_L(r)$. Before describing how to consider the deformation fields in the ART process, let us discuss the image formation process considered in ZART algorithm.

In ZART, we approximate the density map by a series expansion of basis functions of the form:

$$V(\mathbf{r}) = \sum_j x_j b\left(\frac{\mathbf{r} - \mathbf{r}_j}{\sigma}\right) \quad (9)$$

being \mathbf{r} a position in 3D space, \mathbf{r}_j the basis function position, and σ the basis function scale. In ZART, Gaussian functions have been chosen to be the basis function used for the expansion of the density map. The main reason behind the previous choice is that Gaussian functions are easy to handle, and they can almost achieve partition of unity. In addition, since Gaussians are spherically symmetric, their projection does not depend on the projection direction, which becomes very handy during the ART reconstruction process.

If we refer to the Gaussian basis function as G , we can define the projection of (9) along an arbitrary direction as follows:

$$I(\mathbf{s}_i) = \sum_j x_j G\left(\frac{(\mathbf{s}_i - \mathbf{s}_o) - E\mathbf{r}_j}{\sigma}\right) \quad (10)$$

where \mathbf{s}_i refers to a given pixel position, \mathbf{s}_o refers to an image in-plane shift, and E to the projection matrix needed to be applied to the volume $V(\mathbf{r})$.

From the previous equation, we can easily apply a deformation field correction to drive image I towards a different conformational state:

$$I_{deformed}(\mathbf{s}_i) = \sum_j x_j G\left(\frac{(\mathbf{s}_i - \mathbf{s}_o) - E(\mathbf{r}_j + \mathbf{g}_L(\mathbf{r}_j))}{\sigma}\right) \quad (11)$$

Therefore, in the ART process we can generate $I_{deformed}$ during the backward model step to improve the update of the voxel density values in every iteration and yield a heterogeneity corrected volume.

Lastly, ZART also includes a multiresolution reconstruction mode, which allows to converge faster to more meaningful reconstructions without the need for many ART iterations. In a standard ZART reconstruction, the Gaussian basis functions are regularly distributed along the volume grid and have a fixed scale. However, in multiresolution the scale of the Gaussian to be placed in the grid will vary according to the local resolution estimated from the map being updated. In this way, areas with higher local resolution will be expanded by Gaussians with a smaller radius (allowing them to better represent the features in that region). In contrast, areas with lower local resolution (such as the volume

background) will be approximated by Gaussians with a larger radius, making the low frequency component more prominent. The estimation of the local resolution was done with MonoRes [57].

2.2. Zernike3D basis applications

2.2.1. Application to CryoEM density maps

One of the main advantages of working with deformation fields is the versatility they provide when applied to different data types. In the case of Zernike3D deformation fields, their application is also possible independently of the type of data we are working on with, as long as it lives on a 3D space.

CryoEM density maps are the first 3D structural representation that can be obtained from a processing pipeline. As explained in previous sections in the dissertation, it is even possible to obtain several discrete conformational states through a 3D classification.

Even though 3D classification has the potential to resolve several states, the structural relationships and motions connecting those states cannot be easily inferred. Therefore, conformational analysis arising directly from a set of CryoEM density maps can give us more insights on how different transitions occur.

The application of a Zernike3D deformation field \mathbf{g}_L to a map is almost direct, and can be expressed as:

$$V_{deformed}(\mathbf{r}) = V(\mathbf{r} + \mathbf{g}_L) \quad (12)$$

Since the volume 3D space is discretized into a grid of voxels, the previous evaluation can only be achieved by properly interpolating a density value in the position determined by the deformation field. Commonly, backward interpolation is the preferred choice to perform the previous task. However, we found more useful for the Zernike3D case the forward interpolation approach, which is properly handled to avoid the most common interpolation errors associated with this strategy.

Although (12) provides a way to deform a volume according to the Zernike3D deformation field \mathbf{g}_L , it does not provide a proper way to identify such a field. In general, finding the free

form of a deformation field will require a search over a large 3D space, making the process slow. Instead of finding directly the Zernike3D field, we can take advantage of the basis expansion to reduce the number of parameters to look for. As shown in (1), as we already know where and how to evaluate the Zernike3D basis $Z_{l,n,m}$, the only missing piece needed to define the field is to look for the appropriate values for the Zernike3D coefficients $\alpha_{l,n,m}$, which can be found by solving the following optimization problem:

$$\min_{\mathbf{g}_L} \int |V_1(\mathbf{r}) - V_2(\mathbf{r} + \mathbf{g}_L(\mathbf{r}))| d\mathbf{r} \quad (13)$$

Where V_1 is a map representing the conformational state we want to approximate, V_2 is the map we are going to move based on the estimated deformation field, and \mathbf{g}_L is the Zernike3D deformation field expanded in the Zernike3D basis as shown in (1). The previous problem can be solved using standard optimization methods such as Powell optimization [58].

It should be noted that the solution for equation (13) strongly depends on the initial conformation we are starting from, as well as the type of optimization method applied to find the optimal solution. In the case of Powell optimization, if the initial guess is too far from the ideal solution, the local optimization approach Powell is based on might get trapped in a local minimum not meaningful enough. In general, we found that a sensible initial guess for the Zernike3D coefficients $\alpha_{l,n,m}$ is a vector full of zeros (i.e., no deformation field is applied to the initial state).

Although solving (13) should lead to good enough deformation fields, in some cases where maps are noisy or affected by other artefacts the Powell optimization might find fields able to overfit to those errors. To prevent the optimization process from being trapped in unwanted local minima, we propose to add two extra terms to the previous minimization problem.

The first term is the following:

$$\lambda_1 \int \|\mathbf{g}_L(\mathbf{r})\|^2 d\mathbf{r} \quad (14)$$

The previous term penalizes any solution leading to a deformation field with a large magnitude, which usually leads to unwanted excessive deformation of some molecular

regions. The regularization term λ_1 determines the importance of this term in the overall cost function.

The second term takes the form:

$$\lambda_2 \frac{|\int \{V_1(\mathbf{r}) - V_1(\mathbf{r} + \mathbf{g}_L(\mathbf{r}))\}| d\mathbf{r}}{\int V_1(\mathbf{r}) d\mathbf{r}} \quad (15)$$

It accounts for the penalization in those cases where the deformation field leads to changes in the total mass of the deformed volume. Similarly to the previous case, the parameter λ_2 controls the importance of the regularization in the overall cost function.

2.2.2. Application to structural models

Structural models constitute the next direct application of the Zernike3D deformation fields to explore new conformational states. Similarly to the CryoEM density maps application described in the previous sections, as structural models belong to the 3D space, the process of warping the atomic coordinates does not need any extra preprocessing step.

In addition, structural models are represented by a set of continuous 3D dimensional coordinates, unlike density maps that are discretized into voxels. Therefore, the interpolation step needed to recover a new 3D density map after applying a deformation is not needed for the structural model application.

For a set of 3D atomic coordinates \mathbf{S} and a Zernike3D deformation field \mathbf{g}_L , the deformed set of new atomic coordinates can be obtained as:

$$\mathbf{S}_{deformed} = \mathbf{S} + \mathbf{g}_L \quad (16)$$

We can also define an optimization problem that could be solved to estimate a set of Zernike3D coefficients $\alpha_{l,n,m}$ able to expand the deformation field needed to approximate two different structural models as:

$$\min_{\mathbf{g}_L} \sqrt{\frac{\sum_i^N \mathbf{S}_i^1 - \mathbf{S}_i^2 + \mathbf{g}_{L_i}}{N}} \quad (17)$$

Where N is the total number of atoms in the structural models and i is the current atom

index needed to retrieve the atomic coordinates to be compared. Therefore, (17) provides a way to optimize the Zernike3D coefficients $\alpha_{l,n,m}$ if the structural models have the same number of atoms, whose correspondence must be known beforehand. To solve the previous problem, it is possible to apply any optimization algorithm such as Powell optimization.

The previous assumption is, however, not true in many cases (e.g., if pseudoatomic models are used). For those cases, we propose a modified version of the previous problem where correspondence is initially addressed by a KDTree [59] search, allowing to approximate the reference positions needed to define the deformation field as the mean of the closest neighbors of the current displaced position (the number of neighbors to be considered is determined by the user).

Moreover, the cost function proposed in (17) is modified when the neighbor search approximation is applied to:

$$\min_{\mathbf{g}_L^+, \mathbf{g}_L^-} \sqrt[2]{\frac{\sum_i^{N_2} \bar{\mathbf{S}}_k^1 - \mathbf{S}_i^2 + \mathbf{g}_{L_i}^+}{N_2}} + \sqrt[2]{\frac{\sum_i^{N_1} \bar{\mathbf{S}}_k^2 - \mathbf{S}_i^1 + \mathbf{g}_{L_i}^-}{N_1}} \quad (18)$$

where \mathbf{g}_L^+ and \mathbf{g}_L^- stand for the deformation fields needed to approximate S_1 from S_2 and S_2 from S_1 respectively. By looking for the previous two fields at the same time we are preventing Powell optimization from falling into unwanted local minima arising from the neighbor search.

The previous optimization can be extended with the regularization term in (14) to avoid excessive deformations during the Powell optimization process. In addition to the regularization term, we also add a consistency term that drives the search of the deformation fields in such a way that \mathbf{g}_L^+ and \mathbf{g}_L^- are the inverse of each other [60]. The consistency term can be defined as follows:

$$\begin{aligned} E_{cons} &= E_{cons}^+ + E_{cons}^- \\ &= \frac{\sum_i^{N_1} \left\| \mathbf{S}_i^1 - \left((\mathbf{S}_i^1 + \mathbf{g}_{L_i}^-) + \mathbf{g}_{L_i}^+ \right) \right\|^2}{N_1} \\ &\quad + \frac{\sum_i^{N_2} \left\| \mathbf{S}_i^2 - \left((\mathbf{S}_i^2 + \mathbf{g}_{L_i}^+) + \mathbf{g}_{L_i}^- \right) \right\|^2}{N_2} \end{aligned} \quad (19)$$

The combination of all the previous terms leads to the final form of the cost function to

be optimized in the case of working with two structural models with a different number of atoms (and therefore, with unknown atom correspondences).

2.2.3. Application to CryoEM particles

The last application of the Zernike3D basis is the approximation of conformational states captured at the CryoEM particle level. Unlike density maps and structural models, CryoEM particle sets allow to approximate much richer conformational landscapes due to the large number of particles generated during the image processing workflow (commonly, around 100k to 1M images are usually generated).

However, the structural information captured by the particle is collapsed along the projection direction. Therefore, the estimation of the molecular motions observed along the image plane (which is perpendicular to the projection direction) are well defined, but determining the real 3D motion is an ill posed problem.

In addition, since the Zernike3D basis is defined in 3D, its application to the 2D particle case is not immediate as in the previous applications. Instead, Zernike3D deformation fields need to be defined in the 3D space, and posteriorly used to warp and generate a 2D image that can be compared with the experimental projection. The previous process is only possible if the CryoEM particles has angular information associated, as well as CTF to correct for the microscope defocus and improve the accuracy of the estimated motions.

Similarly to the previous applications, two inputs are needed to approximate a conformational state from a particle: the experimental particle and a reference volume providing a starting equilibrium state to perform the search. During the estimation process, Zernike3D deformation fields will be computed and applied to the reference volume, leading to an approximation of the particle conformational state. Posteriorly, the angular alignment information associated with the particle will be used to project the deformed volume, generating a theoretical projection to be compared to the experimental particle. At this step, it is possible to compute a cost function through the comparison of the experimental and theoretical particles, which will be used to update the Zernike3D deformation fields if needed. A summary of the previous process is provided in Figure 6

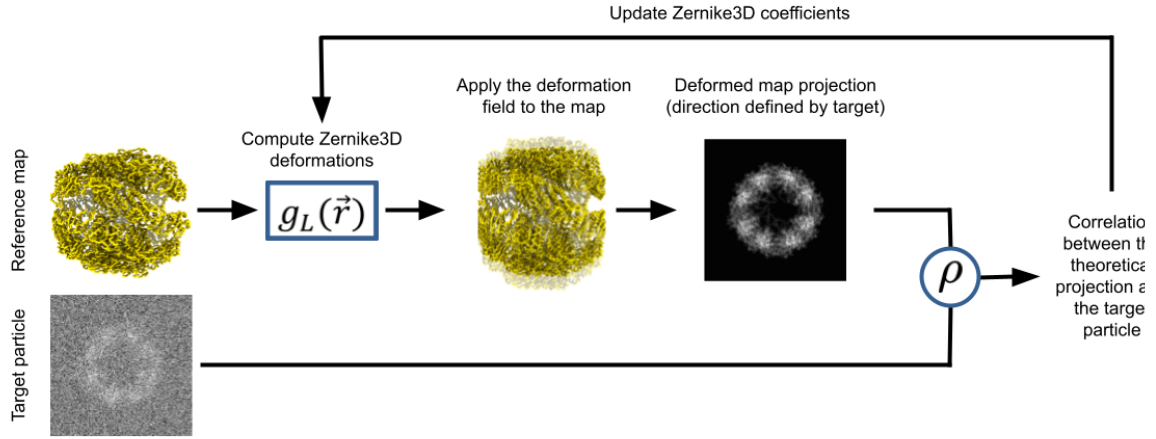


Figure 6: Summary of the application of the Zernike3D basis to find conformational states at particle level [79].

The proposed cost function to be optimized reads as follows:

$$\max_{\mathbf{g}_L} \rho(I, CP_\theta(V(\mathbf{r} + \mathbf{g}_L(\mathbf{r}))) \quad (20)$$

Being I the current experimental image being process, C the CTF associated with the particle, P_θ the projection operator along the direction defined by the angular geometry estimated for the particle, V the reference map, and \mathbf{g}_L the Zernike3D deformation field expanded in term of the Zernike3D coefficients $\alpha_{l,n,m}$ as defined in (1). The comparison between the experimental and simulated image relies on the Pearson correlation coefficient ρ . To optimize the previous problem, Powell optimization is proposed, although other methods could also be applied.

Although the optimization of (20) should be theoretically enough to approximate a given particle structural state, it has an infinite number of solutions equally optimal. As mentioned previously in this section, the approximation of 3D information from an image is an ill posed problem, as the information along the projection direction is missing. Therefore, we can assign any value to the deformation field along the previous projection direction without changing the image content of the theoretical particle generated during the optimization process. A conceivable way to address the previous problem is to regularize the cost function in (20) with an extra term accounting for unwanted deformations, as done for the previous application. Nevertheless, it should be noted that this type of regularization acts in a global manner, which might affect the overall estimation accuracy.

Instead, it is possible to take advantage of the Zernike3D basis properties to completely cancel any deformation along any projection direction. Although this is not completely

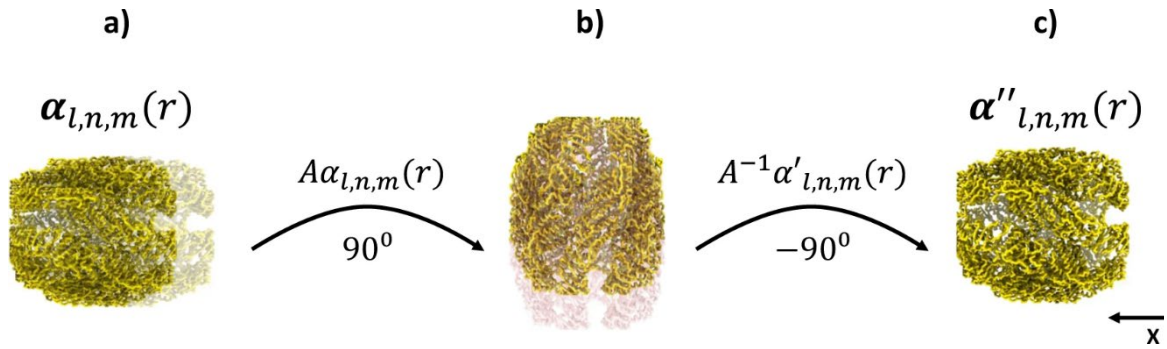


Figure 7: Summary of the correction of the deformation along the projection direction thanks to the closure under rotation property of the Zernike3D basis [79].

realistic, it is the best possible assumption that can be taken when working at the per-particle level (i.e., we do not want to touch the equilibrium state along the projection direction to avoid unwanted guesses).

The previous cancellation can be achieved thanks to the “closure under rotation” property of the basis. As we explained in Section 2.1.2, the Zernike3D deformation fields and Zernike3D coefficients $\alpha_{l,n,m}$ can be easily modified to represent the same field independently of the rotations applied to the reference frame. Thus, we can take advantage of the reference frame rotation to effectively align our deformation field along the projection direction, which translates into a rotation of the coefficients $\alpha_{l,n,m}$. At this point, it is possible to completely cancel the undesired deformation field component.

It should be noted that the closure under rotation property only holds entirely for continuous spaces. However, the reference map is discretized into voxels, meaning that the rotated and corrected fields cannot be applied to the reference map. Before applying the corrected field, it must be rotated back to the map reference frame. After the previous step, unwanted deformation along the projection direction will be effectively corrected. The previous correction step is summarized in Figure 7.

During the optimization of the deformation fields, it is possible to include the previous correction process to prevent the Powell optimizer from falling into unwanted local minima induced by the information collapsed in a particle image.

Due to the high levels of noisy CryoEM particles are buried in, adding an extra regularization to keep the magnitude of the Zernike3D deformation fields low is a good practice to prevent overfitting to noise. To that end, we can extend the cost function equation (20) with the deformation penalization term introduced in equation (14).

Once the optimization of the Zernike3D deformation fields has finished for all the particles in a dataset, a set of Zernike3D coefficients $\alpha_{l,n,m}$ will be obtained, representing the estimated per-particle conformational states. At this point it is a frequent practice to visualize the coefficient space, as it is expected to be a good experimental estimation to the conformational landscape the macromolecule under study has explored. However, the Zernike3D coefficients will have a large dimensionality value (e.g., for the default maximum degree values $N = 3$ and $L = 2$, every vector $\alpha_{l,n,m}$ will end up having a total of 39 coefficients). Therefore, a dimensionality reduction technique such as Principal Components Analysis (PCA) [61] or Uniform Manifold Approximation and Projection (UMAP) [62] needs to be applied to get a meaningful and representable conformational space.

The reduced conformational spaces can be then explored, allowing the identification of the different states a molecule may be visiting. In addition, due to the fast application of the Zernike3D deformation fields, it is possible to generate in real time any conformational state from the landscape as a CryoEM density map or structural model, thus simplifying the understanding of the landscape features.

2.3. Results and discusión

2.3.1. Zernike3D application to density maps

As we discussed in previous sections, the application of the Zernike3D basis to CryoEM density maps is almost straightforward. If we consider a set of maps instead, it is possible to take advantage of the continuous heterogeneity measurement extracted from every map pair in the set to infer heterogeneity information that would be hardly extracted from a visual inspection.

The Zernike3D deformation fields computed from a pair of maps in a different conformational state can be further reduced into a single value, providing a way to compare how different the two conformational states represented by the maps are. For example, it would be possible to compute the root mean square distance of the two maps from the deformation field, or a correlation error between the deformed and reference conformational states.

If instead of a single pair of maps we compared all the maps in a set in a pairwise fashion,

it is possible to arrange the previous comparison metrics in a matrix form. The previous distance matrix is, therefore, a good summary of how different the conformational states explored by a macromolecule are.

However, the matrix representation is usually hard to interpret. Instead, it is possible to apply Multidimensional Scaling (MDS) [63] to recover a point cloud in 2D or 3D whose distances resemble the ones measured from the CryoEM maps and Zernike3D fields. We call the previous low dimensional representation a structure map.

Structure mappings are useful to better understand the possible sequence of conformational states that have been extracted from a set of volumes. For example, a structure mapping would simplify the task of identifying molecular transitions or other structural characteristics captured by the maps.

As an example of the structure mapping capabilities, different volume datasets were analyzed. The first example discussed in this dissertation consist of a synthetic dataset representing the opening and closing of a CCT complex [64]. To that end, 30 structural models were simulated from the transition using an adaptive anisotropic network model [65] implemented in ProDy [66]. The structural models were converted back to electron densities into Coulomb potential maps using the electron atomic scattering factors (EASFs) [67].

The thirty generated maps were posteriorly subjected to the Zernike3D structure map analysis. The results obtained are shown in Figure 8.

As discussed before, once the Zernike3D deformation fields for a pair of volumes has been computed it is possible to compute many comparison measurements. In our case, two different metrics were used:

- **Deformation distance:** Computed as the root mean square value of the deformation field computed from a pair of maps.
- **Correlation distance:** Computed as the Pearson correlation coefficients between the map warped with the deformation field computed for the pair and the reference conformation.

In general, the two previous metrics might lead to slightly different structure mappings showing distinctive features. However, we are lacking a criterion to determine which of the two mappings leads to a more meaningful order of the different states. Therefore, a

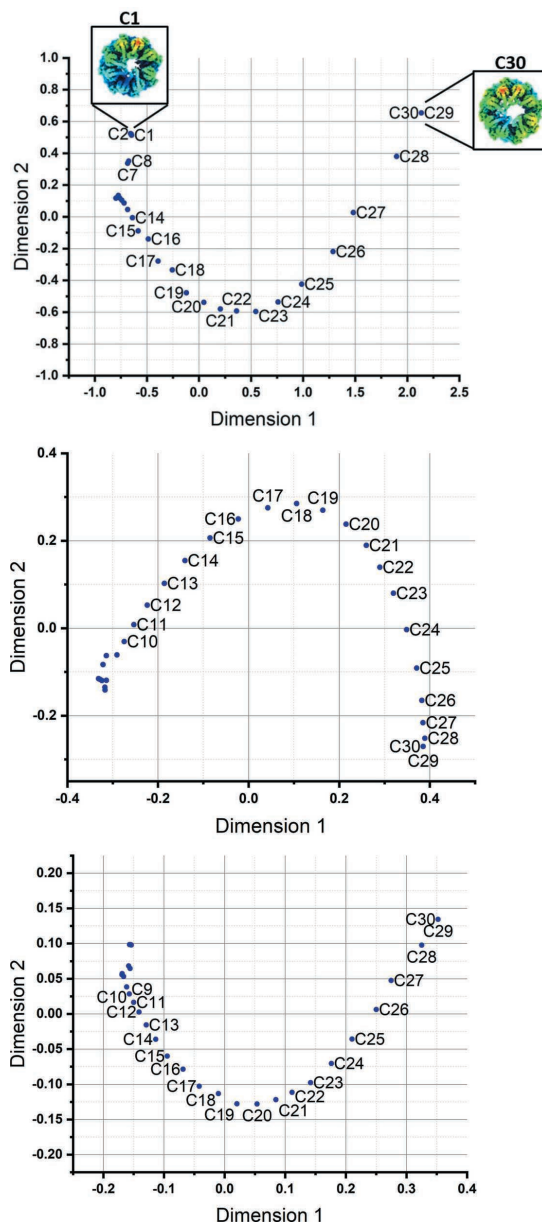


Figure 8: Structure mappings obtained after analyzing a set of 30 CryoEM maps simulated from the opening and closing transition of a CCT complex [64] with the Zernike3D approach. The top mapping shows the results obtained with the deformation distance, middle shows the correlation distance results, and bottom the consensus of the previous two mappings [52].

structure mapping consensus was developed.

The approach followed by the mapping consensus is to find the linear combination of the different mappings obtained in such a way that the entropy of the mapping coming from the linear combination is minimized. The minimum entropy criterium taken in the consensus step enforces the presence of transitions or other structural relationships in the mapping, as this is the kind of information that could be expected from a structure mapping.

As shown in Figure 8 deformation distance, correlation distance, and consensus structure

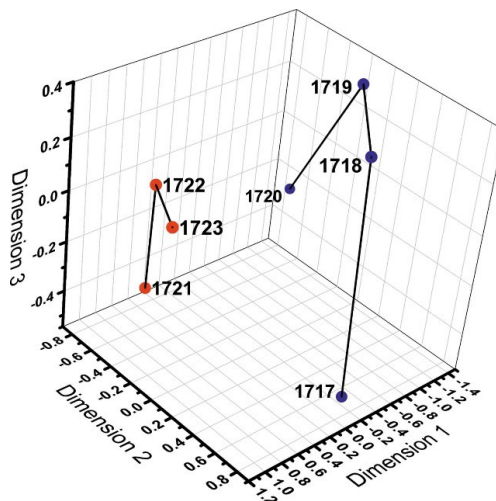


Figure 9: Structure mapping obtained after applying the Zernike3D analysis to a set of 7 human mitochondrial ribosome maps [68]. The structure mapping reveals two different structural transitions captured by the CryoEM density maps [52].

mappings obtained from the synthesized CCT conformations are similar. However, the correlation distance mapping is more sensitive to the presence of small structural variations for this case, providing a better ordering of the maps ranging from 14 to 1. As it can be seen in the bottom image, the consensus is able to produce a better identification of those states in the correlation mapping, giving more weight to it the linear combination. Overall, the three mappings show the expected transition recovered directly from the CryoEM maps, allowing to better identify the structural relationships among the simulated states.

When working with experimental data, 3D classification is usually not able to reconstruct as many good enough conformational states as in the synthetic experiment described previously. Still, the computation of structure mappings might become handy to better identify the molecular motions behind the classified maps.

In order to show the results that could be expected from a structure mapping analysis with experimental data, a set of seven human mitochondrial ribosome maps [68] were retrieved from the EMDB database. It should be noted that experimental maps might have different resolutions, which might lead to an artefactual ordering in the structure map. Thus, it is always convenient to filter the CryoEM maps to be analyzed to the same resolution (i.e., the smallest resolution present in the dataset) before forwarding them into the analysis. The results obtained from the seven experimental maps is shown in Figure 9. The structure mapping reveals the presence of two different structural transitions captured by the CryoEM maps, which would have been hardly identified by a visual inspection of the different

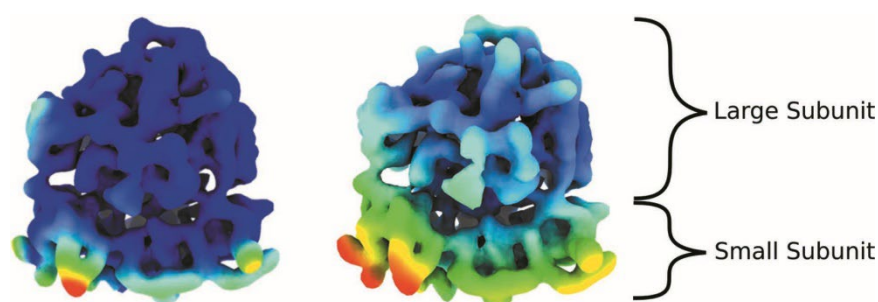


Figure 10: Strain (left) and rotation (right) decomposition obtained from the Zernike3D deformation field computed from the EMDB 1720 and 1723 maps [52].

conformational states.

Apart from analyzing the structural relationships represented by a set of CryoEM maps, the Zernike3D deformation fields can also be further analyzed to extract the strain and rotational forces driving a given molecular motion, as described in [69]. An example of the strain/rotation decomposition obtained from the human mitochondrial ribosome maps 1720 and 1723 is provided in Figure 10.

The strain rotation decomposition is translated into a set of values that can be used to color the original CryoEM maps. Based on the colormap from Figure 10, higher strain/rotational forces are colored in red, while lower forces are colored in blue.

Lastly, the Zernike3D deformation fields computed from a pair of maps can also be applied to a structural model, allowing to easily reproduce the conformational state at atomic level. An example of the application of the Zernike3D fields to structural models is provided in Figure 11.

It should be noted that the Zernike3D deformation field estimation is purely based on geometrical considerations arising from the structural information captured by the CryoEM maps. Therefore, the application of the Zernike3D fields to a structural model might lead to a representation with no stereochemical constraints imposed, which should be posteriorly addressed to correct the Zernike3D model.

2.3.2. Zernike3D application to particles

The extraction of conformational states directly from 2D CryoEM particle images is one of the main and most interesting contributions flexibility algorithms are making to the field. The large amount of structural data captured in a particle dataset allows flexibility algorithms to

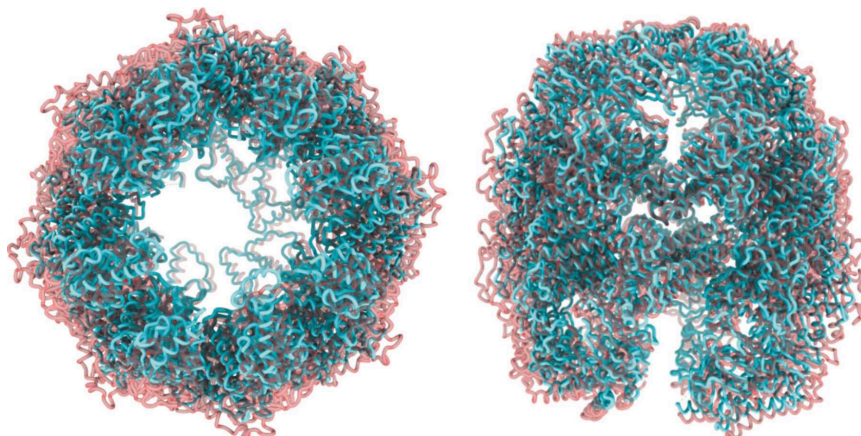


Figure 11: Application of the Zernike3D fields computed from the open to the close CCT maps to the open CCT structural model [52].

approximate the experimental conformational landscape explore by a macromolecule, in contrast to the classical classification approach limited to a reduced number of states.

Therefore, the conformational estimations performed directly at particle level leads in last instance to a mapping much richer and complete than the structure mappings described in the previous sections. The new mappings do not only provide a much better understanding of the structural relationships relating different states, but they might also reveal new structural conformations hidden in the particle data due to their low representation.

As explained in Section 2.2.3, the application of the Zernike3D basis to extract per-particle deformation fields (and therefore, conformational states) is also possible, as long as the projection of the 3D space is handled appropriately. However, structure mappings are no longer a viable option, due to the large dimensions of the distance matrices that would be obtained if all particles are considered.

Instead, it is easier to analyze directly the Zernike3D coefficient space composed by the estimations obtained from the particles. Nevertheless, the high dimensionality of the Zernike3D coefficient vectors makes the interpretation of the estimated space a challenge. Thus, a dimensionality reduction step is advisable to visualize and simplify the interpretation of the Zernike3D conformational landscape.

To assess the performance of the Zernike3D basis on the approximation of per-particle states and landscapes, several experimental datasets were selected and processed. The first dataset discussed in this Section is the EMPIAR 10028 dataset [70], which has become a popular choice for the evaluation of flexibility algorithms.



Figure 12: Zernike3D conformational landscape obtained from the EMPIAR 10028 particles [70] and reduced with UMAP [62]. The landscape reveals two well differentiated regions, marked with an orange and white dot. These two regions correspond to the unrotated and rotated states of the small ribosomal subunit present in the dataset. The colormap used to represent the landscape shows the average deformation magnitude that needs to be applied to the reference map to achieve a new conformational state [79].

The EMPIAR 10028 dataset consists of a series of CryoEM particles of a *P. falciparum* 80 S ribosome bound to emetine. When analyzed with a 3D classification algorithm, two well differentiated conformational states can be identified, which capture a rotational motion of the small ribosomal subunit. The presence of a significant structural change in combination with the large size of the ribosome makes this dataset a good validation choice.

Before subjecting the particles to the Zernike3D analysis, the original dataset downloaded from EMPIAR was further processed and cleaned inside Scipion [10]. In addition, the angular information extracted from different CryoSPARC refinements [51] was consensuated to achieve a more stable dataset.

The consensuated particles were posteriorly analyzed with the Zernike3D algorithm, yielding the Zernike3D landscape. Figure 12 shows the obtained Zernike3D coefficient space reduced with UMAP [62].

As it can be observed from the landscape, two regions are easily identified, which correspond to the rotation of the small ribosomal subunit captured in the dataset. The Zernike3D landscape has been colored according to the average deformation magnitude that needs to be applied to the reference map to reach a different state. In this way, purple colors represent states that are closer to the reference map, while yellow colors are assigned to those states that represent a larger motion compared to the reference state.

Similarly to the volume application, it is possible to apply any Zernike3D coefficient vector

$\alpha_{l,n,m}$ extracted from the previous landscape to the corresponding reference map or its structural model. Figure 13 provides an example of the small ribosomal subunit rotation recovered through the Zernike3D fields extracted from the orange dot in Figure 12.

As shown in the figure, the application of the Zernike3D fields extracted from the landscape can recover the expected conformational change. When compared to a CryoSPARC refinement (blue map in b) panel), it can be observed that there is an agreement between the two states. However, the application of the Zernike3D fields yield a map at the same resolution as the chosen reference state, while the resolution in a refinement is usually more limited due to the lower number of particles used during the reconstruction. Moreover, the Zernike3D fields can also be applied to a structural model as shown in c), allowing to recover a directly the atomic coordinates from the selected state without the need for tracing in a low-resolution map. It should be noted that the Zernike3D fields are estimated based on purely geometrical considerations. Therefore, stereochemical constrains should be posteriorly applied to the recover structure to improve their quality.

The second dataset analyzed by the Zernike3D algorithm was the EMPIAR 10180 dataset [71], representing a catalytic spliceosome exhibiting a large degree of flexibility. In addition, the dataset includes a mixture of compositional and continuous heterogeneity, which might decrease the performance of the Zernike3D algorithm. Therefore, compositional heterogeneity was cleaned beforehand inside Scipion, leading to a cleaned dataset where continuous heterogeneity was more prominent. Posteriorly, the cleaned particle dataset was subjected to an angular consensus analysis to improve the stability of the particle angular assignation.

The flexibility analysis conducted was similar to the workflow followed with the EMPIAR 10028 dataset. The consensuated particles were analyzed through the Zernike3D basis, generating a set of Zernike3D coefficients that approximate the conformational landscape captured by the dataset. The Zernike3D landscape was further reduced with UMAP to simplify its visualization, as shown in Figure 14a.

One of the easiest ways to analyze a conformational landscape is to draw several conformations from it that are uniformly distributed. To that end, it is possible to use a KMeans clustering of the original Zernike3D space, and use the representative coefficients associated with each cluster to recover a set of uniformly distributed states. Figure 14b

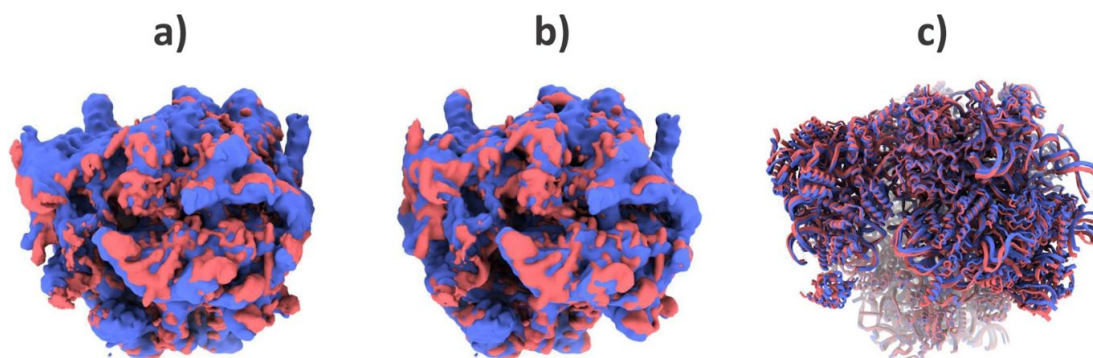


Figure 13: Example of the application of the Zernike3D fields (extracted from the orange dot in Figure 12) to recover the rotated ribosomal state present in the EMPIAR 10028 dataset. *a)* shows the comparison of the Zernike3D rotated state (red) and the reference state (blue). *b)* shows the comparison of the Zernike3D rotated state (red) and the rotated state obtained from a CryoSPARC [51] homogeneous refinement (blue). The comparison shows that the Zernike3D fields can effectively recover the expected conformational state, but without the intrinsic loss in resolution that might induce the refinement. *c)* shows the application of the Zernike3D fields to recover the rotated state (red) at the level of structural models, compared to the structural model coming to the reference map (blue). Thus, versatility of the Zernike3D basis allows to recover directly any state in the landscape at atomic level [79].

shows the resulting representative states extracted from a KMeans analysis with five clusters.

The clustering reveals some of the main motions captured in the dataset, mostly associated with the top and bottom regions of the spliceosome. Figure 14c shows the application of the previous five representative to the structural model deposited with the EMPIAR 10180 dataset. The application of the Zernike3D deformation fields to a structural model simplifies the identification of the relevant and local motions described by the different states obtained from the conformational landscape. We provide in Figure 15 a closer comparison to two of the states shown in Figure 14c.

The example of application of the Zernike3D algorithm to an experimental particle dataset focuses on a harder specimen due to its smaller size and SNR: the SARS-CoV-2 spike. The dataset proposed for this experiment comprises the EMPIAR 10514 and 10516 entries [72].

The two previous datasets were analyzed through a standard 3D classification inside Scipion [10], revealing a slight opening of the Receptor Binding Domain (RBD) in up conformation. Therefore, the analysis of the dataset is a good example to assess the performance of the Zernike3D method when faced with the identification of smaller motions and more localized motions.

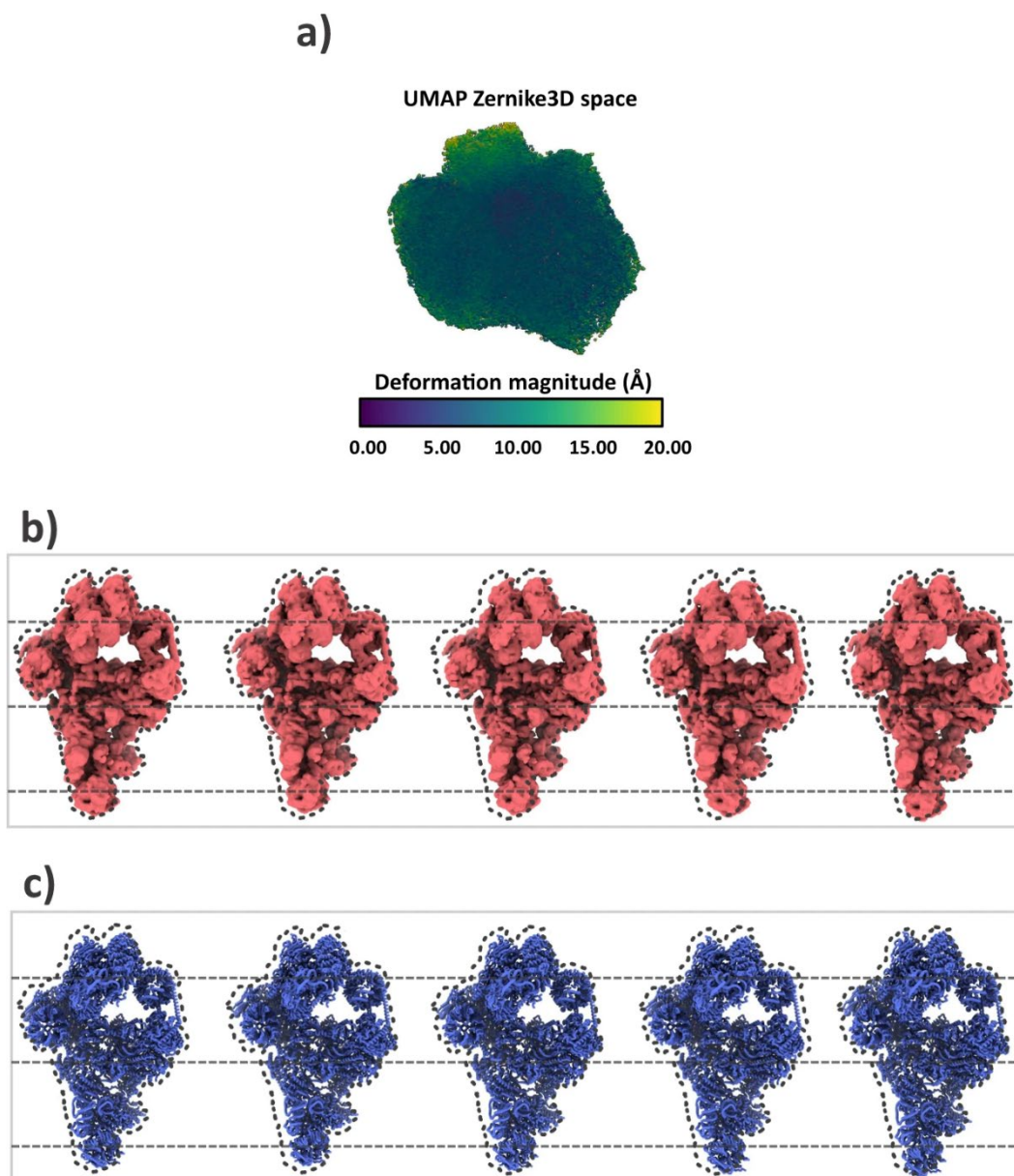


Figure 14: Zernike3D analysis of the EMPIAR 10180 dataset [71]. a) show the Zernike3D conformational landscape recover from the particles and reduced with UMAP [62]. The colormap represents the average deformation of the deformation fields computed for each particle in the dataset. b) shows a set of 5 maps generated through the application of the Zernike3D fields to the reference map. The 5 conformations were uniformly sampled from the Zernike3D landscape with KMeans. c) shows the previous 5 sampled conformational at atomic level, obtained after applying the Zernike3D fields to the structural model deposited with the EMPIAR 10180 dataset [79].

The conformational landscape obtained from the Zernike3D analysis is provided in Figure 16a. To simplify the understanding of the landscape, we can translate the original 3D classified maps obtained from the dataset into the Zernike3D space, allowing to map at the same time the results of the discrete and continuous flexibility analysis. The combined landscape is show in Figure 16b. From the combined landscape, it can be easily seen the advantage of the continuous heterogeneity methods, as they are able to capture a significant amount of structural information that is lost during the classification process.

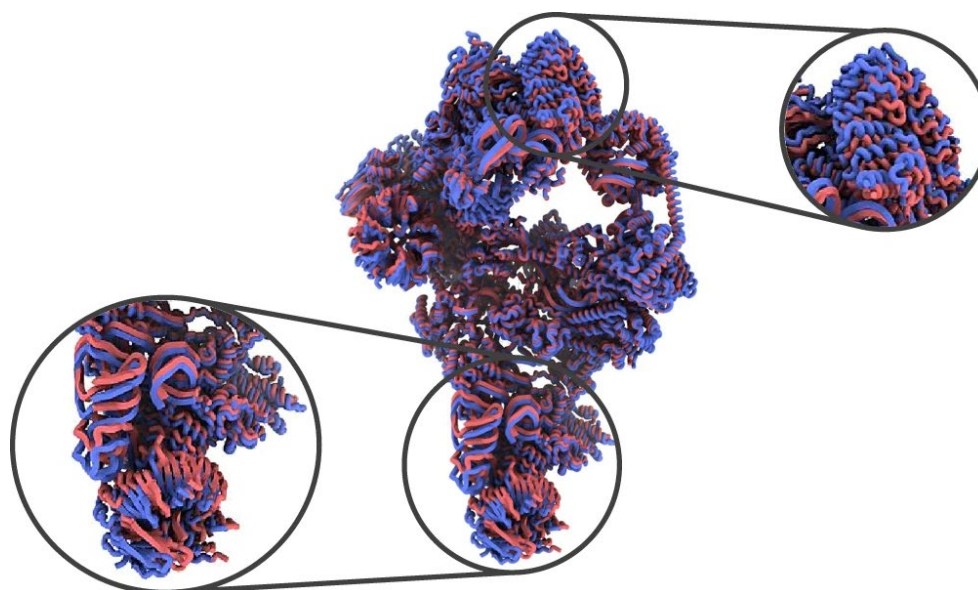


Figure 15: Comparison of two of the structural models recovered after applying the Zernike3D fields obtained from two of the five states uniformly sampled from the estimated Zernike3D conformational landscape in Figure 14a. The possibility of generating states directly at atomic level simplifies the analysis and understanding of the molecular motions captured by the conformational landscape [79].

In addition, the combined landscape reveals a new region in the landscape that was not captured by the discrete heterogeneity analysis, due to the small number of particles defining that conformational landscape region. A thorough analysis of the region of interest shows a new conformational state that is slightly closing, opposite to the changes obtained with the classical discrete analysis as shown in Figure 17. The previous analysis shows the potential of the Zernike3D method in the identification of less stable states that may remain hidden to classical analysis.

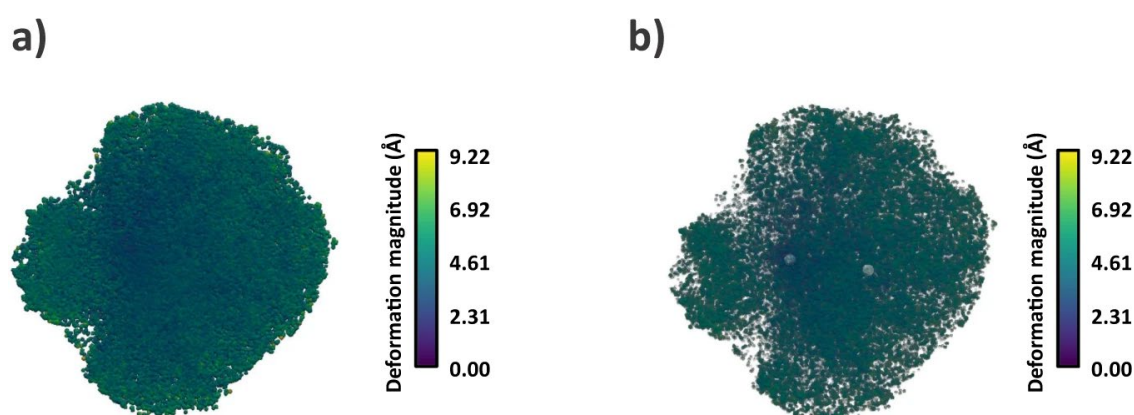


Figure 16: Zernike3D conformational landscapes obtained from the analysis of the EMPIAR 10514 and 10516 datasets [72] and reduced with UMAP [62]. a) shows the estimated landscapes, colored according to the average deformation of the Zernike3D fields obtained for every particle. b) shows the combined landscape obtained after translating the original discrete classification results to the Zernike3D space, represented by the white dots. The combination of the continuous and discrete information simplifies the understanding of the conformational landscape and reveals a new structural region that was not possible to be captured by the classical discrete heterogeneity analysis [79].

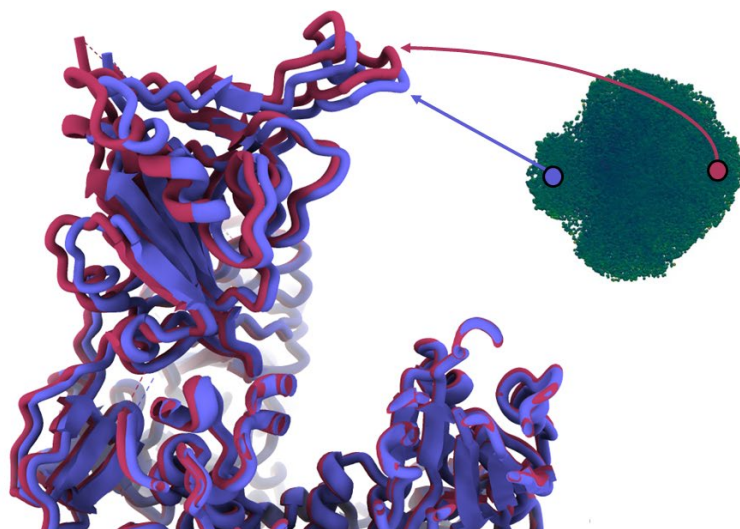


Figure 17: Structural models extracted from the Zernike3D regions shown in the landscape included in the figure, corresponding to the two extreme points of the main transition defined by the discrete classification as shown in Figure 16b. The structural models show the capability of the Zernike3D method in the identification of small protein motions.

2.3.3. ZART reconstruction

The application of the Zernike3D basis in the exploration of conformational landscapes and states is the standard approach in the analysis of conformational variability from CryoEM data. However, it is also possible to consider the estimated Zernike3D deformation fields a non-rigid alignment needed to correct for every particle variability and reduce motion blurring in CryoEM density maps, as discussed in Section 2.1.5.

The development of ZART follows the previous reasoning, as it considers the conformational landscape to correct for the non-rigid alignments associated with every particle during the reconstruction process. The previous correction combined with the sharpening capabilities of the ZART reconstruction algorithm allows to reach more meaningful reconstruction with improved resolution.

Before assessing the performance of the heterogeneity correction, ZART was compared with a standard Fourier reconstruction to discuss the effect of the intrinsic sharpening ZART applies during the reconstruction process. To that end, a synthetic dataset was generated from the SARS-CoV-2 PDB 6VSB [73]. The structural model was converted to a Coulomb potential map and projected to generate a set of 18k ideal CryoEM particles (i.e., with no

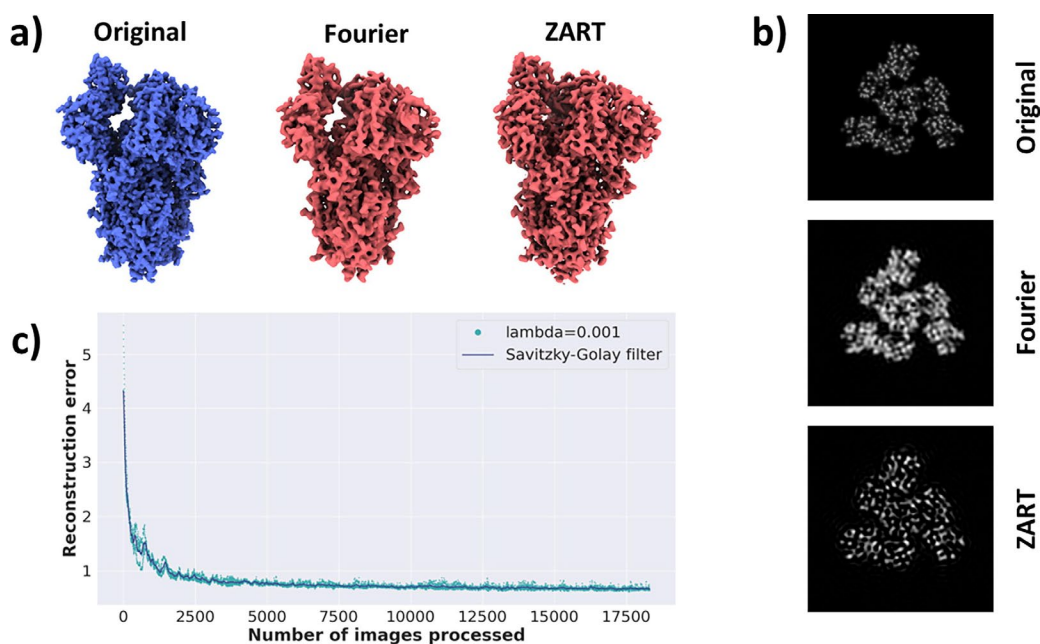


Figure 18: Comparison of ZART and Fourier reconstructions obtained from the synthetic dataset generated from the SARS-CoV-2 PDB 6VSB [73]. a) shows the comparison of the original map simulated from the structural model, Fourier, and ZART. A closer look to the maps suggests a better resolvability of the spike features in the case of ZART, thanks to its sharpening capabilities. b) shows a comparison of the three maps at the level of 2D slices. Similarly to a), ZART enhances the structural features compared to Fourier. c) shows the convergence curve for ZART after iterating through several images [55].

noise or CTF) with sampling rate $2.40 \text{ \AA}/\text{px}$. The particle dataset was then reconstructed with ZART and Fourier reconstruction methods. The results obtained from the reconstructions are summarized in Figure 18. As it can be seen from Figure 18a and Figure 18b, ZART is able to significantly enhance the structural features of the spike compared to Fourier thanks to the sharpening it applies during the reconstruction, reaching a more meaningful CryoEM map. Together with the comparison, the convergence curve of ZART is also provided in Figure 18c, which shows a low enough convergence error after five thousand images have been analyzed.

The previous ZART reconstructions were obtained following the “standard” reconstruction approach implemented in ZART. However, a multi-resolution approach was also implemented as described in Section 2.1.5. Therefore, the next experiment focuses on the analysis of the performance of the method when faced against an experimental dataset, including a comparison of the standard and multi-resolution approaches. The dataset to be analyzed corresponds to the EMPIAR 10391 [74], which was processed inside Scipion [10] to extract a set of 35k particles with angular and CTF information estimated with Relion [40].

The particle dataset was posteriorly reconstructed with ZART following the standard and

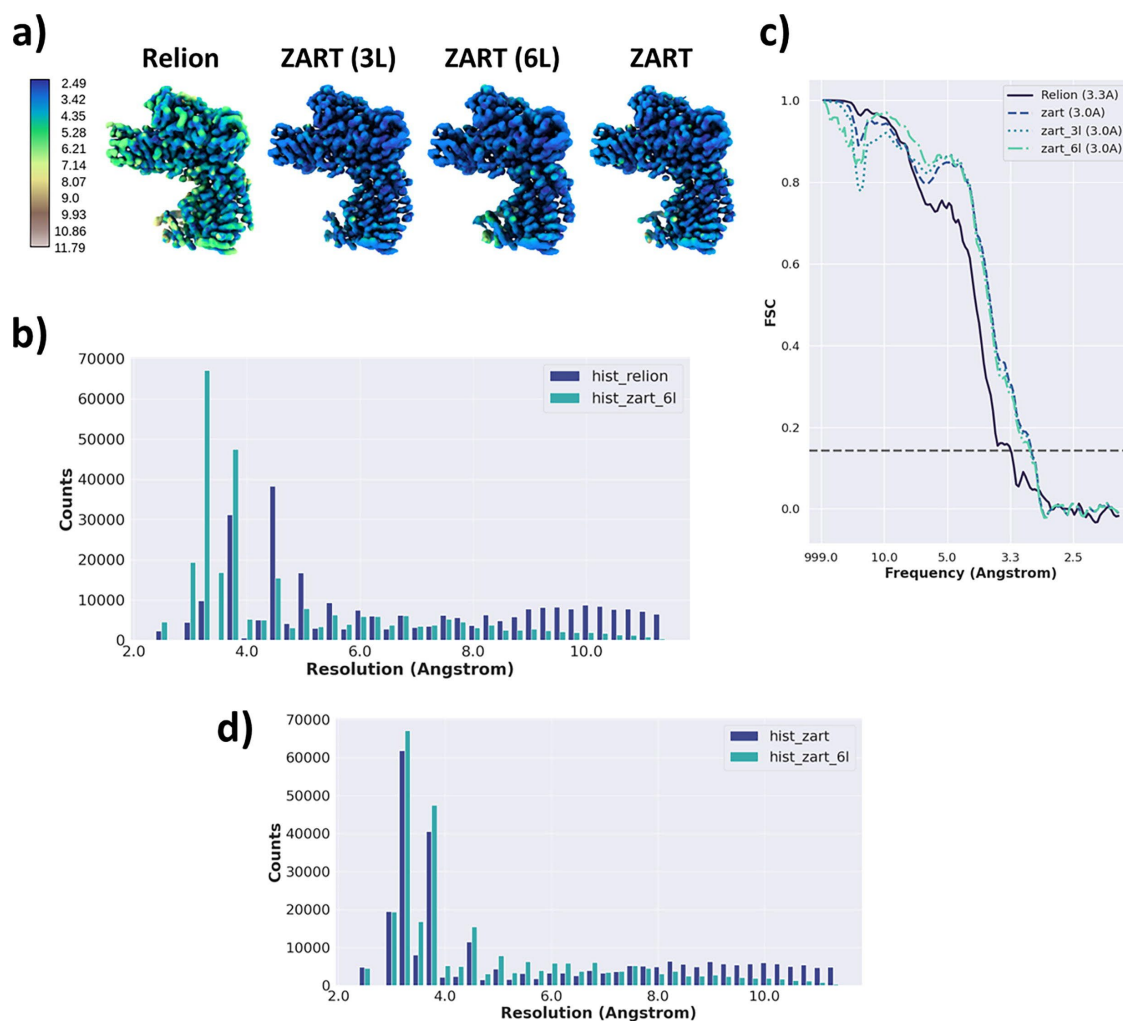


Figure 19: ZART reconstructions obtained from the EMPIAR 10391 dataset [74]. a) shows the reconstructed maps obtained with Relion [40] and ZART multi-resolution with 3 levels, multi-resolution with 6 levels, and standard) sharpened with DeepEMHancer [75]. b) shows a comparison of the local resolution histograms for ZART (multi-resolution with 6 levels) and Relion computed with MonoRes [57]. c) shows the FSC curves obtained from the Relion and ZART maps. d) shows a comparison of the local resolution histograms obtained from the standard and multi-resolution (with 6 levels) ZART reconstructions [55].

multi-resolution approaches, and further analyzed to determine the quality of the maps. The results obtained are summarized in Figure 19. Figure 19a shows the CryoEM reconstructed maps obtained with Relion and ZART (multi-resolution with three levels, multi-resolution with six levels, and standard). The maps are colored according to their local resolution computed with MonoRes [57] and sharpened with DeepEMHancer [75] in wide target mode to reduce the noise. The maps suggest an improvement in resolution and feature resolvability in the case of ZART, similar to the results obtained from the 6VSB SARS-CoV-2 spike.

To compare the ZART (multi-resolution with six levels) more quantitatively and Relion reconstructions, the local resolution histograms were further analyzed as shown in Figure

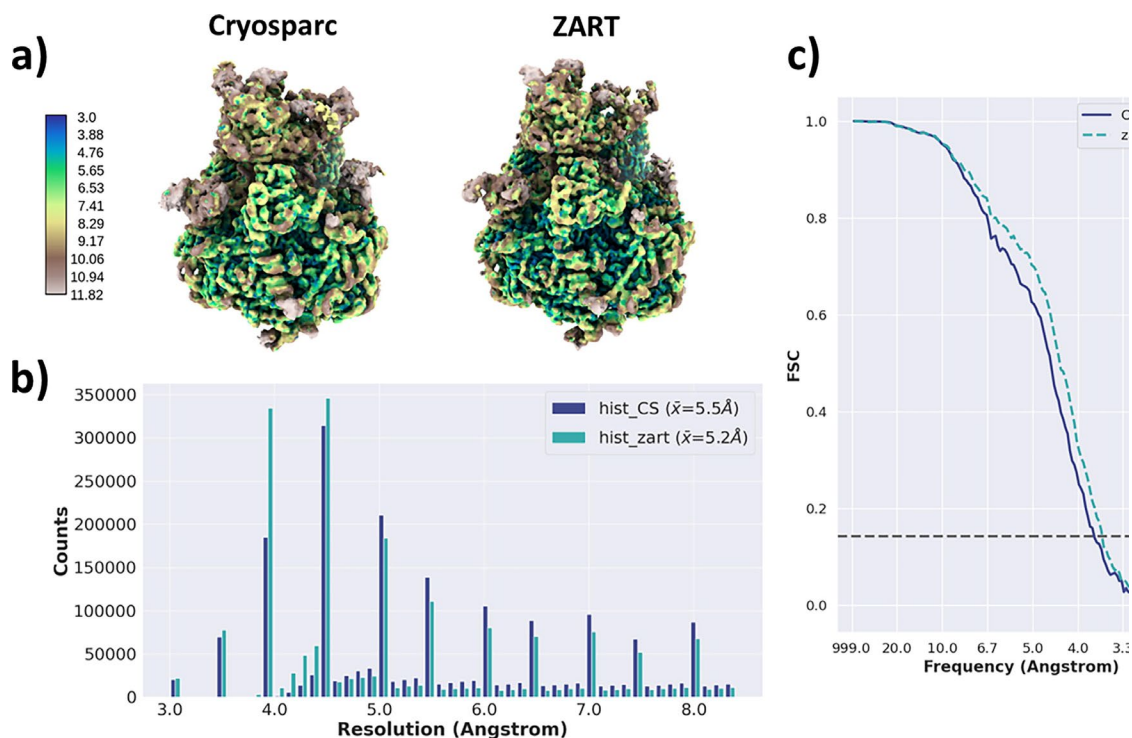


Figure 20: Comparison of CryoSPARC [51] and ZART (with heterogeneity correction) reconstructions obtained from the EMPIAR 10028 dataset [70]. a) shows a comparison of the reconstructed CryoEM maps colored according to their local resolution estimated with MonoRes [57] and sharpened with DeepEMHancer [75]. b) shows the comparison of the local resolution histograms from MonoRes computed from the previous two maps. c) shows the FSC curves obtained from the two reconstructions [55].

19b. The resolution histograms also suggest a significant displacement of ZART voxels towards the high-resolution regime, improving the average local resolution from 7.0 Å to 4.0 Å. In addition, the FSC curves computed from the reconstructed half maps are provided in Figure 19c. Following the trend extracted from the local resolution measurements, ZART maps improved the global resolution of the map around 0.3Å when compared to the Relion reconstruction, and 0.5Å when compared to the original map published with the EMPIAR 10391 dataset.

Together with the quality analysis of Relion and ZART reconstructions, a comparison of the local resolution histograms obtained from the standard and multi-resolution (with six levels) ZART reconstructions is also provided in Figure 19d. Although similar, the multi-resolution approach introduces a slight increment in the local resolution thanks to the consideration of the local resolution estimations and variable Gaussian widths during the reconstruction process.

The last experiment focuses on the analysis of ZART when flexibility correction is considered. To simplify the understanding and discussion of the flexibility correction effect

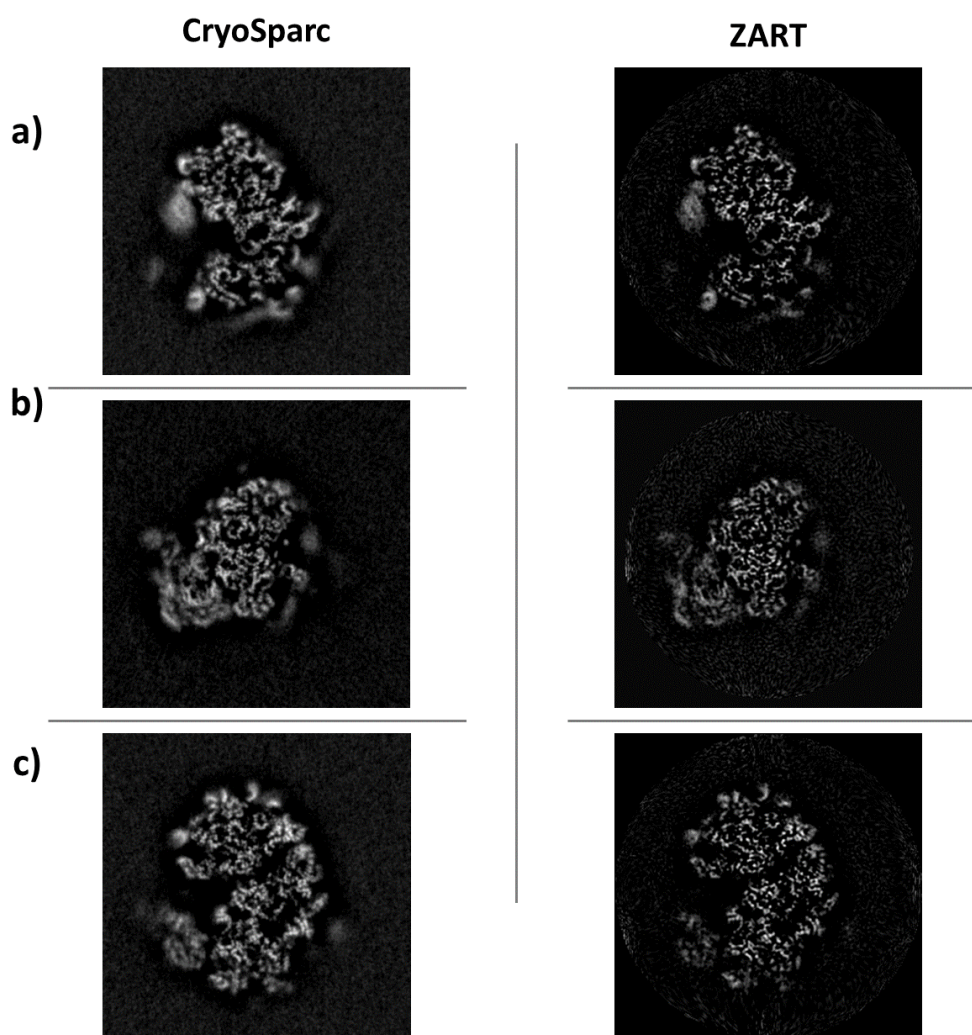


Figure 21: Example of some 2D slices extracted from the CryoSPARC [51] and ZART (with heterogeneity correction) reconstructions from the EMPIAR 10028 dataset [70]. ZART slices provided a better representation of the ribosomal structural features, which is accentuated in the small ribosomal subunit thanks to the heterogeneity correction. For visualization purposes, the negative values of CryoSPARC and ZART slices has been removed [55].

in the reconstruction, the multi-reconstruction approach was not considered for this case. However, it should be noted that it is possible to correct for the conformational variability in both, the standard and multi-resolution approaches.

The EMPIAR 10028 dataset [70] previously analyzed in Section 2.3.2 was also subjected to the ZART analysis. In this case, the estimated Zernike3D conformational landscape provided in Figure 12 was forwarded to the ZART method together with the CryoEM particles to perform a heterogeneity corrected reconstruction. The original map reconstructed from the previous particles in CryoSPARC [51] was also recovered for comparison purposes. The results obtained from the previous analysis are summarized in Figure 20. An initial inspection of the CryoEM maps from Figure 20a shows similar structural features in the core of the large

ribosomal subunit, which is only subjected to residual motions. Instead, the small ribosomal subunit exhibits a better enhancement of its structural features, thanks to the correction of its rotation (as discussed in Figure 13) and other residual motions, which is significant enough to induce a larger degree of motion blurring in the case of CryoSPARC. Both maps in Figure 20a are colored according to their local resolution computed with MonoRes [57] and were sharpened with DeepEMHancer [75] in wide target mode to reduce the noise.

Figure 20b provides a more quantitative comparison of the two maps through the local resolution histograms extracted from MonoRes. Similarly to the EMPIAR 10391 case, ZART pushes a larger number of voxels towards the high-resolution regime, improving the average local resolution from 5.5 Å to 5.2 Å. In the case of the FSC curves show in Figure 20c, it can be observed a similar global resolution, with a slight improvement in the case of ZART of around 0.1Å.

A comparison of CryoSPARC and ZART reconstructions at the level of 2D slices is also provided in Figure 21. The visualization in both cases has been modified to exclude the negative values in both maps. The comparison of the slices follows a similar trend to the one already observed at the level of the maps and resolution histograms, resulting in an enhancement of the structural features of the EMPIAR 10028 ribosome, which is more significant in the small ribosomal subunit thanks to the flexibility correction of ZART.

Conclusions and future work

The main objective of this dissertation, structured as a compendium of articles, was to introduce a new method to analyze continuous conformational variability from various CryoEM datatypes (maps, structural models, and particles) called the Zernike3D algorithm.

The new method is based on the Zernike3D basis, a composition of the Zernike polynomials and spherical harmonics able to expand any function contained in a sphere. Thus, the basis expansion relies on the Zernike3D basis components $Z_{l,n,m}$ and the Zernike3D coefficients $\alpha_{l,n,m}$, that determine the contribution of each component to the expansion. The basis is then used to expand a deformation field describing the transition between two conformational states, which can be computed through the minimization of an appropriate cost function having the Zernike3D coefficients $\alpha_{l,n,m}$ as unknown. Although our basis choice is not unique, we found it to have interesting properties that can be exploited to overcome several challenges during the search of the Zernike3D deformation fields.

The applicability and performance of the basis was assessed on several CryoEM datasets (both, experimental and synthetic), which has proven to be versatile even when subjected to different CryoEM datatypes. The ability to apply the Zernike3D basis to CryoEM maps and particles allows to extract and analyze the estimated heterogeneity information in diverse ways and opens the possibility to the design of combined analysis where discrete and continuous flexibility could be understood together. Moreover, the application of the estimated Zernike3D fields from maps or particles directly to structural models provides a new mechanism to understand conformational landscape at atomic level, as well as analyzing the approximated states with atomic-oriented tools such as Normal Mode Analysis or Molecular Dynamics.

Apart from the identification and exploration of conformational landscapes, the Zernike3D estimations can also be considered as a non-rigid alignment able to partially

reduce blurring artifacts appearing during the CryoEM reconstruction process, leading to a decrease in the motion induce artefacts present in a map. The previous approach was implemented in a new reconstruction method called ZART, which combines reconstruction, sharpening, and heterogeneity correction to yield more meaningful reconstructions with enhanced structural features. The implementation of ZART also includes a multi-resolution reconstruction approach, able to analyze the local resolution of the maps being reconstructed to give a different resolution weight to different areas in the map. In this way, ZART can automatically focus on improving those map regions where the protein is found, further improving the accuracy and convergence speed of the method.

3.1. Future work

In this section we identify the main steps that we would like to follow to improve the Zernike3D and ZART methods described along the dissertation:

- The classical optimization schemes proposed for the search of the Zernike3D coefficients $\alpha_{l,n,m}$ could be improved if a neural network is employed. Using a neural network should provide a much faster search of the coefficients, as well as finding more meaningful conformations thanks to the batch optimization, which should be able to partially recover better the information along the projection direction of a particle by looking at the information of other views.
- ZART is currently implemented on CPU with a multithreading strategy, which improves the performance of the algorithm. However, it suffers from a low performance when correcting for the heterogeneity involving very accurate Zernike3D fields expanded by a larger number of coefficients. To avoid excessive execution times, it would be worth doing a new ZART implementation on GPU to further improve its performance.
- The heterogeneity correction presented with ZART was only evaluated in a single iteration. This implies that conformational landscapes and ZART reconstructions were done just once. Nevertheless, heterogeneity correction might not have fully converged after the previous first iteration. Thus, we would like to explore the applicability of the Zernike3D and ZART algorithms in an iterative manner to

perform a flexible refinement. Flexible refinement would take advantage of successive landscape estimations and ZART reconstructions to help in the simultaneous convergence of both algorithms towards more meaningful structures and conformational landscapes.

- The analysis of conformational landscapes remains a challenging task to users, due to the lack of advanced visualization tools providing an interactive interface that allows to explore in real time and annotate the conformational states captured in the landscape. Therefore, the development of such tools would suppose a great benefit in the usability of the programs described along the dissertation.

Conclusiones y futuro

El objetivo principal de esta tesis, estructurada como un compendio de artículos, consiste en introducir un nuevo método para analizar la variabilidad conformacional continua a partir de varios tipos de datos en CryoEM (mapas, modelos estructurales y partículas) llamado Zernike3D.

El nuevo método se basa en la base Zernike3D, una composición de polinomios de Zernike y armónicos esféricos capaces de expandir cualquier función contenida en una esfera. Por lo tanto, la expansión de la base se basa tanto en los componentes de la base como en los coeficientes de Zernike3D $\alpha_{l,n,m}$, que determinan la contribución de cada componente a la expansión. La base se utiliza para expandir un campo de deformación que describe la transición entre dos estados conformacionales, que pueden ser calcula mediante la minimización de una función de coste que define los coeficientes de Zernike3D $\alpha_{l,n,m}$ como incógnita. Aunque nuestra elección de base no es única, encontramos que tiene propiedades interesantes que pueden ser explotadas para superar varios desafíos durante la búsqueda de los campos de deformación de Zernike3D.

La aplicabilidad y el rendimiento de la base fueron evaluadas en varios conjuntos de datos (tanto experimentales como sintéticos), demostrando la versatilidad del método incluso cuando es sometido a diferentes tipos de datos en CryoEM. La capacidad de aplicar la base de Zernike3D a mapas y partículas permite extraer y analizar la información de heterogeneidad estimada de diversas maneras, abriendo la posibilidad al diseño de análisis combinados donde la flexibilidad discreta y continua pueden entenderse en conjunto. Además, la aplicación de los campos estimados a partir de mapas o partículas directamente a modelos estructurales proporciona un nuevo mecanismo para comprender el paisaje conformacional a nivel atómico, así como analizar los estados aproximados con herramientas orientadas a al estudio de modelos estructurales como el Análisis de Modos Normales o la

Dinámica Molecular.

Además de la identificación y exploración de paisajes conformacionales, las estimaciones de Zernike3D también pueden considerarse como una alineación no rígida capaz de corregir las variabilidades conformacionales durante el proceso de reconstrucción de un volumen, lo que lleva a una disminución en los artefactos inducidos por movimiento presentes en un mapa. El enfoque anterior se implementó en un nuevo método de reconstrucción llamado ZART, que combina reconstrucción, mejora en la nitidez y corrección de heterogeneidad para producir reconstrucciones más significativas con características estructurales mejoradas. La implementación de ZART también incluye un enfoque de reconstrucción de múltiples resoluciones, capaz de analizar la resolución local de los mapas que se están reconstruyendo para dar un peso de resolución diferente a diferentes áreas en el mapa. De esta manera, ZART puede centrarse automáticamente en mejorar aquellas regiones del volumen donde se encuentra la proteína, mejorando aún más la precisión y la velocidad de convergencia del método.

3.2. Trabajo futuro

En esta sección identificamos los principales pasos a seguir para mejorar los métodos Zernike3D y ZART descritos a lo largo de la tesis doctoral:

1. Los esquemas clásicos de optimización propuestos para la búsqueda de los coeficientes de Zernike3D $\alpha_{l,n,m}$ podrían mejorarse si se emplea una red neuronal. El uso de una red neuronal debería proporcionar una búsqueda mucho más rápida de los coeficientes, así como encontrar conformaciones más significativas gracias a la optimización por lotes, que debería ser capaz de recuperar parcialmente mejor la información a lo largo de la dirección de proyección de una partícula a través de la información de otras vistas.
2. ZART está implementado únicamente en CPU siguiendo una estrategia multihilo, lo que mejora el rendimiento del algoritmo. Sin embargo, se puede sufrir una pérdida de rendimiento al corregir la heterogeneidad a través de campos de Zernike3D expandidos por un mayor número de coeficientes. Para evitar tiempos de ejecución excesivos, se propone desarrollar una nueva implementación de ZART en GPU.

3. La corrección de heterogeneidad presentada con ZART fue evaluada a partir de una única estimación de heterogeneidad, lo cual implica que los paisajes conformacionales y las reconstrucciones ZART se hicieron una única vez. Sin embargo, es posible que la corrección de la heterogeneidad no haya convergido completamente después de esta primera iteración. Por lo tanto, explorar la aplicabilidad de los algoritmos Zernike3D y ZART de manera iterativa para realizar un refinamiento flexible podría suponer una mejora significativa en el uso de ZART. El refinamiento flexible aprovecharía las sucesivas estimaciones de paisajes conformacionales y las reconstrucciones ZART para ayudar en la convergencia simultánea de ambos algoritmos hacia mapas y estados estructurales más significativos.
4. El análisis de paisajes conformacionales sigue siendo una tarea desafiante para los usuarios, debido a la falta de herramientas avanzadas de visualización con una interfaz interactiva que permita explorar y anotar en tiempo real los estados conformacionales capturados en el paisaje. Por lo tanto, el desarrollo de tales herramientas supondría un gran beneficio en la usabilidad de los programas descritos a lo largo de la tesis.

References

- [1] D. J. De Rosier y A. Klug, «Reconstruction of three dimensional structures from electron micrographs.,» *Nature*, vol. 217, pp. 130-134, 1968.
- [2] «Method of the Year 2015,» *Nature Methods*, vol. 13, nº 1, 2016.
- [3] A. McPherson, Introduction to macromolecular crystallography, John Wiley & Sons, 2011.
- [4] P. J. Hore, Nuclear magnetic resonance, USA: Oxford University Press, 2015.
- [5] M. Serna, «Hands on methods for high resolution cryo-electron microscopy structures of heterogeneous macromolecular complexes.,» *Frontiers in molecular biosciences*, vol. 6, p. 33, 2019.
- [6] S. H. Scheres y e. al, «Disentangling conformational states of macromolecules in 3D-EM through likelihood optimization.,» *Nature methods*, vol. 4, nº 1, pp. 27-29, 2007.
- [7] G. T. Herman y M. Kalinowski, «Classification of heterogeneous electron microscopic projections into homogeneous subsets,» *Ultramicroscopy*, vol. 108, nº 4, pp. 327-338, 2008.
- [8] F. J. Sigworth, «Principles of cryo-EM single-particle image processing,» *Journal of Electron Microscopy*, vol. 65, nº 1, pp. 57-67, 2015.
- [9] J. M. de la Rosa-Trevín y et al., «Xmipp 3.0: an improved software suite for image processing in electron microscopy.,» *Journal of Structural Biology*, vol. 184, nº 2, pp. 321-328, 2013.
- [10] J. M. de la Rosa-Trevín y e. al., «Scipion: a software framework toward integration, reproducibility and validation in 3D electron microscopy.,» *Journal of Structural Biology*, vol. 195, nº 1, pp. 93-99, 2016.
- [11] D. Střelák y e. al, «FlexAlign: An Accurate and Fast Algorithm for Movie Alignment in Cryo-Electron Microscopy.,» *Electronics*, vol. 9, p. 1040, 2020.
- [12] Q. Z. Shawn y e. al, «MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy,» *Nature Methods*, vol. 14, pp. 331-332, 2017.
- [13] S. S. H., «Beam-induced motion correction for sub-megadalton cryo-EM particles,» *eLife*, vol. 3, p. e03665, 2014.
- [14] K. Zhang, «Gctf: Real-time CTF determination and correction,» *Journal of structural biology*, vol. 193, nº 1, pp. 1-12, 2016.
- [15] J. A. M. y G. Nikolaus, «Accurate determination of local defocus and specimen tilt in electron microscopy,» *Journal of Structural Biology*, vol. 142, nº 3, pp. 334-347, *Journal of Structural Biology*.
- [16] A. Rohou y N. Grigorieff, «CTFFIND4: Fast and accurate defocus estimation from electron micrographs.,» *Journal of Structural Biology*, vol. 122, nº 2, pp. 216-221, 2015.
- [17] C. O. S. Sorzano y e. al., «Automatic particle selection from electron micrographs using machine learning techniques,» *Journal of Structural Biology*, vol. 167, nº 3, pp. 252-260, 2009.
- [18] T. Wagner y e. al., «SPHIRE-crYOLO is a fast and accurate fully automated particle picker for cryo-EM,» *Communications Biology*, vol. 2, nº 1, p. 218, 2019.
- [19] H. C. Wong y e. al, «Model-based particle picking for cryo-electron microscopy.,» *Journal of*

- Structural Biology*, vol. 145, nº 1-2, pp. 157-167, 2004.
- [20] C. O. S. Sorzano y e. al, «A clustering approach to multireference alignment of single-particle projections in electron microscopy.,» *Journal of Structural Biology*, vol. 171, nº 2, pp. 197-206, 2010.
- [21] Y.-C. C. y e. al, «Fast Cryo-EM Image Alignment Algorithm Using Power Spectrum Features,» *Journal of Chemical Information and Modeling*, vol. 61, nº 9, pp. 4795-4806, 2021.
- [22] S. H. W. S. y e. al, «Maximum-likelihood Multi-reference Refinement for Electron Microscopy Images,» *Journal of Molecular Biology*, vol. 348, nº 1, pp. 139-149, 2005.
- [23] J. Vargas y e. al., «Efficient initial volume determination from electron microscopy images of single particles,» *Bioinformatics*, vol. 30, nº 20, pp. 2891-2898, 2014.
- [24] J. Gomez-Blanco y e. al., «A robust approach to ab initio cryo-electron microscopy initial volume determination,» *Journal of Structural Biology*, vol. 208, nº 3, p. 107397, 2019.
- [25] H. Elmlund y e. al, «PRIME: Probabilistic Initial 3D Model Generation for Single-Particle Cryo-Electron Microscopy,» *Structure*, vol. 21, nº 8, pp. 1299-1306, 2013.
- [26] A. Punjani y e. al., «Non-uniform refinement: adaptive regularization improves single-particle cryo-EM reconstruction,» *Nature methods*, vol. 17, nº 12, pp. 1214-1221, 2020.
- [27] M. Zehni y e. al, «Joint Angular Refinement and Reconstruction for Single-Particle Cryo-EM,» *IEEE Transactions on Image Processing*, vol. 29, pp. 6151-6163, 2020.
- [28] E. Ramírez-Aportela y e. al., «Automatic local resolution-based sharpening of cryo-EM maps,» *Bioinformatics*, vol. 36, nº 3, pp. 765-772, 2020.
- [29] E. Ramírez-Aportela y e. al, «Automatic local resolution-based sharpening of cryo-EM maps.,» *Bioinformatics*, vol. 36, pp. 765-772, 2020.
- [30] F. J. J. y e. al, «Sharpening high resolution information in single particle electron cryomicroscopy.,» *Journal of Structural Biology*, vol. 164, nº 1, pp. 170-175, 2008.
- [31] S. Jonic y e. al., «Denoising of high-resolution single-particle electron-microscopy density maps by their approximation using three-dimensional Gaussian functions,» *Journal of Structural Biology*, vol. 194, nº 3, pp. 423-433, 2016.
- [32] K. Ramlal y e. al, «Local Agreement Filtering Algorithm for Transmission EM Reconstructions.,» *Journal of Structural Biology*, vol. 205, pp. 30-40, 2019.
- [33] S. Jonic y e. al, «Denoising of high-resolution single-particle electron-microscopy density maps by their approximation using three-dimensional Gaussian functions.,» *Journal of Structural Biology*, vol. 194, pp. 423-433, 2016.
- [34] J. L. Vilas y e. al, «Measuring local-directional resolution and local anisotropy in cryo-EM maps,» *Nature communications*, vol. 11, nº 55, 2020.
- [35] E. Ramírez-Aportela y e. al, «DeepRes: a new deep-learning- and aspect-based local resolution method for electron-microscopy maps,» *IUCRj*, vol. 6, pp. 1054-1063, 2019.
- [36] G. Harauz y M. van Heel, «Exact filters for general geometry three dimensional reconstruction,» *Optik*, vol. 73, pp. 146-156, 1986.
- [37] K. Jamali y e. al., «ModelAngelo: automated model building in Cryo-EM maps,» *arXiv preprint*, 2022.
- [38] C. K. Cassidy y e. al, «CryoEM-based hybrid modeling approaches for structure determination,» *Current opinion in microbiology*, vol. 43, pp. 14-23, 2018.
- [39] J. Pfab y e. al, «DeepTracer for fast de novo cryo-EM protein structure modeling and special studies on CoV-related complexes,» *PNAS*, vol. 118, p. e2017525118, 2021.
- [40] S. H. W. Scheres, «Relion: Implementation of a bayesian approach to cryo-em structure determination,» *Journal of Structural Biology*, vol. 180, pp. 519-530, 2012.
- [41] S. H. W. Scheres y et al., «Image processing for electron microscopy single-particle analysis

- using XMIPP,» *Nature Protocols*, vol. 3, pp. 997-990, 2008.
- [42] E. D. Zhong y et al., «CryoDRGN: reconstruction of heterogeneous cryo-EM structures using neural networks,» *Nature Methods*, vol. 18, pp. 176-185, 2021.
- [43] M. Chen y S. J. Ludtke, «Deep learning-based mixed-dimensional Gaussian mixture model for characterizing variability in cryo-EM,» *Nature Methods*, vol. 18, pp. 930-936, 2021.
- [44] J. Frank y A. Ourmazd, «Continuous changes in structure mapped by manifold embedding of single-particle data in cryo-EM,» *Methods*, vol. 100, nº 1, pp. 61-67, 2016.
- [45] A. Punjani y D. J. Fleet, «3DFlex: determining structure and motion of flexible proteins from cryo-EM,» *Nature Methods*, 2023.
- [46] Q. Jin, «Iterative elastic 3D-to-2D alignment method using normal modes for studying structural dynamics of large macromolecular complexes.,» *Structure*, vol. 22, pp. 496-506, 2014.
- [47] G. Tang y et al., «EMAN2: an extensible image processing suite for electron microscopy,» *Journal of Structural Biology*, vol. 157, pp. 38-46, 2007.
- [48] S. Jonic y C. O. S. Sorzano, «Versatility of Approximating Single-Particle Electron Microscopy Density Maps Using Pseudoatoms and Approximation-Accuracy Control,» *Biomed Research International*, 2016.
- [49] M. Harastani, «HEMNMA-3D: Cryo Electron Tomography Method Based on Normal Mode Analysis to Study Continuous Conformational Variability of Macromolecular Complexes,» *Frontiers in Molecular Biosciences*, vol. 8, 2021.
- [50] M. Harastani, «ContinuousFlex: Software package for analyzing continuous conformational variability of macromolecules in cryo electron microscopy and tomography data.,» *Journal of Molecular Biology*, vol. 214, nº 4, 2022.
- [51] A. Punjani y et al., «cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination,» *Nature Methods*, vol. 14, pp. 290-296, 2017.
- [52] D. Herreros y et al., «Approximating deformation fields for the analysis of continuous heterogeneity of biological macromolecules by 3D Zernike polynomials,» *IUCrJ*, vol. 8, nº 6, pp. 992-1005, 2021.
- [53] I. Bahar y et al., «Global dynamics of proteins: bridging between structure and function,» *Annual Review of Biophysics*, vol. 39, pp. 23-42, 2010.
- [54] D. Strelak y et al., «A GPU acceleration of 3-D Fourier reconstruction in cryo-EM,» *The International Journal of High Performance Computing Applications*, vol. 33, nº 5, 2019.
- [55] D. Herreros y et al., «ZART: A Novel Multiresolution Reconstruction Algorithm with Motion-blur Correction for Single Particle Analysis,» *Journal of Molecular Biology*, vol. 435, nº 9, 2023.
- [56] C. O. S. Sorzano y e. al., «A survey of the use of iterative reconstruction algorithms in electron microscopy,» *BioMed Res. Int.*, vol. 2017, 2017.
- [57] J. L. Vilas y et al., «Monores: Automatic and accurate estimation of local resolution for electron microscopy maps,» *Structure*, vol. 26, pp. 337-344, 2018.
- [58] M. J. D. Powell, «An efficient method for finding the minimum of a function of several variables without calculating derivatives,» *The Computer Journal*, vol. 7, nº 2, pp. 155-162, 1964.
- [59] J. L. Bentley, «Multidimensional binary search trees used for associative searching,» *Communications of the ACM*, vol. 18, nº 9, pp. 509-517, 1975.
- [60] I. Arganda-Carreras y et al., «Consistent and elastic registration of histological sections using vector-spline regularization,» de *Computer Vision Approaches to Medical Image Analysis: Second International ECCV Workshop*, Graz, Austria, 2006.
- [61] I. T. Jolliffe y J. Cadima, «Principal component analysis: a review and recent developments,» *Philosophical transactions of the royal society A: Mathematical, Physical and Engineering*

- Sciences*, vol. 374, nº 2065, p. 20150202, 2016.
- [62] L. McInnes y et al., «UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction,» *Arxiv*, 2018.
- [63] W. K. Härdle y L. Simar, «Multidimensional Scaling,» de *Applied Multivariate Statistical Analysis*, Heidelberg: Springer, 2012, pp. 397-412.
- [64] Y. Zhang y et al., «State-dependent sequential allostery exhibited by chaperonin TRiC/CCT revealed by network analysis of Cryo-EM maps,» *Progress in Biophysics and Molecular Biology*, vol. 160, pp. 104-120, 2021.
- [65] Y. Zhang y et al., «Allosteric Transitions of Supramolecular Systems Explored by Network Models: Application to Chaperonin GroEL,» *PLOS Computational Biology*, vol. 5, nº 4, 2009.
- [66] S. Zhang y et al., «ProDy 2.0: increased scale and scope after 10 years of protein dynamics modelling with Python,» *Bioinformatics*, vol. 37, nº 20, pp. 3657-3659, 2021.
- [67] C. O. S. Sorzano, «Fast and accurate conversion of atomic models into electron density maps,» *AIMS Biophysics*, vol. 2, nº 1, pp. 8-20, 2015.
- [68] A. Amunts y et al., «The structure of the human mitochondrial ribosome,» *Science*, vol. 348, nº 6230, pp. 95-98, 2015.
- [69] C. O. S. Sorzano y et al., «Local analysis of strains and rotations for macromolecular electron microscopy maps,» *Journal of Structural Biology*, vol. 195, nº 1, pp. 123-128, 2016.
- [70] W. Wong y et al., «Cryo-EM structure of the 80S ribosome bound to the anti-protozoan drug emetine,» *eLife*, 2014.
- [71] C. Plaschka y et al., «Structure of a pre-catalytic spliceosome,» *Nature*, vol. 564, pp. 617-621, 2017.
- [72] R. Melero y et al., «Continuous flexibility analysis of SARS-CoV-2 spike prefusion structures,» *IUCr*, vol. 7, pp. 1059-1069, 2020.
- [73] D. Wrapp y et al., «Cryo-em structure of the 2019-ncov spike in the prefusion conformation,» *Science*, vol. 367, nº 6483, pp. 1260-1263, 2020.
- [74] Y. Z. Tan y et al., «Cryo-em structures and regulation of arabinofuranosyltransferase aftd from mycobacteria,» *Molecular Cell*, vol. 78, nº 4, pp. 683-699, 2020.
- [75] R. Sanchez-Garcia y et al., «Deepemhancer: a deep learning solution for cryo-em volume post-processing,» *Communications Biology*, vol. 4, p. 874, 2021.
- [76] T. Bendory y et al., «Single-Particle Cryo-Electron Microscopy: Mathematical Theory,» *IEEE Signal Processing*, vol. 37, pp. 58-76, 2019.
- [77] T. R. D. Costa y et al., «Structural analysis of protein complexes by Cryo Electron Microscopy,» de *Bacterial Protein Secretion Systems*, New York, Humana Press, 2017, pp. 377-413.
- [78] G. J. Georges y et al., «The Contorsbody, an antibody format for agonism: Design, structure, and function,» *Computational and Structural Biotechnology Journal*, vol. 18, pp. 1210-1220, 2020.
- [79] D. Herreros y et al., «Estimating conformational landscapes from Cryo-EM particles by 3D Zernike polynomials,» *Nature Communications*, vol. 14, nº 154, 2023.
- [80] F. Abbot, Artist, *This infographic explains the basics of cryo-EM..* [Art]. SLAC, 2018.

Appendices

Approximating deformation fields for the analysis of continuous heterogeneity of biological macromolecules by 3D Zernike polynomials

Herreros, D., Lederman, R. R., Krieger, J., Jiménez-Moreno, et al. Approximating deformation fields for the analysis of continuous heterogeneity of biological macromolecules by 3D Zernike polynomials. *IUCrJ*, 8, 992-1005 (2021). <https://doi.org/10.1107/S2052252521008903>



Approximating deformation fields for the analysis of continuous heterogeneity of biological macromolecules by 3D Zernike polynomials

David Herreros,^{a*} Roy R. Lederman,^b James Krieger,^c Amaya Jiménez-Moreno,^a Marta Martínez,^a David Myška,^d David Strelak,^{a,e} Jiri Filipovic,^d Ivet Bahar,^c Jose Maria Carazo^a and Carlos Oscar S. Sanchez^a

Received 26 March 2021

Accepted 25 August 2021

Edited by E. Bullitt, Boston University School of Medicine, USA

Keywords: multi-dimensional scaling (MDS); 3D reconstruction and image processing; single-particle cryo-EM; spherical harmonics; Zernike polynomials; conformations.

Supporting information: this article has supporting information at www.iucrj.org

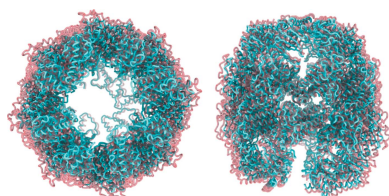
^aCentro Nacional de Biotecnología-CSIC, C/ Darwin 3, Cantoblanco, Madrid 28049, Spain, ^bDepartment of Statistics and Data Science, Yale University, New Haven, Connecticut, USA, ^cDepartment of Computational and Systems Biology, University of Pittsburgh, Pennsylvania, USA, ^dInstitute of Computer Science, Masaryk University, Botanická 68a, 60200 Brno, Czech Republic, and ^eFaculty of Informatics, Masaryk University, Botanická 68a, 60200 Brno, Czech Republic. *Correspondence e-mail: dherreros@cnb.csic.es

Structural biology has evolved greatly due to the advances introduced in fields like electron microscopy. This image-capturing technique, combined with improved algorithms and current data processing software, allows the recovery of different conformational states of a macromolecule, opening new possibilities for the study of its flexibility and dynamic events. However, the ensemble analysis of these different conformations, and in particular their placement into a common variable space in which the differences and similarities can be easily recognized, is not an easy matter. To simplify the analysis of continuous heterogeneity data, this work proposes a new automatic algorithm that relies on a mathematical basis defined over the sphere to estimate the deformation fields describing conformational transitions among different structures. Thanks to the approximation of these deformation fields, it is possible to describe the forces acting on the molecules due to the presence of different motions. It is also possible to represent and compare several structures in a low-dimensional mapping, which summarizes the structural characteristics of different states. All these analyses are integrated into a common framework, providing the user with the ability to combine them seamlessly. In addition, this new approach is a significant step forward compared with principal component analysis and normal mode analysis of cryo-electron microscopy maps, avoiding the need to select components or modes and producing localized analysis.

1. Introduction

The application in electron microscopy of techniques such as cryo-electron microscopy (cryo-EM), single-particle analysis (SPA) (Carroni & Saibil, 2016) or electron cryo-tomography (Schur, 2019) has proven to be a versatile tool to trace high-resolution structures. In particular, cryo-EM SPA has proven to be especially good at providing not only one structure, but a series of them, with most methods aiming to resolve stable states that are referred to as classes. In this way, we get a first approximation to the conformational landscape of the macromolecule, albeit restricted to these stable states.

However, the limited number of classes that can be extracted from a 3D classification is usually not enough to unveil fully the dynamics of a given macromolecule. The complete characterization of a conformational landscape can only be achieved through the analysis of multiple transient and stable states needed to describe the molecular flexibility in a more accurate manner. The knowledge of these transient and stable states leads to a better description of how structural



changes might affect molecular function or interaction affinity, among other properties of interest.

The formulation we introduce here is oriented towards modelling continuous flexibility (Sorzano *et al.*, 2019), which can be used to characterize the motions undergone by a molecule when exploring different states. We have already addressed this problem in our previous work on continuous heterogeneity using normal mode analysis (NMA) (Sanchez Sorzano *et al.*, 2016). However, this process relied on manual selection of the modes describing the structural changes reflected by two cryo-EM maps, thus making the analysis of molecular flexibility more complex for the user. The new algorithm that we propose in this work tries to address this problem by simplifying the analysis for the user.

In our new methodology, there is no longer a normal modes space where some choices have to be made. Instead, a totally new approach is presented here, based on an expansion on a 3D basis that does not require user intervention at all. We have also improved the analysis of pairwise comparisons by introducing a multidimensional scaling algorithm that automatically combines the outputs from two different metrics. Finally, the new algorithm also allows the analysis of local strains and rotations, as done by us earlier (Sorzano *et al.*, 2016), with the advantage of having all the analyses integrated into a single mathematical framework. We provide a more in-depth comparison with alternative methods in Section 2.1.

The paper makes the following major contributions.

(i) The development of an automatic algorithm to analyse continuous heterogeneity of macromolecules through cryo-EM maps.

(ii) Representation of the strain and rotation components defining a transition between two different conformational states.

(iii) Representation of a series of conformations in a structure mapping and consensus of different mappings defined by different comparison metrics.

(iv) A methodology to compare cryo-EM maps with simulated data.

(v) The application of deformation fields to atomic structures to predict different conformations given by a series of cryo-EM maps.

2. Methods

2.1. Determining structural deformations

In order to detect the movements defining a conformational transition between two states of the same macromolecule, we need to determine the displacements that each region of the molecule will undergo between the two states. The key development in this work is the successful expression of the maps in terms of a mathematical basis on which the displacements are calculated. Although full details are provided in Appendices A and B, here it suffices to say that we use a generalized form of Zernike polynomials to expand functions on a ball (as the macromolecule we are interested in is defined inside a spherical volume). This is not the only possible choice of basis functions [for example, it would have

been possible to use the Laguerre polynomials described by Provencher & Voguel (2010) or the prolate spheroidal functions (Greengard & Serkh, 2018)], and we do not expect superiority of any of these possible bases as long as all of them are bases of functions defined within a sphere. Additionally, we find that Zernike polynomials have some appealing mathematical properties especially well suited to our problem. Indeed, these Zernike polynomials allow for the expansion of functions on a sphere which do not vanish at the boundaries (so that the more external parts of the macromolecule can move). Moreover, the basis is closed under rotations. In Appendix B we further explore its properties and its relationship to spherical harmonics.

Considering a pair of electron-density maps representing two conformational states of a macromolecule, it is possible to pose the displacement-finding problem as

$$\min_{\mathbf{g}_L} \int |V_1(\mathbf{r}) - V_2[\mathbf{r} + \mathbf{g}_L(\mathbf{r})]| \, d\mathbf{r}, \quad (1)$$

where V_1 and V_2 represent two conformations of a given molecule. Here it is important to note that we are measuring the distance between the target and the distorted volumes in terms of the L_1 norm. Although it would also have been possible to use the L_2 norm, we have chosen this definition as it is more robust to outliers (*i.e.* it is more robust to those cases where the maps do not match completely or have missing regions). The displacement to be applied to the coordinates of V_1 is defined by the deformation field $\mathbf{g}(\mathbf{r})$ parameterized through the expansion in Zernike polynomials $Z_{l,n,m}(\mathbf{r})$ (see Appendix A),

$$\mathbf{g}_L(\mathbf{r}) = \sum_{l=0}^L \sum_{n=0}^N \sum_{m=-l}^l \begin{pmatrix} \alpha_{l,n,m}^x \\ \alpha_{l,n,m}^y \\ \alpha_{l,n,m}^z \end{pmatrix} Z_{l,n,m}(\mathbf{r}), \quad (2)$$

where N and L represent the maximum allowed degrees for the Zernike polynomials and the corresponding spherical harmonics, respectively.

The amount of displacement at every point is controlled by the deformation coefficients $\alpha_{l,n,m}$. Our objective is to find the deformation coefficients that minimize the goal function in equation (1). This is achieved through a Powell's conjugate direction method starting from an initial guess of $\alpha_{l,n,m} = 0$ for all indices l, n, m and directions x, y, z (that is, no deformation). This initialization of the minimization method assumes that the identity/equilibrium solution ($\alpha_{l,n,m} = 0$) is close enough to the real solution defining the structural transition represented by the cryo-EM maps. Since in most of the cases this assumption is fulfilled, this initial guess allows the minimization method to find the set of coefficients that appropriately describes the motion between the two maps. However, it is important to note that there are many local minima where the minimization process could be trapped. In this respect, and although in our experience the initialization conditions proposed in this work provide results close enough to the ideal solution, there could be cases in which other ways to initialize the algorithm could be more beneficial in terms of minima search.

The deformation field estimated above can be submitted to the local strain and rotation analysis described by Sorzano *et al.* (2016). This analysis reveals the nature (stretching, compression or rotation) of the local forces acting on V_1 to transform it into V_2 as well as their local intensity.

In our deformation model, it is possible to divide the movements that a molecule may undergo into ‘low’- and ‘high’-frequency movements, depending on how localized these movements are, *e.g.* a transition from an open to a closed state can be considered a low-frequency movement, while the rotation of a specific α -helix might be a high-frequency movement. Parameters L and N specify the maximum degree of the polynomials used in the description of molecular flexibility. In this way, we may control the maximum frequency of the movements that could be analysed by the basis. Obviously, analysing larger L and N will result in a longer computational time, because more $\alpha_{l,n,m}$ coefficients will need to be determined and there is a higher risk of overfitting. However, the larger the values given to the parameters L and N , the higher the frequencies the algorithm will be able to analyse (although in general, global motions dominate the conformational change; Bahar *et al.*, 2010).

Although in many cases analysing global motions is enough to describe in a precise manner the structural changes a macromolecule may undergo, it can be the case that the motions of interest are focused on a very localized area of the molecule. In that case, being able to go to higher degrees on the basis will allow the algorithm to study those motions specifically, without modifying the areas that should remain still. Another possibility is direct restriction of the structural analysis to any specific region in the macromolecule by centring a sphere on that area and selecting an appropriate radius. In this way, it will not be necessary to reach very high degrees in the basis (thus reducing the computational complexity). However, by imposing these kinds of restrictions the algorithm might include artefacts in the surface of the sphere as the molecular regions outside of it will remain untouched. Depending on the molecule and motions to be analysed, the researcher can decide which analysis will be more appropriate for a specific case.

It is important to mention that only in a very few cases did we need to increase the degree of the basis to analyse a localized motion that we were interested in, or have to play with the regularization parameter to get a better approximation of the deformation fields, since the default values were good enough for most of the experiments we have performed so far.

To reduce the possibility of overfitting as much as possible, we regularize the cost function by adding two penalty terms,

$$\min_{\alpha_{l,n,m}} \int |V_1(\mathbf{r}) - V_2[\mathbf{r} + \mathbf{g}_L(\mathbf{r})]| \, d\mathbf{r} + \lambda_1 \int \|\mathbf{g}_L(\mathbf{r})\|^2 \, d\mathbf{r} + \lambda_2 \frac{\int \{V_1(\mathbf{r}) - V_1[\mathbf{r} + \mathbf{g}_L(\mathbf{r})]\} \, d\mathbf{r}}{\int V_1(\mathbf{r}) \, d\mathbf{r}}. \quad (3)$$

The first term of the regularization penalizes excessive deformation and the second penalizes changes in the mass of V_1 due to the deformation. Regularization terms λ_1 and λ_2 are usually given low values to prevent large deviations from the ideal solution. Nevertheless, both can be set by the user to any value they consider appropriate for their specific analysis. The guideline for their selection should be that the three terms in the goal function should have values of the same order of magnitude. In our implementation, we report the three contributions helping the user to choose these multipliers.

2.2. Relationship to other continuous deformation models

Probably the two most widely continuous deformation models used by the structural biology community in mapping the conformational space of biomolecules (or in analysing cryo-EM images) are principal component analysis (PCA) (Tagare *et al.*, 2015) and normal mode analysis (NMA) (Cui & Bahar, 2006). The three models (PCA, NMA and 3D Zernike) claim to be bases for continuous movements. However, as will be clarified below, they define bases of different mathematical entities.

PCA considers a volume of size N^3 voxels as a vector in \mathbb{R}^{N^3} . Due to the continuous heterogeneity and the uncertainty in the 3D reconstruction process, the reconstructed map can be considered as the mean of a set of other vectors (maps) whose projections are acquired by the microscope. If we consider the covariance matrix associated with that set of maps (a matrix of size $N^3 \times N^3$), then the principal components form a basis (if the covariance matrix is not degenerate) in which the set of maps can be linearly expressed. The PCA approach approximates the deformed volume by a linear combination of volumes (the principal directions),

$$V_2(\mathbf{r}) \simeq V_1(\mathbf{r}) + \sum_n \alpha_n V_n(\mathbf{r}), \quad (4)$$

where V_n are the eigenvolumes of the PCA decomposition. The undeformed model is then obtained by subtracting the appropriate amount of each of the eigenvolumes,

$$V_1(\mathbf{r}) \simeq V_{2 \rightarrow 1}(\mathbf{r}) = V_2(\mathbf{r}) - \sum_n \alpha_n V_n(\mathbf{r}). \quad (5)$$

Due to the low-frequency nature of the PCA principal directions (Sorzano & Carazo, 2021), the undeformed volume is necessarily of low resolution.

In our model, we assume that any deformed volume V_2 can be undeformed by applying \mathbf{g}_L ,

$$V_1(\mathbf{r}) \simeq V_{2 \rightarrow 1}(\mathbf{r}) = V_2[\mathbf{r} + \mathbf{g}_L(\mathbf{r})]. \quad (6)$$

Zernike polynomials provide a basis for $\mathbf{g}_L(\mathbf{r})$, not the volumes. Our model revolves around the location of the voxel (which implies a nonlinear relationship between V_1 and V_2), providing an intrinsically better handling of the local characteristics of the map, while in PCA there is a linear model at the level of the volumes themselves (not their internal coordinates).

Our approach has another potential advantage over the PCA model: it can easily be applied to atomic structures fitted

into V_1 . For any given atom in the atomic structure at a position \mathbf{r}_1 , that is defined in the same coordinate system as V_1 , we simply have to move it to the location $\mathbf{r}_1 + \mathbf{g}_L(\mathbf{r}_1)$.

In NMA, volumes are approximated by a set of P pseudoatoms with weights c_p and basis function $b(\mathbf{r})$ located at the locations \mathbf{r}_p (Jonić & Sorzano, 2016),

$$V_2(\mathbf{r}) = \sum_p c_p b(\mathbf{r} - \mathbf{r}_p). \quad (7)$$

NMA is based on a second-order Taylor approximation of the energy landscape of the macromolecule, starting with a description of the interactions between the pseudoatoms. This is typically treated using an elastic network model where pseudoatoms within a distance criterion are connected by harmonic springs (Bahar *et al.*, 2010). The associated Hessian is of size $3P \times 3P$ and the normal modes are its eigenvectors (sorted by increasing eigenvalue) and a basis of the \mathbb{R}^{3P} space. Let us call $\mathbf{u}_k \in \mathbb{R}^{3P}$ the k th normal mode, and $\mathbf{u}_{k,p} \in \mathbb{R}^3$ the part of the normal mode corresponding to the p th pseudoatom. To deform V_2 to make it similar to V_1 we consider the first K normal modes with different weights α_k ,

$$V_{2 \rightarrow 1}(\mathbf{r}) = \sum_{p=1}^P c_p b \left[\mathbf{r} - \left(\mathbf{r}_p + \sum_k \alpha_k \mathbf{u}_{k,p} \right) \right]. \quad (8)$$

Similar to our method, NMA acts by displacing the location of the pseudoatoms (our model acts by displacing the location at which we must interpolate V_2). However, an advantage of our new method with respect to NMA is that the NMA deformation is only known at the location of the pseudoatoms, while our new method is fully defined within the sphere containing the macromolecule. In this way, the NMA would be a discretized version of the underlying continuous deformation field, while 3D Zernike polynomials would be an estimate of that continuous field.

Summarizing, each of the methods described so far (PCA, NMA and 3D Zernike polynomials) has a basis in different mathematical entities (vectors in \mathbb{R}^{N^3} , \mathbb{R}^{3P} or the set of square integrable functions defined within the sphere of a given radius). 3D Zernike polynomials have the advantage that they are defined for every point in the macromolecule (as opposed to NMA) and the undeformed volumes do not lose resolution (as opposed to PCA).

Elastic deformations have also become popular for the alignment of frames within a movie (Abrishami *et al.*, 2015; Tegunov & Cramer, 2019; Zheng *et al.*, 2017). Although they have not been explicitly used to deform volumes, one could envision that they could be easily extended to three dimensions. This would certainly be a possible approach and we earlier used cubic splines for this purpose (Sorzano *et al.*, 2016). However, the basis used in this paper, which is defined exclusively within a sphere, is more appropriate for the task at hand (describing a function whose support is fully contained within that sphere) than for a more generic set of functions that constitute a basis of functions defined within a cube. This ‘greater appropriateness’ translates into requiring fewer

coefficients to express the same deformation field to the same level of accuracy.

2.3. Distances between a set of maps

In most practical cases, the number of states that can be reconstructed by cryo-EM SPA is larger than two, which naturally implies the generalization of the case presented above to a number of pairwise operations capturing the different structural relationships among the set of maps under consideration. This information is summarized in a graph known as a structure map (Sanchez Sorzano *et al.*, 2016) or conformational landscape (Zhang *et al.*, 2021b), which represents each conformation as a point in conformational space. The closer two points are in the structure map, the more similar they are.

By estimating the Zernike polynomial deformation for all possible pair combinations in a set of N cryo-EM maps, a distance matrix can be computed in which we measure how far two cryo-EM maps are from each other. The deformation field between the two cryo-EM maps $\mathbf{g}_L(\mathbf{r})$ provides a mechanism for calculating such a distance. For instance, we may define the distance between two cryo-EM maps V_1 and V_2 as the sum of the lengths of the deformations at each point,

$$d_1(V_1, V_2) = \int \|\mathbf{g}_L(\mathbf{r})\|^2 d\mathbf{r}. \quad (9)$$

Besides equation (9), there are additional sensible ways of defining the distance between two cryo-EM maps. One of them consists of measuring the correlation between V_1 and V_2 once V_2 is undeformed to resemble V_1 ,

$$d_2(V_1, V_2) = 1 - \rho\{V_2[\mathbf{r} + \mathbf{g}_L(\mathbf{r})], V_1(\mathbf{r})\}, \quad (10)$$

where ρ is Pearson’s correlation coefficient.

By comparing all cryo-EM maps, we would construct a matrix of the distances of all versus all maps.

It is worth mentioning here that, in order to get accurate comparison measurements, it is desirable to have a set of cryo-EM maps with similar characteristics. In particular, it is important to filter the maps in the set so that all their resolutions match the lowest value present in the data set. In this way, the structure mappings and distance matrices will not be affected by resolution changes, leading to a more meaningful projection of the different maps in the low-dimensional space resulting from the application of this method.

2.4. Embedding of conformations using multiple multi-dimensional scaling

Once we have the above-mentioned distance matrix, we may use multidimensional scaling (MDS) (Härdle & Simar, 2012) to find points in a low-dimensional space of dimension p (typically $p = 2$ or $p = 3$ for ease of representation) such that the distances between points in the low-dimensional space represent in some form the distances between the cryo-EM maps in the full dimensionality space [*e.g.* equation (9)]. For a detailed description of MDS, see Härdle & Simar (2012). If we have N cryo-EM maps to compare, let us refer to the matrix

collecting all the points in the low-dimensional space as X_1 [$X_1 \in \mathbb{M}(N, p)$], that is, the set of cryo-EM maps of size $N \times p$]. The subscript 1 indicates that we used d_1 to perform the low-dimensional mapping.

If instead of equation (9) we use equation (10), then this would give us another MDS representation X_2 . While the distance d_1 concentrates on the amount of deformation required to transform V_1 into V_2 , d_2 describes the distance between V_1 and V_2 after applying the inverse deformation to V_2 .

We could similarly conceive other strategies to measure the distance between any pair of cryo-EM maps V_1 and V_2 . None of them should necessarily be better than the others, since each one addresses the problem from a different perspective. In this regard, it is impossible to favour any one of the different metrics without a specific task to accomplish. However, it is still sensible to combine the different mappings induced by each one of the distances as a way of producing a single summary of all their information. For the task of producing such a summary, we propose to construct a combination of the embeddings that minimizes the entropy of the result, understanding that the entropy is reduced when more order is found.

At this point and following the aforementioned idea, we may want to combine all those low-dimensional mappings into a single set of points to summarize the relative distances derived from each distance definition. For doing so we have found useful the following procedure that we call multiple multidimensional scaling:

(i) We take one of the mappings as reference, for instance, X_1 .

(ii) We look for the affine transformation T_i that minimizes the Frobenius norm [for an arbitrary matrix A , its Frobenius norm is defined as $\|A\|_F = (\sum_{i,j} |a_{ij}|^2)^{1/2}$] between each X_i transformed mapping and the reference mapping (since the MDS mappings of different distances, performed in an independent way, normally result in mappings of different scales, central locations, rotations and mirrors),

$$\operatorname{argmin}_{T_i} \|X_1 - T_i(X_i)\|_F. \quad (11)$$

For convenience of notation, let us define $T_1(X_1) = X_1$.

(iii) The consensus mapping is constructed as the convex combination of all transformed mappings (the determination of the specific α_i coefficients for the combination will be addressed in the following step),

$$X_a = \sum_i \alpha_i T_i(X_i), \quad (12)$$

with the constraints $\alpha_i \geq 0$ and $\sum_i \alpha_i = 1$ (with these constraints X_a is said to be a convex combination of the input matrices). Note that the j th row of the matrix X_a (referred to as $\mathbf{x}_{j,a}$) indicates the position of the j th cryo-EM map in the low-dimensional space (whose dimension is p). For each one of the consensus candidates we associate the probability density function

$$p_a(\mathbf{x}) = \sum_j \frac{1}{N} G_\sigma(\mathbf{x} - \mathbf{x}_{j,a}), \quad (13)$$

where G_σ is a p -multivariate spherical Gaussian whose covariance matrix is $\sigma^2 I$ [in our experiments, we chose $\sigma = \max[\text{range}(X_1), \dots, \text{range}(X_i)]/20$, where $\text{range}(X_i)$ is the difference between the maximum and minimum values of any of the components of the mapped vectors].

(iv) Since the best combination of coefficients α_i is not known beforehand, each possible convex combination has to be analysed. The criterion followed was to look for the convex combination that minimized the Shannon entropy of the probability density function defined above,

$$\operatorname{argmin}_a \left\{ - \int p_a(\mathbf{x}) \log[p_a(\mathbf{x})] d\mathbf{x} \right\}. \quad (14)$$

The rationale is that we are looking for the convex combination that brings maximum order to the low-dimensional mapping.

We observe that the procedure described above normally finds a good balance between the properties of the different low-dimensional mappings, resulting in well structured summaries.

3. Results

This algorithm has been implemented in *Xmipp* (de la Rosa-Trevín *et al.*, 2013) and it is available through *Scipion* (de la Rosa-Trevín *et al.*, 2016) under the protocols named `volume deform - Zernike3D` and `struct map - Zernike3D`.

We performed some tests with a pair of maps to compare these two implementations to analyse the performance improvement. The maps used for the tests had dimensions of 250 in X , Y and Z , leading to averaged execution times of 1 h and 20 min (CPU) and 39.5 s (GPU). The tests were performed with an Intel i7-9750H and a Nvidia 2060 with Cuda 10.1, respectively.

3.1. Experiment 1: cryo-EM maps of the human mitochondrial ribosome

We first tested our approach using a small data set covering a range of conformational states of a human mitochondrial ribosome (Amunts *et al.*, 2015), as previously described by Sanchez Sorzano *et al.* (2016). To check whether the structure map suggested two independent (pre-translocation and post-translocation) states following different conformational transitions as found in our previous study (Sanchez Sorzano *et al.*, 2016), we applied the methodology described above with $N = 3$ and $L = 2$ (the maximum allowed degrees for the Zernike polynomials and the spherical harmonics, respectively). As expected, the structure map indeed suggests two independent arrangements following their own conformational transitions, grouped as red and blue dots in Fig. 1 (the black line segments joining the dots are just provided to enhance visualization). We thus conclude that the new approach is capable of reproducing the results of previous supervised methods that perform similar analyses and accurately groups the seven

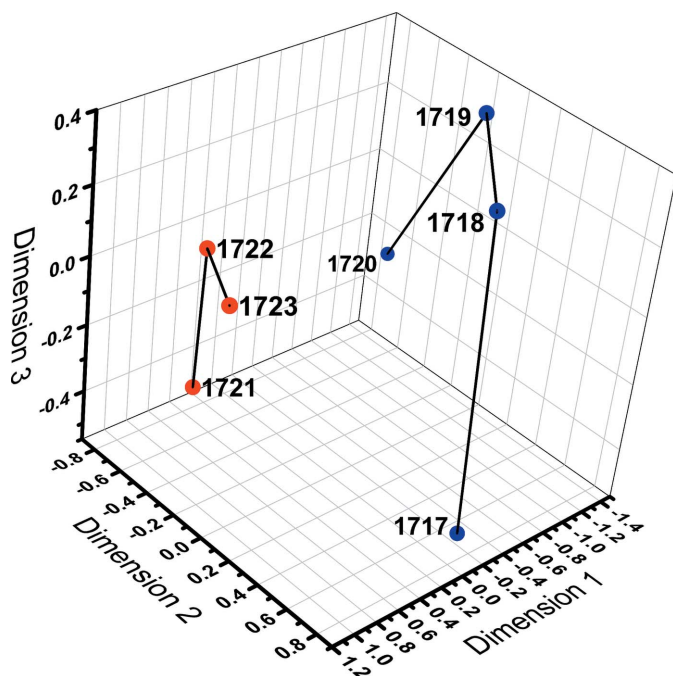


Figure 1

Structure mapping recovered a set of seven maps of the human mitochondrial ribosomes (Amunts *et al.*, 2015) from the data set retrieved from the EMDB after running the *Zernike3D* algorithm. Two trajectories are suggested that might correspond to two independent states (pre-translocation and post-translocation) present in the data set, consistent with results from a previous normal-mode-based structure mapping algorithm (Sanchez Sorzano *et al.*, 2016). The labels refer to the EMDB entries.

cryo-EM structures [indicated by their EMDB (Electron Microscopy Data Bank, <https://www.ebi.ac.uk/emdb/>) identification numbers] into two groups of conformers, each representative of a different functional state.

Additionally, the new approach also allows for the local decomposition of the deformation field into local strains and local rotations, as was done by Sorzano *et al.* (2016). The representation of these two components is shown in Fig. 2 for one of the pairs of ribosomes (EMDB entries 1720 and 1723). In addition, Video 1 in the supporting information shows the conformational changes described by these two maps. For this video, we coloured the ribosomes using the rotation component represented in Fig. 2 to simplify their comparison. According to this analysis, the rotations appear to be distributed through the whole structure of the ribosome, although the larger rotations (shown in red) are mostly found in the small subunit. Similarly, the strains are mainly localized in the small subunit and appear to be less distributed. This reveals that the basis is capable of deforming in a localized fashion, leading to a better description and identification of the different movements that define the transition between the two conformations. It is also possible to see that we are obtaining results comparable with those found by Sorzano *et al.* (2016), with the advantage of having all these analyses unified in the same framework, which implies an overall simplification leading to more complete studies.

3.2. Experiment 2: trajectory recovery of the CCT complex

Our next experiment is aimed at characterizing the ability of the method to recover the sequence of events present in a set of conformations defining a certain trajectory in conformational space. Such conformations can be created computationally by taking advantage of biophysical methods such as molecular dynamics simulations (MD) (Adcock & McCammon, 2006) and normal mode analysis (NMA) (Bahar *et al.*, 2010), simulating the movements defining a transition between two conformations. In this case, we used a trajectory from a recent study (Zhang *et al.*, 2021b), which was generated using a purely NMA-based approach called the adaptive anisotropic network model (adaptive ANM; Yang *et al.*, 2010) implemented in *ProDy* (Zhang *et al.*, 2021a). This gave us 30 different models along an open–closed transition of the mammalian chaperonin CCT complex between two atomic models derived from a previous cryo-EM study (Cong *et al.*, 2012), taken from the Protein Data Bank (PDB) (wwPDB Consortium, 2019), as described by Zhang *et al.* (2021b). The starting structure with one ring open and one ring closed (PDB entry 4a0w) (Cong *et al.*, 2012) corresponded to an ATP-bound state and the target structure with both rings in an intermediate conformation (PDB entry 4a13) (Cong *et al.*, 2012) corresponded to the ADP-bound state, allowing us to explore the conformational changes triggered by ATP hydrolysis. In the adaptive ANM method, all steps are based on coarse-grained normal modes calculated using the anisotropic network model (ANM) (Atilgan *et al.*, 2001; Doruker *et al.*, 2000; Eyal *et al.*, 2006), providing coordinate changes for C^α atoms only. At each step, normal modes were selected that had the highest directional overlap (correlation cosine) with the deformation vector between the current conformation and the target structure up until the sum of the squared overlaps exceeded a threshold of 0.4. The contribution of each mode to the deformation was chosen so as to take 20% of the maximum provided by the unnormalized dot products (a scaling factor of 0.2) so as to avoid unphysical deformations while maintaining efficiency. The normal modes were recalculated until the root-mean-square deviation (r.m.s.d.) from the target structure fell below 1 Å, resulting in a total of 30 steps. Each step recruited a larger number of modes and had a smaller total size as the required deformation became less

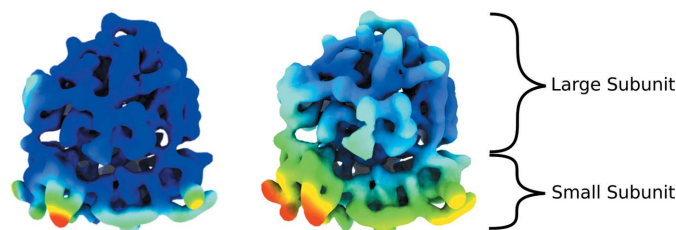


Figure 2

Mitochondrial ribosome subunits 28S and 39S (from EMDB entry 1720) coloured using the strain (left) and rotation (right) components extracted from the deformation coefficients obtained when analysing the motion described by EMDB 1720 and EMDB 1723. The conformational change described by these two maps is represented in Video 1.

cooperative and more local (see Video 2). We focus our discussion on the ring that goes from open to intermediate-closed for simplicity.

We then transformed these atomic structures into Coulomb potential maps using the electron atomic scattering factors (EASFs) as described in previous work (Sorzano *et al.*, 2015). Fig. 3 shows the structure maps recovered after applying our methodology. We can see that the sequential order of the 30 intermediate conformers along the trajectory was successfully recovered by our approach. The direction, however, is arbitrary and in this case the start of the trajectory was numbered as conformer 30 and the end as conformer 1.

With this example, we additionally illustrate the distinct MDS mappings obtained when the distances d_1 [amount of deformation, Fig. 3 (top)] and d_2 [similarity after deformation, Fig. 3 (middle)] are used. Although the trajectory was successfully recovered by both distances, the correlation distance d_2 was slightly more accurate in this case. The reason is that most of the changes between the structures at the end of the transition (labelled 1 to 13 by the algorithm) are high-frequency movements (*i.e.* movements of loops or small α -helices and β -sheets) that cannot be fully captured by the Zernike 3D basis with $N = 3$ and $L = 2$ (although larger N and L would allow one to express these small-detail movements, they would also increase the computational cost). Fortunately, the consensus mapping [Fig. 3 (bottom)] is able to identify the existence of high-frequency movements and gives more weight to the d_2 mapping (correlation distance) automatically, resulting in an almost exact recovery of the volume sequence along the trajectory.

At least in this case we can conclude that d_1 is very good for describing the low-frequency movements (*e.g.* C23–C30), while d_2 is very good for characterizing the high-frequency differences (*ca* C1–C13), and both perform well in the intermediate-frequency regime. Depending on whether our set of input maps are related by large or small movements, one distance or the other will be better suited to capturing the overall set of relationships. The consensus mapping will thus analyse both mappings and automatically determine the optimal weight that results in a low-dimensional mapping that can be readily interpreted.

3.3. Experiment 3: comparison of atomic models and cryo-EM maps from the rabbit ryanodine receptor RyR1

In the following example, we explored the possibility of matching (pseudo/simulated) cryo-EM maps derived from atomic models with experimental electron microscopy maps in the same low-dimensional space. For this purpose we selected five experimental cryo-EM maps deposited for the ryanodine receptor 1 (RyR1) from rabbit (EMDB entries 8379, 8385, 8390, 8395 and 8373) and their respective atomic models in the PDB (PDB entries 5tam, 5tau, 5taz, 5tb4 and 5tn).

First, we converted the atomic models into density maps using EASFs, as described in the previous section. Then, to make the cryo-EM maps and atomic models comparable, we also filtered all volumes in the analysis to a common resolution

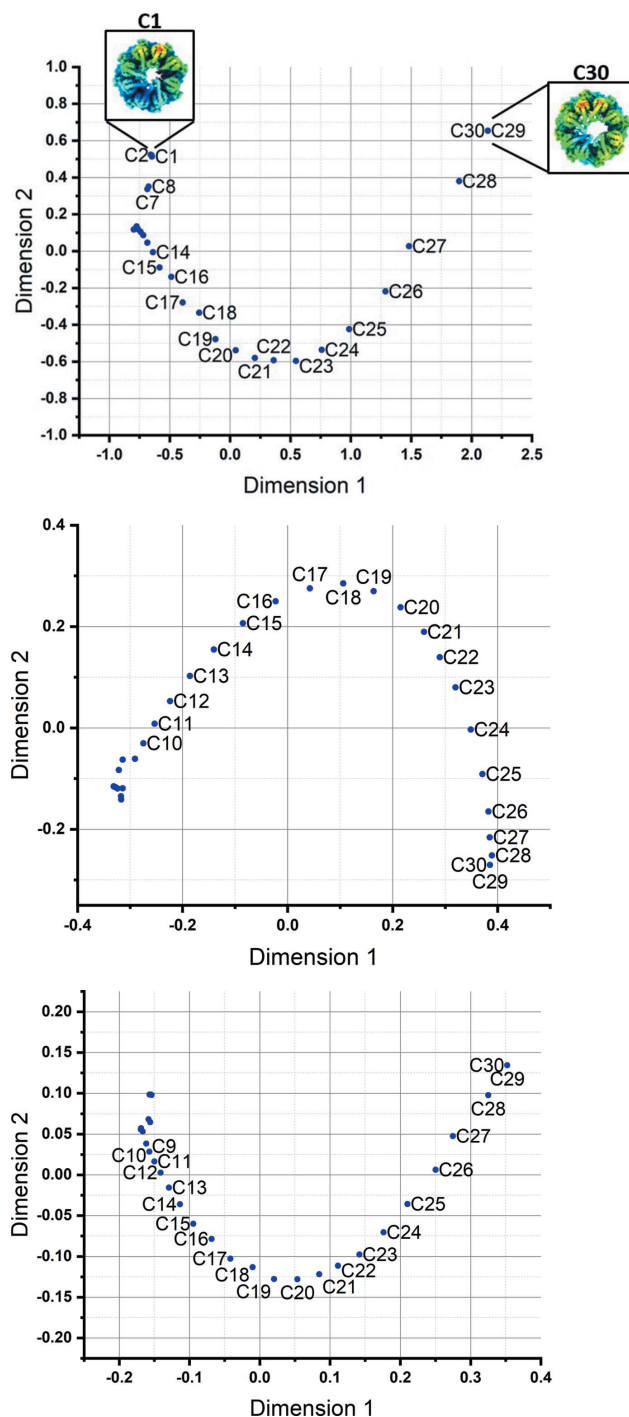


Figure 3 Structure maps of a set of 30 models obtained by an NMA-based approach called adaptive ANM over an open–closed transition of the chaperonin CCT from our previous study (Zhang *et al.*, 2021b) using (top) the deformation distance d_1 , (middle) the correlation distance d_2 and (bottom) the minimum entropy consensus followed by an MDS analysis of the corresponding distance matrices. The open conformation is labelled as C30 and the closed one is C1. The intermediates predicted along low-frequency modes starting from the open state are labelled C29, C28 *etc.*, whereas the vicinity of C1 populates conformers reached by high-frequency modes. The latter is relatively more sensitive to the metric used in the *Zernike3D*-based evaluation (compare d_1 in the top panel and d_2 in the middle panel). The consensus path (bottom) provides an optimal solution based on the convex combination of the structure mappings shown in the top and middle plots in such a way that the entropy of the final mapping is minimized.

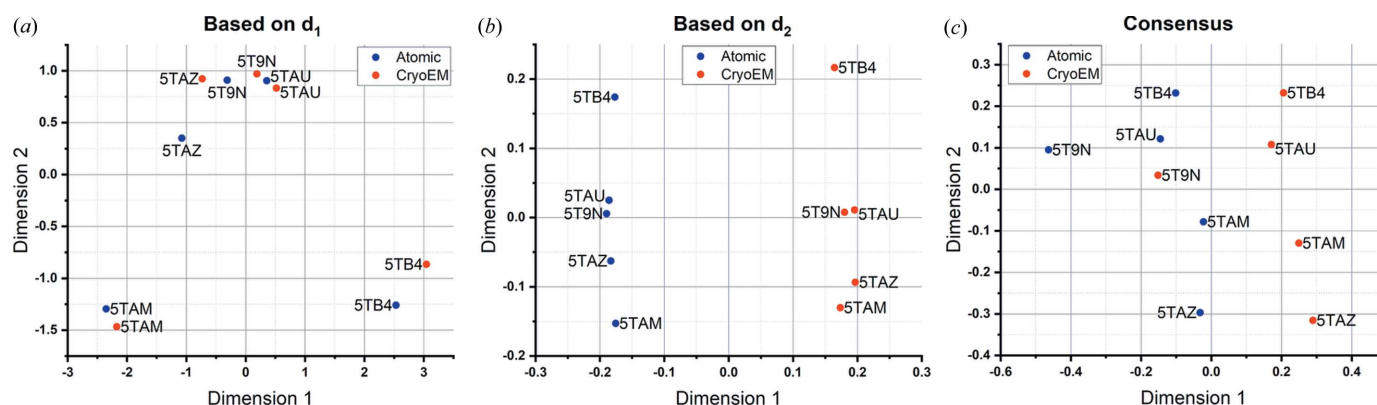


Figure 4

Results obtained after applying the *Zernike3D* algorithm to a set of cryo-EM maps from the ryanodine receptor 1 (RyR1). The data set was constructed in such a way that there are always two maps corresponding to the same conformational state: an experimental cryo-EM map and a cryo-EM map simulated from the atomic structure associated with the previous experimental map. (a) A structure map obtained when comparing experimental cryo-EM maps (red dots) and atomic models (blue dots) for RyR1 through the deformation distance d_1 . The results show that the method succeeded in recovering most of the pairs defined by the experimental cryo-EM maps and atomic structures. (b) A structure map obtained when comparing experimental cryo-EM maps (red dots) and atomic models (blue dots) through the correlation distance. In this case, the correlation metric fails to recover the pairs but it identifies correctly the two different map types used for this analysis. (c) A consensus structure map resulting from the combination of (a) and (b). The consensus provides an optimal solution that helps to identify the map pairs and the map types by keeping a similar structural relationship in the blue and red branches. In these cases, none of these approaches are sufficient for creating a meaningful structure map based on conformation alone, leading us to apply the improvement in Fig. 5.

(specifically, to the lowest of the reported resolutions of the cryo-EM maps). Note that without applying this low-pass filter the minimization process of equation (1) might not reach a meaningful minimum. Finally, we applied the method presented in this work to this combined data set.

Our results, shown in Fig. 4, report the main difficulties that appear when mixing simulated and experimental cryo-EM maps. While the structure map based on d_1 (the distance based on the amount of deformation) illustrates that many pairs are correctly placed together, the structure map based on d_2 (the distance based on the similarity after undeforming) discriminated between maps derived from atomic models and maps coming from cryo-EM experiments. However, the point of this example was to intermix maps from different origins, so discrimination by origin was to be minimized, requiring a further adjustment to our approach. To tackle this problem, we extended our methodology by analysing separately the sub-blocks of the distance matrix including only atomic or only cryo-EM maps [see Fig. 5 (top)]. We thus performed the MDS of each one of the sub-blocks independently, obtaining the low-dimensional mappings X_{AA} and X_{CC} (the subscript indicates whether it corresponds to atomic/computational or cryo-EM/experimental maps). These two low-dimensional mappings were the input into the consensus procedure described in Section 2.4. Focusing on the consensus, we can see that the information provided by the two mappings X_{AA} and X_{CC} is combined into two different trajectories corresponding to each dimension in the distance matrices (simulated and experimental cryo-EM maps) that show a similar distance relationship among their points, illustrating that both trajectories correspond to the same states of RyR1. Therefore, the counterpart of each other, and their relative distances/positions, are retained [Fig. 5 (bottom)].

3.4. Experiment 4: application of the deformation field to atomic models of the CCT complex

We described our deformation field \mathbf{g}_L as a function that deforms V_1 to let it become similar to V_2 , that is, as we have done in previous cases, acting only on two cryo-EM maps. However, since the deformation field is defined in the coordinate system of V_1 , it can also be applied to atomic models defined in the same coordinate system and not only to maps. In this way, we can also deform an atomic model defined for V_1 and use it as starting point for a model of V_2 . Obviously, since the new atomic model defined in the coordinate system of V_2 has been constructed purely based on geometrical considerations, all the stereochemical constraints have to be further imposed.

An example of an atomic model deformed following the previous procedure is presented in Fig. 6. The example was taken from the same data set as used in Experiment 2, which shows an open–closed transition of the CCT complex. The figure illustrates how the deformation applied to the atomic model of the open conformation results in an approximation to the closed conformation. Naturally, we can now compare this deformed model representing the closed conformation with the one obtained directly from the experimental map of the closed conformation. The r.m.s.d. (computed with *ChimeraX*) between these two models was 5.29 Å, certainly high, but substantially reduced compared with that between the open–closed models without applying any deformation, which was 7.90 Å. This r.m.s.d. reduction suggests that the deformation applied is appropriately reproducing a conformational change in the right direction, from open towards the closed state.

However, the overall scores obtained for the two deformed structures still show a high value, as many stereochemical

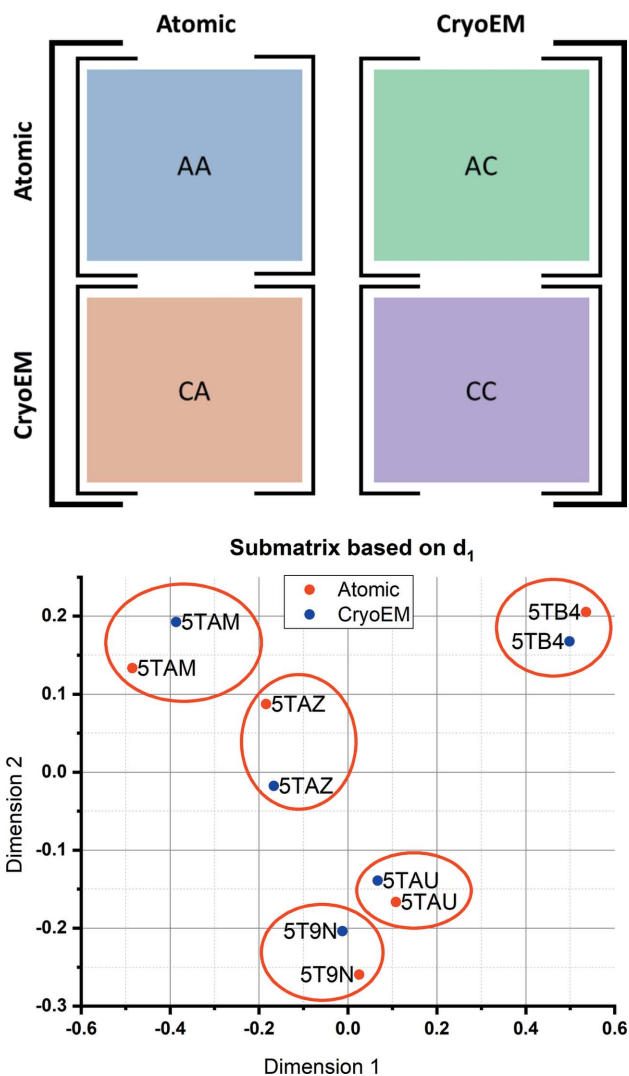


Figure 5 Results obtained after applying the *Zernike3D* algorithm to a set of cryo-EM maps from the ryanodine receptor 1 (RyR1) followed by a decomposition of the distance matrix computed by the algorithm into different blocks to recover more meaningful structure mappings. (Top) A partition of the distance matrix into 2×2 blocks. Each block stores the distances obtained when comparing the different map types used in this test (pairs of experimental cryo-EM maps and maps derived from atomic structures representing the same conformational state): AA (atomic versus atomic), AC (atomic versus cryo-EM), CA (cryo-EM versus atomic) and CC (cryo-EM versus cryo-EM). (Bottom) A consensus structure map for pairs of RyR1 conformations (from atomic model-derived simulated maps and from cryo-EM maps) resulting from the analysis of the blocks. The red circles are used to enhance the visualization of the different pairs. When compared with Fig. 4(a), it is possible to see that this decomposition of the distance matrix leads to a proper recovering of all the pairs found in the data set.

features are not taken into account when computing the deformations. In order to improve the geometry of the deformed structures, we applied a real-space refinement [executing *Phenix* software (Liebscherm *et al.*, 2019) with the default parameters] to the predicted structures using their respective electron-density maps. After this refinement, the r.m.s.d. value measured before decreased further to 4.52 Å. As a conclusion, the combination of deformation and refinement

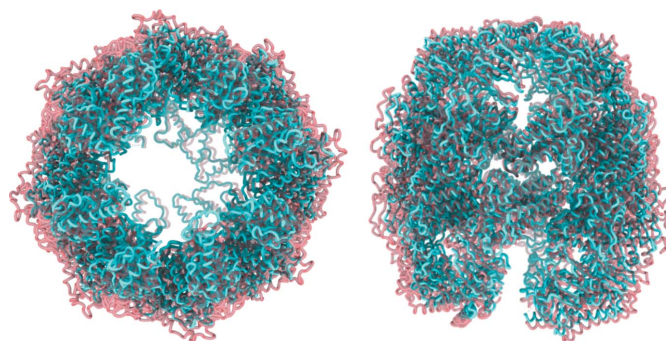


Figure 6 Deformation applied to one of the 30 CCT models obtained by the NMA-based approach called adaptive ANM described in Experiment 2. The deformation was computed using the cryo-EM maps simulated from the 30 models. The original atomic structure in the open state is shown in pink and the deformed version in the closed state in cyan. The results show that the deformation coefficients $\alpha_{l,n,m}$ computed with maps can be effectively applied to the atomic space of the model to approximate geometrically the conformation represented by the cryo-EM map at the level of atoms.

of atomic structures enables us to achieve predictions of different structure conformations on the path between two end points, suitable for performing further studies, though there is clearly room for improvement. For example, refining in between smaller deformations could be of benefit, *e.g.* in hybrid simulation methods where local refinement/simulation complements global deformations (Krieger *et al.*, 2020).

4. Conclusions

The development of automatic algorithms to study continuous flexibility presented in this work results in simplified yet precise procedures, avoiding the need for user interference with the software and increasing the reproducibility of the results. It is also a significant step forward with respect to approaches like PCA and NMA of cryo-EM maps, avoiding the need to select components or modes and producing localized analysis.

The way this new approach works is by defining a new 3D basis where all deformation occurs. It is conceptually similar to the Fourier transform. The movements defining a transition between two different conformational states are decomposed into different components (that can be regarded as low-, medium- and high-frequency movements). Those components will depend on the degree of the basis used in the calculations. The displacements needed along each different component to minimize the distances between two electron-density maps are stored in a series of deformation coefficients $\alpha_{l,n,m}$, which can be further analysed to obtain the local strains and rotations undergone by the macromolecules during conformational transitions. The new approach thus unifies two of our previous developments (NMA and strain/rotation component extraction) for the analysis of continuous heterogeneity.

Apart from the information extracted from the deformation coefficients, our method allows for the definition of a distance measure based on the deformed electron-density maps, which

Table 1
List of real-valued spherical harmonics $y_l^m(\mathbf{r}/|\mathbf{r}|)$.

Degree (l)	Order (m)						
	-3	-2	-1	0	1	2	3
0				$\frac{1}{2} \left(\frac{1}{\pi}\right)^{1/2}$			
1			$\left(\frac{3}{4\pi}\right)^{1/2} \frac{y}{r}$	$\left(\frac{3}{4\pi}\right)^{1/2} \frac{z}{r}$	$\left(\frac{3}{4\pi}\right)^{1/2} \frac{x}{r}$		
2		$\frac{1}{2} \left(\frac{15}{\pi}\right)^{1/2} \frac{xy}{r^2}$	$\frac{1}{2} \left(\frac{15}{\pi}\right)^{1/2} \frac{yz}{r^2}$	$\frac{1}{4} \left(\frac{5}{\pi}\right)^{1/2} \times \frac{-x^2 - y^2 + 2z^2}{r^2}$	$\frac{1}{2} \left(\frac{15}{\pi}\right)^{1/2} \frac{xz}{r^2}$	$\frac{1}{4} \left(\frac{15}{\pi}\right)^{1/2} \frac{x^2 - y^2}{r^2}$	
3	$\frac{1}{4} \left(\frac{35}{2\pi}\right)^{1/2} \times \frac{(3x^2 - y^2)y}{r^3}$	$\frac{1}{2} \left(\frac{105}{\pi}\right)^{1/2} \frac{xyz}{r^3}$	$\frac{1}{4} \left(\frac{21}{2\pi}\right)^{1/2} \times \frac{y(4z^2 - x^2 - y^2)}{r^3}$	$\frac{1}{4} \left(\frac{7}{\pi}\right)^{1/2} \times \frac{z(2z^2 - 3x^2 - 3y^2)}{r^3}$	$\frac{1}{4} \left(\frac{21}{2\pi}\right)^{1/2} \times \frac{x(4z^2 - x^2 - y^2)}{r^3}$	$\frac{1}{2} \left(\frac{105}{\pi}\right)^{1/2} \times \frac{(x^2 - y^2)z}{r^3}$	$\frac{1}{4} \left(\frac{35}{2\pi}\right)^{1/2} \times \frac{(x^2 - 3y^2)x}{r^3}$

is useful for building distance matrices. These distance matrices can be used afterwards to recover structure mappings that show the structural relationships existing among the diverse conformational states. Different definitions of the distance measures may focus on different aspects of the comparison. For this reason, we have devised a new procedure to combine several low-dimensional mappings into a single consensus mapping based on a minimum entropy criterion that tends to produce well ordered low-dimensional mappings and outperforms the results obtained by individual distance metrics.

The possibility of converting atomic models back to electron densities opens the possibility of a combined analysis on maps and models in the same conformational space. An illustrative example has been provided in Experiment 4, where cryo-EM maps, together with their respective structural models between two end points, have been represented in the same space as a set of experimental cryo-EM maps.

In the future, it may be interesting to explore alternative bases for the deformation field and the distance between volumes (like the Wasserstein distance).

APPENDIX A 3D real-valued generalized Zernike polynomials

In this section we discuss the functions that we use as basis functions of the deformations in the unit ball B . We use the generalized Zernike polynomials defined on the 3D ball; in Appendix B1, we briefly review the relation of this basis to the better-known 2D form of Zernike polynomials.

In general, the expansion of any real valued function $g(\mathbf{r}) \in L_2(B)$ in this basis is defined by the formula

$$g(\mathbf{r}) = \sum_{l=0}^{\infty} \sum_{n=0}^{\infty} \sum_{m=-l}^l \alpha_{l,n,m} Z_{l,n,m}(\mathbf{r}), \quad (15)$$

where $\alpha_{l,n,m}$ are real-valued coefficients, and $Z_{l,n,m}(\mathbf{r})$ are the 3D real-valued (normalized) generalized Zernike polynomials defined by the formula

$$Z_{l,n,m}(\mathbf{r}) = \bar{R}_{l,n}^1(r) y_l^m(\theta, \phi), \quad (16)$$

where r is the radial component of the 3D coordinate \mathbf{r} , θ and ϕ are its polar and azimuthal angles, respectively, in spherical coordinates, n and l are non-negative integers, and m is an integer such that $-l \leq m \leq l$. We refer to l as the *spherical frequency*. y_l^m is the real-valued spherical harmonic defined by the formula

$$y_l^m(\theta, \phi) = (-1)^m \left[\frac{2l+1}{4\pi} \frac{(l-|m|)!}{(l+|m|)!} \right]^{1/2} P_l^{|m|} \times (\cos \theta) \begin{cases} 1 & \text{if } m = 0 \\ 2^{1/2} \cos(m\phi) & \text{if } m > 0 \\ 2^{1/2} \sin(|m|\phi) & \text{if } m < 0 \end{cases}, \quad (17)$$

where P_l^m are the associated Legendre polynomials defined by the formula

$$P_l^m(x) = \frac{(-1)^m}{2^l l!} (1-x^2)^{m/2} \frac{d^{l+m}}{dx^{l+m}} (x^2-1)^l. \quad (18)$$

The real and imaginary parts of the complex-valued spherical harmonics are available in standard textbooks such as that by Abramowitz & Stegun (1966). For completeness, Table 1 shows these spherical harmonics in Cartesian coordinates. Before defining the normalized generalized radial Zernike polynomials denoted by $\bar{R}_{l,n}^p$ above, we define the (unnormalized) generalized radial Zernike polynomials $R_{l,n}^1$ as follows:

Table 2
Generalized and normalized radial Zernike polynomials.

$\bar{R}_{0,0}^1(r) = 3^{1/2}$	$\bar{R}_{0,1}^1(r) = 7^{1/2} \left(\frac{5r^2}{2} - \frac{3}{2} \right)$	$\bar{R}_{0,2}^1(r) = (11)^{1/2} \left(\frac{63r^4}{8} - \frac{35r^2}{4} + \frac{15}{8} \right)$	$\bar{R}_{0,3}^1(r) = (15)^{1/2} \left(\frac{429r^6}{16} - \frac{693r^4}{16} + \frac{315r^2}{16} - \frac{35}{16} \right)$	$\bar{R}_{0,4}^1(r) = (19)^{1/2} \left(\frac{12155r^8}{128} - \frac{6435r^6}{32} + \frac{9009r^4}{64} - \frac{1155r^2}{32} + \frac{315}{128} \right)$
$\bar{R}_{1,0}^1(r) = (5^{1/2})r$	$\bar{R}_{1,1}^1(r) = \frac{21r^3}{2} - \frac{15r}{2}$	$\bar{R}_{1,2}^1(r) = (13)^{1/2} \left(\frac{99r^5}{8} - \frac{63r^3}{4} + \frac{35r}{8} \right)$	$\bar{R}_{1,3}^1(r) = (17)^{1/2} \left(\frac{715r^7}{16} - \frac{1287r^5}{16} + \frac{693r^3}{16} - \frac{105r}{16} \right)$	$\bar{R}_{1,4}^1(r) = (21)^{1/2} \left(\frac{20995r^9}{128} - \frac{12155r^7}{32} + \frac{19305r^5}{64} - \frac{3003r^3}{32} + \frac{1155r}{128} \right)$
$\bar{R}_{2,0}^1(r) = (7^{1/2})r^2$	$\bar{R}_{2,1}^1(r) = (11)^{1/2} \left(\frac{9r^4}{2} - \frac{7r^2}{2} \right)$	$\bar{R}_{2,2}^1(r) = (15)^{1/2} \left(\frac{143r^6}{8} - \frac{99r^4}{4} + \frac{63r^2}{8} \right)$	$\bar{R}_{2,3}^1(r) = (19)^{1/2} \left(\frac{1105r^8}{16} - \frac{2145r^6}{16} + \frac{1287r^4}{16} - \frac{231r^2}{16} \right)$	$\bar{R}_{2,4}^1(r) = (23)^{1/2} \left(\frac{33915r^{10}}{128} - \frac{20995r^8}{32} + \frac{36465r^6}{64} - \frac{6435r^4}{32} + \frac{3003r^2}{128} \right)$
$\bar{R}_{3,0}^1(r) = 3r^3$	$\bar{R}_{3,1}^1(r) = (13)^{1/2} \left(\frac{11r^5}{2} - \frac{9r^3}{2} \right)$	$\bar{R}_{3,2}^1(r) = (17)^{1/2} \left(\frac{195r^7}{8} - \frac{143r^5}{4} + \frac{99r^3}{8} \right)$	$\bar{R}_{3,3}^1(r) = (21)^{1/2} \left(\frac{1615r^9}{16} - \frac{3315r^7}{16} + \frac{2145r^5}{16} - \frac{429r^3}{16} \right)$	$\bar{R}_{3,4}^1(r) = \frac{260015r^{11}}{128} - \frac{169575r^9}{32} + \frac{314925r^7}{64} - \frac{60775r^5}{32} + \frac{32175r^3}{128}$
$\bar{R}_{4,0}^1(r) = (11)^{1/2}r^4$	$\bar{R}_{4,1}^1(r) = (15)^{1/2} \left(\frac{13r^6}{2} - \frac{11r^4}{2} \right)$	$\bar{R}_{4,2}^1(r) = (19)^{1/2} \left(\frac{255r^8}{8} - \frac{195r^6}{4} + \frac{143r^4}{8} \right)$	$\bar{R}_{4,3}^1(r) = (23)^{1/2} \left(\frac{2261r^{10}}{16} - \frac{4845r^8}{16} + \frac{3315r^6}{16} - \frac{715r^4}{16} \right)$	$\bar{R}_{4,4}^1(r) = 3(3^{1/2}) \left(\frac{76475r^{12}}{128} - \frac{52003r^{10}}{32} + \frac{101745r^8}{64} - \frac{20995r^6}{32} + \frac{12155r^4}{128} \right)$
$\bar{R}_{5,0}^1(r) = (13)^{1/2}r^5$	$\bar{R}_{5,1}^1(r) = (17)^{1/2} \left(\frac{15r^7}{2} - \frac{13r^5}{2} \right)$	$\bar{R}_{5,2}^1(r) = (21)^{1/2} \left(\frac{323r^9}{8} - \frac{255r^7}{4} + \frac{195r^5}{8} \right)$	$\bar{R}_{5,3}^1(r) = (25)^{1/2} \left(\frac{3059r^{11}}{16} - \frac{6783r^9}{16} + \frac{4845r^7}{16} - \frac{1105r^5}{16} \right)$	$\bar{R}_{5,4}^1(r) = (29)^{1/2} \left(\frac{108675r^{13}}{128} - \frac{76475r^{11}}{32} + \frac{156009r^9}{64} - \frac{33915r^7}{32} + \frac{20995r^5}{128} \right)$

$$R_{l,n}^p(x) = (-1)^n x^l P_n^{[l+(p/2), 0]}(1 - 2x^2), \quad (19)$$

$$\bar{R}_{l,n}^p(x) = 2^{1/2} \left[2n + l + \frac{p}{2} + 1 \right]^{1/2} R_{l,n}^p(x). \quad (21)$$

where $P_n^{(\alpha,\beta)}$ are the Jacobi polynomials,

$$P_n^{(\alpha,\beta)}(x) = \frac{(-1)^n}{2^n n!} (1-x)^{-\alpha} (1+x)^{-\beta} \times \frac{d^n}{dx^n} [(1-x)^\alpha (1+x)^\beta (1-x^2)^n]. \quad (20)$$

The definitions and properties of the standard associated Legendre polynomials and Jacobi polynomials are available, *inter alia*, in the book by Abramowitz & Stegun (1966).

Finally, while the radial polynomials are orthogonal (with the appropriate norm), they are not orthonormal. This is easily corrected by replacing the radial polynomials with the normalized generalized radial Zernike polynomials, denoted by $\bar{R}_{l,n}^p$,

The parameter p is associated with the choice of inner product and the dimensionality of the balls; in the case of a 3D ball, the natural choice of p is $p = 1$, which yields the basis in (16) which is orthonormal in the natural inner product on $L_2(B)$. We give in Table 2 the explicit list of radial functions $\bar{R}_{l,n}^1$ that we use.

We recognize that our choice of basis functions for the expansion is certainly not the only possible choice. We used the Zernike polynomials (in the generalized form presented here) to obtain an expansion of functions in a ball which do not vanish at the boundaries. Our use of Zernike polynomials also yields a basis that is closed under rotations. The graphical representation of some components of the basis is shown in Fig. 7.

APPENDIX B

Properties and relationships of this basis

B1. Properties of the polynomials involved – Zernike polynomials

We note that slightly different definitions and normalization are used in different sources; the most commonly used form of Zernike polynomials is associated with 2D functions on the unit disc, whereas we are interested in 3D functions on the unit ball. The better-known traditional radial Zernike polynomials, denoted here by $\tilde{R}_m^l(x)$, are a special case of the generalized radial Zernike polynomials $R_{n,l}^p(x)$,

$$\tilde{R}_l^m(x) = R_{m,(l-m)/2}^0(x), \tag{22}$$

with $\tilde{R}_l^m(x) = 0$ if $l - m$ is odd or if $m > l$. The definition of the 3D real-valued Zernike polynomial [equation (16)] is analogous to the definition of the traditional 2D Zernike polynomials Z_l^m on the unit disc,

$$Z_l^m(r, \phi) = \tilde{R}_l^{|m|}(r) \begin{cases} \cos(m\phi) & \text{if } m \geq 0, \\ \sin(|m|\phi) & \text{if } m < 0. \end{cases} \tag{23}$$

Unfortunately, the common notation for the 2D and 3D cases can be misleading: the parameter m here plays the role of the parameter l in the analogous 3D case; the parameter m in the 3D case is related to the existence of both sine and cosine for each m here, but does not otherwise have an immediate counterpart here.

Some properties of these generalized Zernike polynomials, including the higher-dimensional cases, are discussed in further detail by Slepian (1964), Serkh (2015), Greengard & Serkh (2018) and Lederman (2017).

The radial Zernike polynomials in equation (19) are orthogonal with respect to the inner product,

$$\langle f(x), g(x) \rangle = \int_0^1 x^{p+1} f(x) g(x) dx, \tag{24}$$

so that $\langle R_{l,n_1}^p(x), R_{l,n_2}^p(x) \rangle = 0$ if $n_1 \neq n_2$. Note that they are not necessarily orthogonal for different l , that is, $\langle R_{l_1,n}^p(x), R_{l_2,n}^p(x) \rangle$ is not 0, in general. It follows that the Zernike polynomials $Z_{l,n,m}$ (16) are orthogonal (across all different combinations of n, l and m) on the natural inner product on the unit ball.

B2. A remark on numerical evaluation

As is the case with many orthogonal polynomials, the direct computation using the explicit sum of monomials is generally unstable and not recommended in numerical computation. However, in this work, since we truncate the polynomials at low n , the explicit form has been found experimentally to be sufficiently stable. For additional details on computation see Lederman (2017).

B3. Closure under rotations

We recall that we use basis functions defined in equation (15), which are composed of a radial component and an angular component. Furthermore, we truncate the expansion

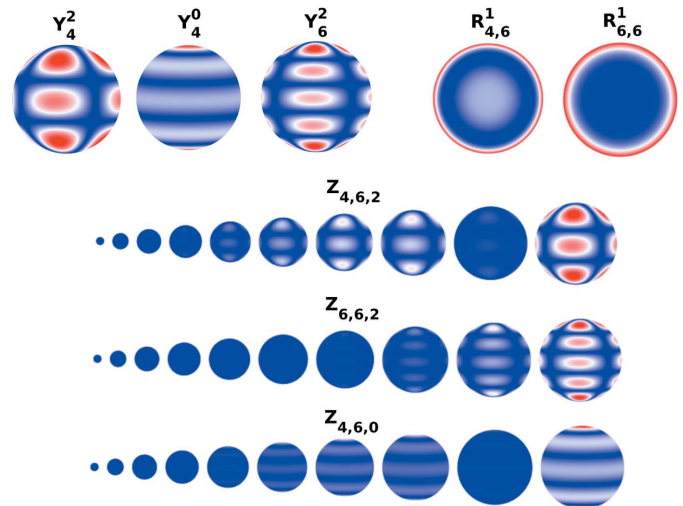


Figure 7

Representation of some components of the basis $Z_{l,n,m}$ regarding their former angular and radial components Y_l^m and $R_{l,n}^1$. Since the spherical harmonics Y_l^m are only defined on the surface of the sphere, the representation of the basis components $Z_{l,n,m}$ includes several spheres whose radius is contained in the interval $[0, 1]$ to have a better graphical representation of the whole component. The real component would be obtained by stacking all the spheres (whose radii belong to the interval $[0, 1]$) concentrically. Each point in the three representations corresponds to the value obtained when evaluating the corresponding functions on a grid.

in equation (2) such that if $Z_{l,n,m}(\mathbf{r})$ is in the expansion, then $Z_{l,n,m'}(\mathbf{r})$ is also in the expansion for any $-l \leq m, m' \leq m$. As is well known, the rotation of the frame of reference of spherical harmonics of a given spatial frequency l is a unitary operation and closed rotations. It follows that the linear combination of $\sum_{n=0}^N \sum_{l=0}^L \sum_{m=-l}^l \alpha_{l,n,m}^x Z_{l,n,m}(\mathbf{r})$ is closed under rotations. In other words, regardless of the frame of axis we choose for our spherical harmonics, we can represent the same functions using our choice of basis.

We recall that the deformation field $\mathbf{g}_L(\mathbf{r})$ is a 3D vector defined in equation (2) at any point r . If we rotate the axes x, y and z we must rotate the vector $\mathbf{g}_L(\mathbf{r})$ to obtain a vector in the new coordinate system. As is well known,

$$A \mathbf{g}_L(\mathbf{A}^{-1} \mathbf{r}) = \sum_{n=0}^N \sum_{l=0}^L \sum_{m=-l}^l A \begin{pmatrix} \alpha_{l,n,m}^x \\ \alpha_{l,n,m}^y \\ \alpha_{l,n,m}^z \end{pmatrix} \tilde{Z}_{l,n,m}(\mathbf{A}^{-1} \mathbf{r}), \tag{25}$$

where A is the appropriate unitary rotation matrix.

It follows that the basis we have chosen is closed under rotations; regardless of the orientation of the frame of reference we choose (but depending on the position of the centre), we can represent the same fields. Furthermore, the transformation between frames of reference is unitary. We note that, since the volume is defined on the grid, the problem definition is not entirely closed under rotations, although the definition of the field is closed under rotations.

APPENDIX C

Complex-valued Zernike 3D basis

We may extend our basis function to a complex-valued Zernike 3D basis. This second basis uses spherical harmonics, which are well known basis functions for functions defined over the surface of a unit sphere. For doing so, we should define $Z'_{l,n,m}$ [see equation (16)],

$$\bar{Z}'_{l,n,m}(\mathbf{r}) = \bar{R}'_{l,n}(r) Y_l^m(\theta, \phi), \tag{26}$$

where $Y_l^m(\theta, \phi)$ are the standard spherical harmonics,

$$Y_l^m(\theta, \phi) = \left[\frac{2l+1}{4\pi} \frac{(l-m)!}{(l+m)!} \right]^{1/2} P_l^m(\cos \theta) \exp(im\phi). \tag{27}$$

It is well known that the spherical harmonics are a complete orthonormal basis on the surface of the unit sphere such that

$$\int_{\Omega} Y_{l_1}^{m_1}(\mathbf{r}) [Y_{l_2}^{m_2}(\mathbf{r})]^* d\mathbf{r} = \begin{cases} 1 & \text{if } l_1 = l_2, m_1 = m_2, \\ 0 & \text{otherwise,} \end{cases} \tag{28}$$

where Ω is the surface of the unit sphere.

In this new basis the expansion would be expressed as

$$g(\mathbf{r}) = \sum_{n=0}^{\infty} \sum_{l=0}^{\infty} \sum_{m=-l}^l \beta_{l,n,m} \bar{Z}'_{l,n,m}(\mathbf{r}). \tag{29}$$

The relationship between the expansion that uses the real-valued basis functions and the expansion that uses the complex-valued basis functions is

$$\alpha_{l,n,m} = \begin{cases} (-1)^{m+1} \frac{i}{2^{1/2}} (\beta_{l,n,m} - \beta_{l,n,-m}) & \text{if } m < 0, \\ \beta_{l,n,0} & \text{if } m = 0, \\ (-1)^m \frac{1}{2} (\beta_{l,n,m} + \beta_{l,n,-m}) & \text{if } m > 0, \end{cases} \tag{30}$$

and for the basis functions

$$\bar{Z}'_{l,n,m}(\mathbf{r}) = \begin{cases} \frac{i}{2^{1/2}} [\bar{Z}'_{l,n,m}(\mathbf{r}) - (-1)^m \bar{Z}'_{l,n,-m}(\mathbf{r})] & \text{if } m < 0, \\ \bar{Z}'_{l,n,0}(\mathbf{r}) & \text{if } m = 0, \\ \frac{1}{2^{1/2}} [\bar{Z}'_{l,n,-m}(\mathbf{r}) + (-1)^m \bar{Z}'_{l,n,-m}(\mathbf{r})] & \text{if } m > 0. \end{cases} \tag{31}$$

Using the fact that the radial Zernike polynomials are a complete radial orthonormal basis, and the fact that spherical harmonics are a complete orthonormal basis on the the surface of a sphere, one can show that the generalized Zernike polynomials are a complete orthonormal basis of function on the unit ball with respect to the natural L^2 norm on the unit ball.

Suppose that \mathbf{r} is a 3D coordinate of a point in real space and \mathbf{R} is a 3D frequency. Then the Fourier transform of the complex-valued basis function would be

$$\begin{aligned} (\mathcal{F}Z'_{l,n,m})(\mathbf{R}) &= \int_{\mathbf{r} \in B} \exp(-i(\mathbf{R}, \mathbf{r})) Z'_{l,n,m}(\mathbf{r}) d\mathbf{r} \\ &= Y_l^m\left(\frac{\mathbf{R}}{R}\right) \frac{1}{i^{l(2\pi)^{l+1/2}}} \frac{(-1)^n J_{2n+l+(1/2)+1}(R)}{R}, \end{aligned} \tag{32}$$

where (\mathbf{r}, \mathbf{R}) is the usual inner product in \mathbb{R}^3 , R is the modulus of \mathbf{R} and $J_{\alpha}(x)$ is the Bessel function of the first kind and order α .

Funding information

The project that gave rise to these results received the support of a fellowship from ‘la Caixa’ Foundation (No. LCF/BQ/DI18/11660021). This project has received funding from the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement (No. 713673) and the European Regional Development Fund Project ‘CERIT Scientific Cloud’ (No. CZ.02.1.01/0.0/0.0/16_013/0001802). This work was also partially supported by the National Institutes of Health (P41GM103712 to IB), by a MolSSI Seed Software Fellowship (to JK) and by the NIH/NIGMS (No. 1R01GM136780-01 to RRL). Funding is also acknowledged from the Ministry of Science and Innovation through grants: SEV 2017-0712, PID2019-104757RB-I00/AEI48/10.13039/501100011033; the ‘Comunidad Autonoma de Madrid’ through grant: S2017/BMD-3817; and the European Union (EU) and Horizon 2020 through grant: HighResCells (ERC - 2018 - SyG, Proposal: 810057) and iNEXT-Discovery (Proposal: 871037).

References

Abramowitz, M. & Stegun, I. A. (1966). *Handbook of Mathematical Functions*. New York: National Bureau of Standards.

Abrishami, V., Vargas, J., Li, X., Cheng, Y., Marabini, R., Sorzano, C. S. & Carazo, J. M. (2015). *J. Struct. Biol.* **189**, 163–176.

Adcock, S. A. & McCammon, J. A. (2006). *Chem. Rev.* **106**, 1589–1615.

Amunts, A., Brown, A., Toots, J., Scheres, S. H. W. & Ramakrishnan, V. (2015). *Science*, **348**, 95–98.

Atilgan, A. R., Durell, S. R., Jernigan, R. L., Demirel, M. C., Keskin, O. & Bahar, I. (2001). *Biophys. J.* **80**, 505–515.

Bahar, I., Lezon, T. R., Yang, L. W. & Eyal, E. (2010). *Annu. Rev. Biophys.* **39**, 23–42.

Carroni, M. & Saibil, H. R. (2016). *Methods*, **95**, 78–85.

Cong, Y., Schröder, G. F., Meyer, A. S., Jakana, J., Ma, B., Dougherty, M. T., Schmid, M. F., Reissmann, S., Levitt, M., Ludtke, S. L., Frydman, J. & Chiu, W. (2012). *EMBO J.* **31**, 720–730.

Cui, Q. & Bahar, I. (2006). *Normal Mode Analysis: Theory and Applications to Biological and Chemical Systems*. Boca Raton: CRC press.

Doruker, P., Atilgan, A. R. & Bahar, I. (2000). *Proteins*, **40**, 512–524.

Eyal, E., Yang, L. W. & Bahar, I. (2006). *Bioinformatics*, **22**, 2619–2627.

Greengard, P. & Serkh, K. (2018). arXiv:1811.02733.

Härdle, W. K. & Simar, L. (2012). *Multidimensional Scaling*, pp. 397–412. Heidelberg: Springer.

Jonić, S. & Sanchez Sorzano, C. O. (2016). *IEEE J. Sel. Top. Signal. Process.* **10**, 161–173.

Krieger, J. M., Doruker, P., Scott, A. L., Perahia, D. & Bahar, I. (2020). *Curr. Opin. Struct. Biol.* **64**, 34–41.

Lederman, R. R. (2017). arXiv:1710.02874.

Liebscherm, D., Afonine, P. V., Baker, M. L., Bunkóczi, G., Chen, V. B., Croll, T. I., Hintze, B., Hung, L. W., Jain, S., McCoy, A. J., Moriarty, N. W., Oeffner, R. D., Poon, B. K., Prisant, M. G., Read, R. J., Richardson, J. S., Sammito, M. D., Sobolev, O. V., Stockwell, D. H., Terwilliger, T. C., Urzhumtsev, A. G., Videau, L. L., Williams, C. J. & Adams, P. D. (2019). *Acta Cryst.* **65**, 861–877.

Provencher, S. W. & Vogel, R. H. (1988). *Ultramicroscopy*, **25**, 209–221.

Rosa-Trevín, J. M. de la, Otón, J., Marabini, R., Zaldívar, A., Vargas, J., Carazo, J. M. & Sorzano, C. O. S. (2013). *J. Struct. Biol.* **184**, 321–328.

- Rosa-Trevín, J. M. de la, Quintana, A., del Cano, L., Zaldívar, A., Foche, I., Gutiérrez, J., Gómez-Blanco, J., Burguet-Castell, J., Cuenca-Alba, J., Abrishami, V., Vargas, J., Otón, J., Sharov, G., Vilas, J. L., Navas, J., Conesa, P., Kazemi, M., Marabini, R., Sorzano, C. O. S. & Carazo, J. M. (2016). *J. Struct. Biol.* **195**, 93–99.
- Sanchez Sorzano, C. O., Alvarez-Cabrera, A. L., Kazemi, M., Carazo, J. M. & Jonić, S. (2016). *Biophys. J.* **110**, 1753–1765.
- Schur, F. K. M. (2019). *Curr. Opin. Struct. Biol.* **58**, 1–9.
- Serh, K. (2015). arXiv:1811.02733.
- Slepian, D. (1964). *Bell Labs Technical J.*, **43**, 3009–3057.
- Sorzano, C. O. S. & Carazo, J. M. (2021). *Acta Cryst. D* **77**, 835–839.
- Sorzano, C. O. S., Jiménez, A., Mota, J., Vilas, J. L., Maluenda, D., Martínez, M., Ramírez-Aportela, E., Majtner, T., Segura, J., Sánchez-García, R., Rancel, Y., del Caño, L., Conesa, P., Melero, R., Jonic, S., Vargas, J., Cazals, F., Freyberg, Z., Krieger, J., Bahar, I., Marabini, R. & Carazo, J. M. (2019). *Acta Cryst. F* **75**, 19–32.
- Sorzano, C. O. S., Martín-Ramos, A., Prieto, F., Melero, R., Martín-Benito, J., Jonic, S., Navas-Calvente, J., Vargas, J., Otón, J., Abrishami, V., de la Rosa-Trevín, J. M., Gómez-Blanco, J., Vilas, J. L., Marabini, R. & Carazo, J. M. (2016). *J. Struct. Biol.* **195**, 123–128.
- Sorzano, C. O. S., Vargas, J., Otón, J., Abrishami, V., de la Rosa-Trevín, J. M., Riego, S., Fernández-Alderete, A., Martínez-Rey, C., Marabini, R. & Carazo, J. M. (2015). *AIMS Biophys.* **2**, 8–20.
- Tagare, H. D., Kucukelbir, A., Sigworth, F. J., Wang, H. & Rao, M. (2015). *J. Struct. Biol.* **191**, 245–262.
- Tegunov, D. & Cramer, P. (2019). *Nat. Methods*, **16**, 1146–1152.
- wwPDB Consortium (2019). *Nucleic Acids Res.* **47**, D520–D528.
- Yang, Z., Májek, P. & Bahar, I. (2010). *PLoS Comput. Biol.* **5**, e1000360.
- Zhang, S., Krieger, J. M., Zhang, Y., Kaya, C., Kaynak, B., Mikulska-Ruminska, K., Doruker, P., Li, H. & Bahar, I. (2021a). *Bioinformatics*, doi: 10.1093/bioinformatics/btab187.
- Zhang, Y., Krieger, J., Mikulska-Ruminska, K., Kaynak, B., Sorzano, C. O. S., Carazo, J. M., Xing, J. & Bahar, I. (2021b). *Prog. Biophys. Mol. Biol.* **160**, 104–120.
- Zheng, S. Q., Palovcak, E., Armache, J. P., Verba, K. A., Cheng, Y. & Agard, D. A. (2017). *Nat. Methods*, **14**, 331–332.

Estimating conformational landscapes from Cryo-EM particles by 3D Zernike polynomials


Herreros, D., Lederman, R.R., Krieger, J.M. et al. Estimating conformational landscapes from Cryo-EM particles by 3D Zernike polynomials. *Nat Commun* 14, 154 (2023). <https://doi.org/10.1038/s41467-023-35791-y>

Estimating conformational landscapes from Cryo-EM particles by 3D Zernike polynomials

Received: 1 June 2022

Accepted: 29 December 2022

Published online: 11 January 2023

 Check for updates

D. Herreros¹✉, R. R. Lederman², J. M. Krieger¹, A. Jiménez-Moreno¹, M. Martínez¹, D. Myška³, D. Strelak^{1,4}, J. Filipovic³, C. O. S. Sorzano^{1,5} & J. M. Carazo^{1,5}

The new developments in Cryo-EM Single Particle Analysis are helping us to understand how the macromolecular structure and function meet to drive biological processes. By capturing many states at the particle level, it is possible to address how macromolecules explore different conformations, information that is classically extracted through 3D classification. However, the limitations of classical approaches prevent us from fully understanding the complete conformational landscape due to the reduced number of discrete states accurately reconstructed. To characterize the whole structural spectrum of a macromolecule, we propose an extension of our Zernike3D approach, able to extract per-image continuous flexibility information directly from a particle dataset. Also, our method can be seamlessly applied to images, maps or atomic models, opening integrative possibilities. Furthermore, we introduce the ZART reconstruction algorithm, which considers the Zernike3D deformation fields to revert particle conformational changes during the reconstruction process, thus minimizing the blurring induced by molecular motions.

Cryo-electron microscopy (Cryo-EM) single particle analysis (SPA)¹ has proven to be a powerful technique to understand the structure of macromolecules. By capturing individual images of the specimen in different poses, it is not only possible to reconstruct the average macromolecular conformation of the specimen under study, but it also brings to light the challenging problem of identifying several conformational states from the acquired dataset.

Generally, compositional heterogeneity, as well as flexibility, have been addressed through 3D classification². This approach allows reconstructing a given number of different states from the particle images based on the assumption that there is a defined number of discrete conformational states being explored by the specimen. This methodology has been very successful in the study of many systems, being recently expanded to increase the number of states being resolved³.

However, the explicit modeling assumption of the existence of discrete motions has obvious limitations in most experimental cases, depending on the actual biological system under study. Clearly, removing this constraint is methodologically very challenging, although the pay-offs are clear, both in terms of obtaining richer conformational landscapes than currently done, and in providing improved algorithmic stability and objectivity, removing many assumptions and trial and error tests.

Limitations faced with discrete flexibility can only be solved at the image processing level by a paradigm change introducing methods able to handle continuous flexibility: the ability to extract macromolecular conformational information at the individual particle level to get a sufficiently rich and populated landscape of molecular states. Several approaches have been previously proposed to face continuous flexibility, each from a different perspective^{4–9}.

¹Centro Nacional de Biotecnología-CSIC, C/Darwin, 3, 28049 Cantoblanco, Madrid, Spain. ²The Department of Statistics and Data Science, Yale University, New Haven, CT, USA. ³Institute of Computer Science, Masaryk University, Botanická 68a, 60200 Brno, Czech Republic. ⁴Faculty of Informatics, Masaryk University, Botanická 68a, 60200 Brno, Czech Republic. ⁵These authors jointly supervised this work: C.O.S. Sorzano, and J.M. Carazo.

✉ e-mail: dherreros@cnb.csic.es

In this work, we extend our recent Zernike3D algorithm¹⁰ (specifically designed to deal with continuous heterogeneity) to precisely accomplish the latter task starting from Cryo-EM images with some unique properties, such as (1) the possibility to work with images, maps, and atomic models in the same space, (2) a clear mathematical design that intrinsically helps avoiding over-deformations in projection directions, and (3) a reconstruction algorithm (that we name ZART - Zernike3D-based Algebraic Reconstruction Technique) that takes into account individual particle conformational information, reverts the structural changes, and obtains a new map in which flexible regions have intrinsically increased resolution. Note that property (1) indicates that one can work at the level of structural models, avoiding multiple fitting steps, property (2) drastically reduces flexibility estimation errors that would be very difficult to consider in other mathematical frameworks, and property (3) makes it possible to explore states with a small number of classes while still reconstructing maps with large datasets, though at improved resolution since motion blurring is substantially reduced.

We note that the full derivation of ZART is rather technical, so we present in this work its main properties in the context of continuous flexibility, while the derivation of the algorithm in itself is presented as a separate technical work.

Results

Conformational landscape of EMPIAR-10028 dataset

The following experiment is aimed at assessing the capacity of the Zernike3D algorithm to identify conformational variability on real Cryo-EM data. To that end, we analyzed the EMPIAR-10028 dataset¹¹ corresponding to the *P. falciparum* 80 S ribosome bound to emetine. This dataset has been extensively studied by other methods^{5,6}, becoming a popular validation dataset for continuous heterogeneity algorithms.

In this work we have reprocessed that dataset inside Scipion¹², leading to a total of 50,000 particles. The workflow followed included several cleaning steps to reduce as much as possible the number of unwanted particles, followed by some consensus protocols to compare the parameters estimated by different algorithms (angular assignment, shifts, Contrast Transfer Function...) and keep only the particles consistently estimated.

The previous particles were subjected to the Zernike3D analysis, translating them to a set of Zernike3D coefficients. The maximum basis degrees were set to $N = 3$ and $L = 2$ for the estimation of the deformation fields. In addition, the particles were downsampled to a box size of 125 voxels to increase the performance of the algorithm. Apart from the Zernike3D analysis, the particles were not subjected to other heterogeneity workflows such as classical 3D classification.

The resulting UMAP (Uniform Manifold Approximation and Projection)¹³ representation of the Zernike3D coefficient space is shown in Fig. 1. As it can be seen from the representation, the Zernike3D coefficient space leads to an informative representation of the heterogeneity present in the dataset. Two clear states are well differentiated, representing the two rotation states of the small subunit of the ribosome, as well as some other more localized movements. The colormap used to represent the embedding describes the amount of deformation associated with each deformation field: purple colors correspond to small deformations and are usually associated with conformational changes similar to the reference map, while yellow colors are associated with bigger changes. The possibility of coloring the coefficient space adds another dimension of information helping in the analysis of the heterogeneity of the dataset.

There are two different possibilities to recover conformational changes from the previous embedding: (1) applying a deformation field to the reference or (2) exploring, by refinement and reconstruction, different areas of the conformational space. Option 1 represents an almost instant and interactive exploration of the conformational

space, in which just by placing the cursor on any point of the representation conformation we obtain a Zernike3D synthesis of a map, while Option 2 goes to the original images and aims at exploring whether there are residual errors not accounted for by Zernike3D. In all cases tested so far, the differences between the two options are minimal, as it is shown in Fig. 2a, b. However, the application of the deformation field leads to a higher resolution representation of the conformational change (equal to the resolution of the reference map), while the refinement resolution is intrinsically limited to the number of particles selected from the space.

In addition, the Zernike3D coefficients extracted from the conformational space can be applied simultaneously to the reference map and to a structural model traced (or aligned) from the reference. This allows obtaining a rigid fitting of the atomic positions that match the conformation of any particle in the dataset. An example of the application of the Zernike3D coefficients to the ribosome atomic structure can be found in Fig. 2c. However, it is worth mentioning that the Zernike3D coefficients are computed exclusively based on geometrical considerations, so the approximated structural models might need to be refined to correct for stereochemistry artifacts. Indeed, it should always be considered that the estimation of the deformation fields describing a given transition only depends on the rigid alignment of the reference towards the conformation represented by a given particle. Therefore, the estimated deformation field does not consider any stereochemistry constrains, which should be posteriorly imposed to avoid atomic clashes or improve Ramachandran outliers among others.

An example of the simultaneous exploration of the coefficient space performed with the reference map and its structural model is provided in Supplementary Movie 1 (we are aware that this and subsequent videos are only graphical means to make more obvious conformational changes, and that they are not to be considered as suggesting molecular trajectories at all). The different states were obtained by grouping the coefficients with KMeans into 5 clusters. Then, the cluster representatives were used to generate the deformed maps/structures, which were afterward morphed with ChimeraX software¹⁴.

The next step we followed in the analysis of the dataset is to use the estimated deformation fields and the particles to reconstruct a higher-resolution map by correcting the conformational changes of each image with ZART. The comparison between the map reconstructed with CryoSparc¹⁵ and ZART reconstruction algorithm is shown

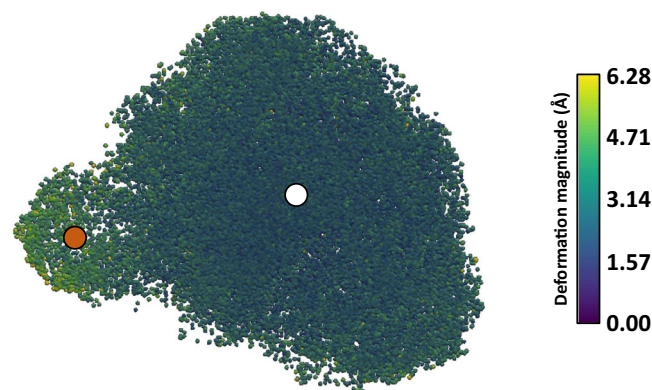


Fig. 1 | EMPIAR-10028 Zernike3D conformational landscape. UMAP representation of the Zernike3D coefficient space for the *P. falciparum* 80 S ribosome (EMPIAR-10028 dataset). The colormap represents the modulus of the deformation field that has to be applied to the reference map to match the conformational state of each particle projection image. Purple colors represent lower deformations (close to the reference state). The representation shows a clear distinction between two different states marked by the white (reference map) and orange (rotated Pf80S state) dots.

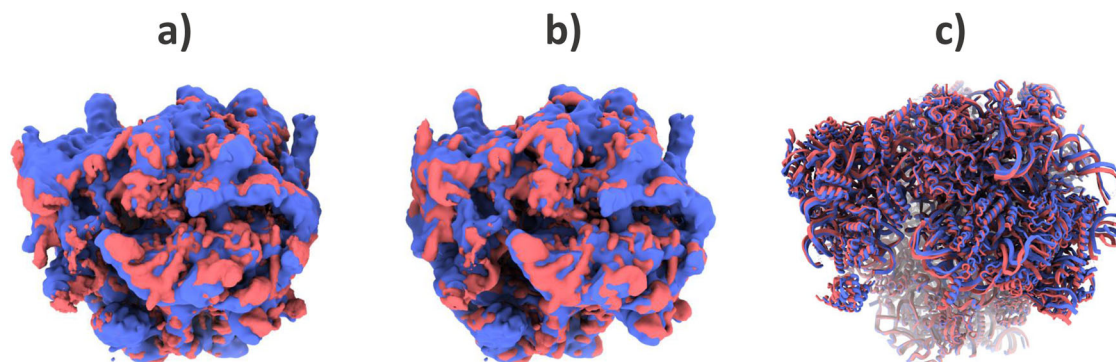


Fig. 2 | Example of Pf80S Zernike3D states. **a** Comparison of the reference conformation required by the Zernike3D algorithm (red) and the rotated Pf80S state recovered from a homogeneous refinement with CryoSparc (blue). **b** Comparison of the rotated Pf80S state recovered from the Zernike3D deformation fields (red) and the rotated Pf80S state recovered from a homogeneous refinement with CryoSparc (blue). The particles processed by CryoSparc are taken from the coefficient space area defined by the orange dot in Fig. 1, and the deformation field is computed with the coefficients associated with this dot. The comparison between the maps displayed in **a** and **b** show that the Zernike3D conformation (**b** - red map)

is consistent with the experimental conformation refined from the particles selected from that region of the coefficient space (blue). In addition, the application of the deformation field does not decrease the resolution of the reference map. **c** Comparison of the atomic structure associated with the Zernike3D reference map (red) and the structure deformed with the Zernike3D deformation fields (blue). Since the Zernike3D can work indistinguishably with maps, atomic structures, and particles, the rotated state can be appropriately reproduced at the atomic level using the deformation fields estimated from the particles.

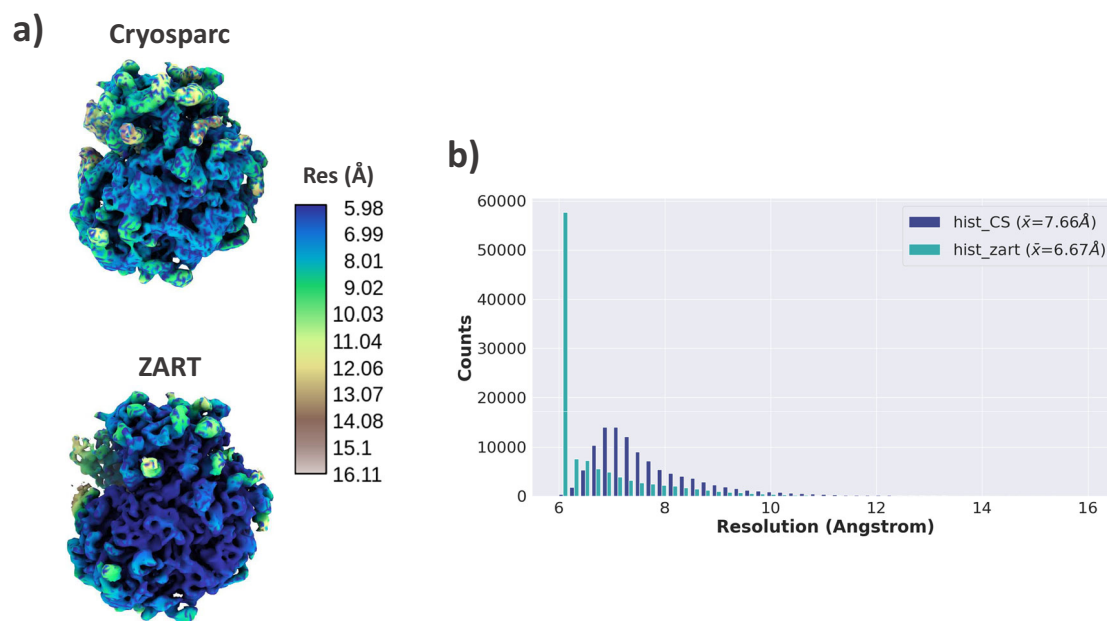


Fig. 3 | Analysis of EMPIAR-10028 ZART reconstruction. **a** Comparison of *P. falciparum* 80 S ribosome map refined with CryoSparc (blue) and the motion-corrected map recover with our ZART algorithm. The colormap represents the local resolution estimation for each voxel computed with BlocRes²³. The ZART reconstructed map shows an overall improvement in resolution thanks to the deformation fields considered during the reconstruction process. **b** Resolution histogram

comparison for CryoSparc and ZART reconstructions obtained from the resolution map computed by BlocRes. The histogram shows a clear displacement of the local resolutions towards higher resolutions. The value provided in the legend of the histograms shows the mean value of the local resolution estimations for both reconstructions.

in Fig. 3a. The comparison of the two maps shows a clear improvement at both, the level of maps (a) and slices (b), in the moving and still areas of the molecule. In order to make a more quantitative comparison of the maps, we computed the local resolution histograms of both reconstructions, which are compared in Fig. 3b. Similarly to the visual inspection of the maps, the resolution histograms confirm the improvement in local resolution, being the average resolution of ZART pushed 1.01 Å compared to the mean resolution of CryoSparc.

Conformational landscape of EMPIAR-10180 dataset

The EMPIAR-10180¹⁶ dataset has become another standard dataset to test continuous heterogeneity algorithms due to the large degree of

flexibility information it contains. The dataset corresponds to a pre-catalytic spliceosome exhibiting an extensive heterogeneity already observed by classical methods such as 3D classification.

Since the Zernike3D algorithm focuses on the analysis of continuous heterogeneity rather than compositional heterogeneity, the dataset was preprocessed inside Scipion¹² to clean as much as possible the original deposited particles. The original dataset is composed of around 320k particles, which were reduced to around 180k after the cleaning steps.

The cleaned particles were afterward subjected to the Zernike3D analysis to extract the different conformational changes suffered by the pre-catalytic spliceosome. As we did in the previous

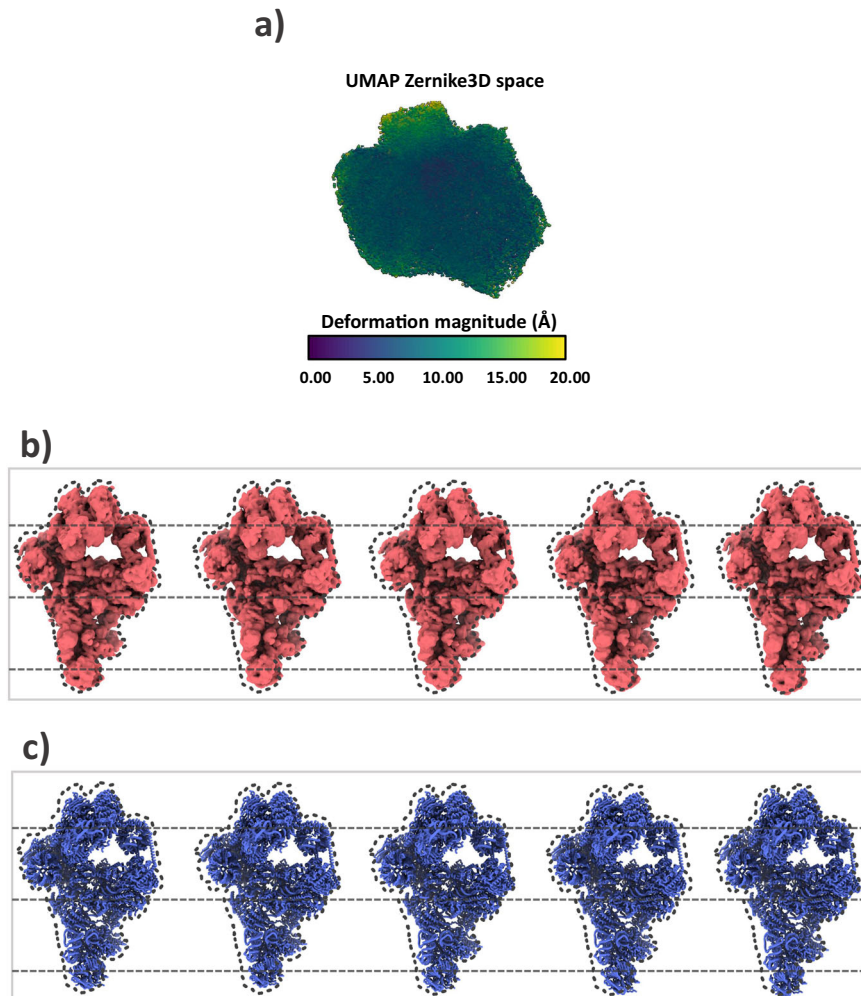


Fig. 4 | Analysis of the EMPIAR-10180 Zernike3D conformational landscape. **a** UMAP representation of the Zernike3D coefficient space obtained from the estimation of the per-particle conformational changes associated with EMPIAR-10180. The color map represents the deformation field magnitude associated with each particle involved in the analysis. **b** Example of five conformations extracted from the landscape shown in **a** after clustering by KMeans. The different conformations correspond to the representatives of each KMean cluster. The dotted gray lines are provided to enhance the visualization of the conformational changes.

The dotted map corresponds to the reference conformation, provided also to aid in the visualization of the conformational changes. **c** Example of the conformational changes shown in **b** at atomic levels. The conformations were obtained after applying the deformation fields associated with the KMean representatives in **b** to the atomic structure deposited with the EMPIAR-10180 dataset. The dotted lines and the contour of the reference map are also shown to aid in the visualization of the motions.

experiment, we set the maximum basis degrees to $N = 3$ and $L = 2$, and particles were binned to a box size of 128 pixels. The resulting Zernike3D coefficient space is represented in Fig. 4a. The Zernike3D space obtained is similar to the continuous heterogeneity region described by other software like CryoDrngn (Fig. 6 of their manuscript). However, the representation of the conformational changes followed in the Zernike3D approach provides a more versatile manner to assess structural variability.

An example of the versatility of the Zernike3D results is shown in Fig. 4b, c, and Supplementary Movie 2. The maps and structures shown in both Figures were obtained by clustering the Zernike3D space with KMeans into 5 different regions. Then, the representative Zernike3D coefficients of each cluster were extracted to represent the different conformational changes.

Similarly to other algorithms, the conformational changes can be represented at the level of Cryo-EM maps, although the Zernike3D representation will keep the same resolution as the reference map used for the analysis. In addition, the Zernike3D deformation fields can also be applied directly to an atomic structure traced or fitted to the

reference map. In this way, it is also possible to compare the different conformations at an atomic level.

An example of the comparison between two of the previous structures is provided in Fig. 5. Thanks to the Zernike3D approach, it is possible to analyze both, the local and global motion of the atoms present in the structure, which provides a more accurate and informative representation of the conformational changes suffered by the spliceosome.

SARS-CoV-2 spike one RBD up the conformational landscape

We next applied the Zernike3D algorithm to a set of particles acquired from the SARS-CoV-2 spike. In our previous work¹⁷, we followed a discrete classification approach followed by a PCA (Principal Component Analysis)¹⁸ to study the presence of flexibility in these images, revealing two different open conformations of one of the Receptor Binding Domains (RBDs). The conformations represent small motions around an open RBD state.

The analysis of this dataset is useful to assess the ability of the Zernike3D algorithm to detect small motions from the noisy Cryo-EM

images. Thus, we estimated the deformation fields for each particle starting from one of the conformations reported in ref. ¹⁷. The parameters set for this execution were the same as those used in the previous experiments ($N = 3$ and $L = 2$, yielding a total of 39 components per coefficient set). The particles were also downsampled to a box size of 125 voxels). The UMAP representation of this space is shown in Fig. 6a. The resulting space displays several interesting regions to be analyzed, and it is much richer than the space explored by discrete classification.

In addition, we can integrate the results of the previous discrete classification analysis, resulting in two main classes, with our continuous flexibility approach, by projecting all this information into the same Zernike3D space (in practice, in the reduced representation of the conformational landscape), effectively combining maps and images. The combined space is shown in Fig. 6b. The new representation simplifies the analysis of the embedding, aiding in the identification of the possible conformational changes of the spike by comparing the continuous states to the information of the discrete classification. Clearly, there is much more flexibility than the one originally accounted for by the discrete classification.

An exploration of the conformational space shown in Fig. 6a is provided in Supplementary Movie 3. The different states presented in the video were obtained by applying a set of 20 Zernike3D coefficients

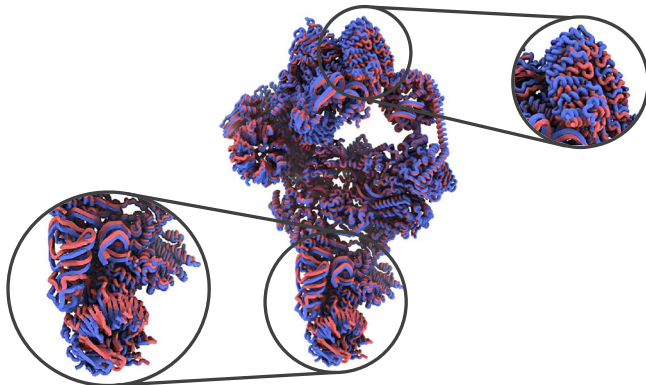


Fig. 5 | Example of recovered Zernike3D spliceosome states at atomic level. Example of two conformations obtained from the KMeans clustering of the Zernike3D space in Fig. 4. The conformations were obtained after applying the Zernike3D deformation fields to the atomic structure associated with the reference map used during the Zernike3D analysis. The versatility of the Zernike3D results allows following both, the local and global conformational changes due to the atom's motion, as shown from the different zooms in the recovered structures.

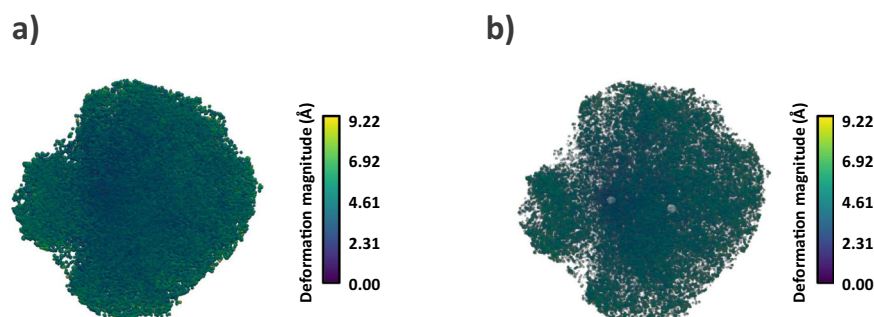


Fig. 6 | SARS-CoV-2 Zernike3D conformational landscapes. **a** UMAP representation of the Zernike3D coefficient space for the SARS-CoV-2 coronavirus spike open state obtained from the particles analyzed previously in ref. ¹⁷. Each point in the space represents a different particle conformation. **b** Combined analysis of particles and volumes (in white dots) corresponding to the two RBD states

to the reference map and its traced structure, followed by morphing in ChimeraX. The representatives were obtained by clustering the space with KMeans.

The embedding shows an interesting region (composed of a low number of particles) along the direction defined by the white dots representing each classified map. The analysis of this region reveals a conformational change moving in the opposite direction to the one defined by the two discrete classes, which was not previously identified. Supplementary Movie 4 shows the whole motion of the 1Up RBD defined by the main transition identified in the coefficient space. This result shows the importance of analyzing the heterogeneity on a per-particle basis, as discrete classification might not have the ability to resolve low-represented states.

The next step we followed in the analysis of the dataset was to use the estimated deformation fields and the particles to reconstruct a higher-resolution map by “undoing” the conformational changes of each image. The motion-corrected map reconstructed with ZART is provided in Fig. 7a. As expected, the information available in the deformation fields leads to a better resolvability of the moving areas of the spike (the RBDs and N-terminal domains (NTDs) for this specific case), increasing the local resolution of these regions. Fig 7b shows a comparison of the local resolution histograms associated with the maps shown in Fig. 7a. The correction of the per-particle conformational changes leads to a significant increment of the local resolution in the case of ZART, thanks to the reduction of the motion induced blurring present in the CryoSparg reconstruction.

Discussion

Continuous heterogeneity is widely considered to be a significant breakthrough in the Cryo-EM field, progressively becoming more popular, as shown by the several new software developments to analyze this information from the acquired particle images.

In this regard, we have introduced an extension of the Zernike3D algorithm during this work, which has proven to be a versatile tool to study the continuous motion of macromolecules at the level of maps, structures, and particle images. The extension focuses on the extraction of per-particle conformations, leading to a much more detailed description of the conformational landscape of a molecule compared to classical 3D classification approaches. Furthermore, the versatility of the Zernike3D basis unites maps, particles, and structures into a common framework, opening new possibilities to perform combined heterogeneity analysis with all data available.

Moreover, we have proven that the resulting coefficient space can be applied simultaneously to Cryo-EM maps and atomic coordinates to approximate a new conformational state. The approximation of the

described in ref. ¹⁷. Thanks to the combined analysis, we can detect a clear group of particles corresponding to an unidentified conformation in the dataset. The colormap in both images represents the modulus of the deformation field associated with the conformation estimated for every particle.

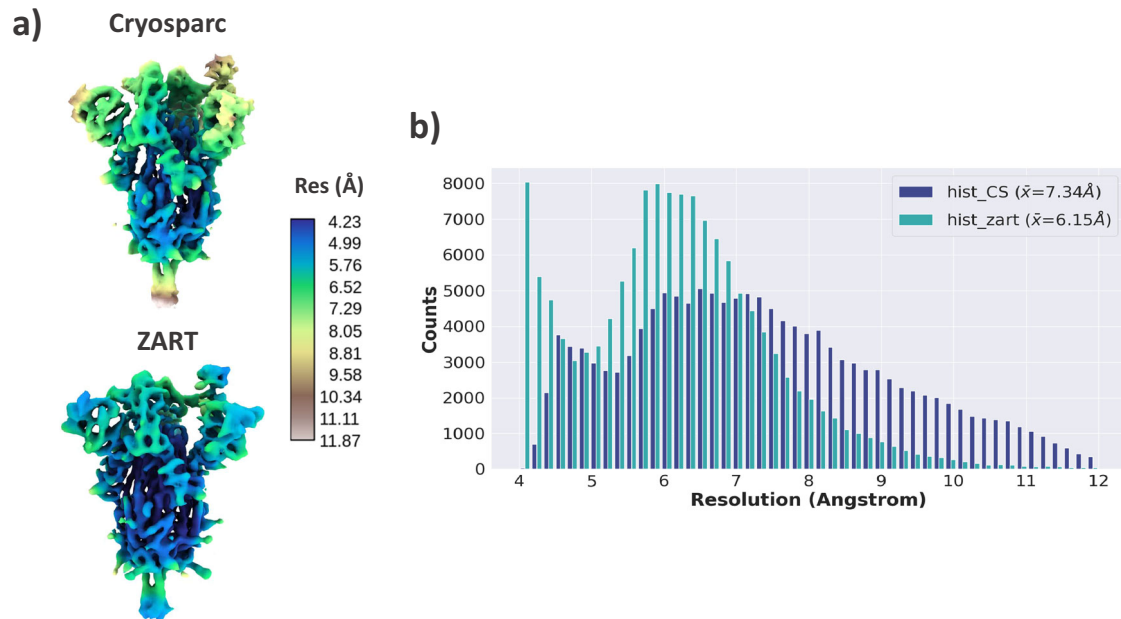


Fig. 7 | Analysis of SARS-CoV-2 ZART reconstruction. **a** Comparison of SARS-CoV-2 coronavirus spike refined with CryoSparc and the motion-corrected map recovered with our ZART algorithm. The colormap represents the local resolution computed from BlocRes²³. ZART reconstruction presents a clear improvement in map quality in the RBD, NTDs, and other regions of the spike thanks to the correction of the

motions. **b** Resolution histogram comparison for CryoSparc and ZART reconstructions. The histogram shows a clear displacement of the local resolutions towards the high-resolution regime in the case of ZART. The resolution value provided in the legend of the histograms corresponds to the mean of the local resolution measurements.

conformational changes at the atomic level supposes another step in the connection of the Cryo-EM landscapes with molecular dynamics. This connection will allow getting real energetic landscapes directly based on experimental data in the future.

In addition, we have developed the ZART reconstruction algorithm, which considers deformation fields during the reconstruction to “undo” conformational changes. In this way, it is possible to model the blurring artifacts induced by molecular motions and increase the local resolution of the reconstructed volumes.

Methods

This section is organized starting with general presentations of the Zernike3D basis and its use for the case of particles exhibiting continuous flexibility (first two subsections), and then dedicating several subsections to useful properties of our proposed method, see also Supplementary Methods.

We also provide some metrics regarding the performance of the Zernike3D algorithm in Table 1.

Zernike3D basis definition

We use the Zernike3D to estimate the deformation field associated with a given conformational transition, as we previously described in our work¹⁰.

The Zernike3D basis is an infinite-dimensional function space defined over the unit ball. Thus, it is convenient to express it as the combination of a radial and an angular component. For this basis, we have chosen the normalized and generalized definition of the Zernike

polynomials as the radial component:

$$\bar{R}_{l,n}^p(x) = \sqrt{2} \sqrt{2n+l + \frac{p}{2} + 1} R_{l,n}^p(x) \quad (1)$$

p being a parameter associated with the inner product and dimensionality of the polynomials. For example, in a 3D scenario, the appropriate value of p should be 1.

The previously mentioned angular component is defined in terms of the real spherical harmonics:

$$y_l^m(\theta, \phi) = (-1)^m \sqrt{\frac{2l+1(l-|m|)!}{4\pi(l+|m|)!}} P_l^{|m|}(\cos\theta) \begin{cases} 1 & \text{if } m=0 \\ \sqrt{2}\cos(m\phi) & \text{if } m>0 \\ \sqrt{2}\sin(|m|\phi) & \text{if } m<0 \end{cases} \quad (2)$$

By combining the previous two components, we obtain the final definition of the Zernike3D basis:

$$Z_{l,n,m}(r) = \bar{R}_{l,n}^1(r) y_l^m(\theta, \phi) \quad (3)$$

Estimating deformation fields from particles

As we explained in the previous section, the Zernike3D basis was initially formulated to be applied to 3D spaces. Therefore, it is quite direct to estimate conformational transitions from maps or atomic structures, as they live in a three-dimensional space. However, this is not the case for Cryo-EM particles, as we are intrinsically losing information along the projection direction during the acquisition process in the reduction from the three-dimensional space where the Coulomb potential of the specimen is defined to the two-dimensional space of the projection images being acquired in the microscope. In other words, conformational changes along the projection direction cannot be extracted from an individual image, since an infinite number of them would be compatible with the image information.

Table 1 | Execution times for the Zernike3D algorithm

Performance metrics for the Zernike3D algorithm				
Image size	N	L	Time per-particle (s)	Time for 10^6 particles (min-150 threads)
128	3	2	0.1076	39.93
300	3	2	0.182	300.00

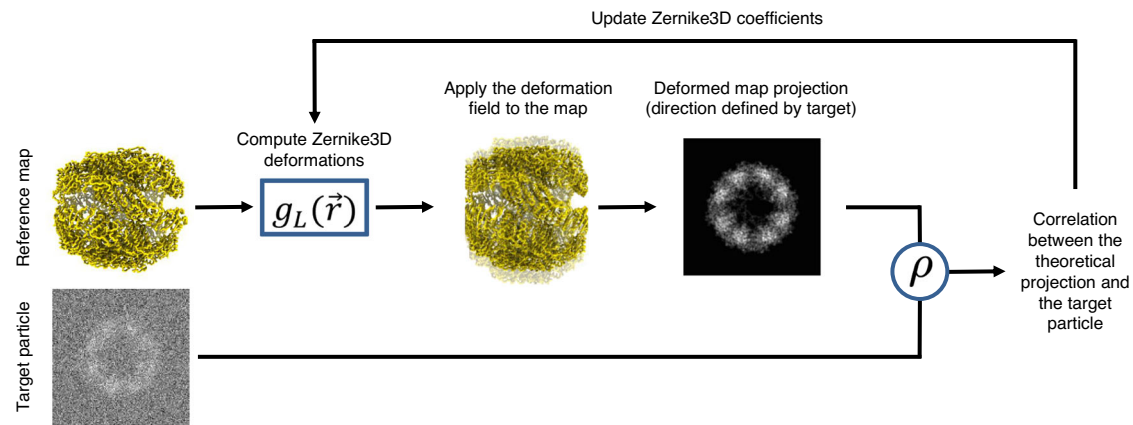


Fig. 8 | Zernike3D workflow at particle level. Estimation of the Zernike3D deformation fields from a particle image. The process requires the deformation of a reference map to be consistent with the dimension of the deformation field, which is defined by the set of Zernike3D coefficients. Each coefficient component is

estimated by Powell's conjugate direction method, so Pearson's correlation coefficient ρ between the experimental and theoretical projection obtained from the deformed map is maximized.

The algorithm we present in this work starts by computing a reference map/model (in practice, and continuing the presentation for the case of maps, it is common to either use an average map or one of the discrete classes). This map will be the origin (reference) to obtain the deformation fields from the parameters of the Zernike3D basis. The approach is summarized in Fig. 8, and it is a common procedure in optimization. In brief, it is an iterative procedure in which deformation fields are applied to the reference map and the resulting projection images are compared with the experimental ones until convergence.

Following the aforementioned method, finding the deformation field to describe the state represented by a given particle can be expressed as:

$$\max_{g_L} \rho(I, CP_{\theta}(V(r + g_L(r)))) \quad (4)$$

ρ being the Pearson's correlation coefficient, I an experimental Cryo-EM particle, C the CTF estimated for that particle, P_{θ} the projection operator along the 3D direction and in-plane shift defined by the parameters θ , V the reference volume needed to apply the Zernike3D deformation field, and g_L the displacement suffered by each voxel due to the deformation field. The vector g_L depends on each Zernike3D component, and it is expressed as:

$$g_L(r) = \sum_{l=0}^L \sum_{n=0}^N \sum_{m=-l}^l \begin{pmatrix} \alpha_{l,n,m}^x \\ \alpha_{l,n,m}^y \\ \alpha_{l,n,m}^z \end{pmatrix} Z_{l,n,m}(r) \quad (5)$$

where the $\alpha_{l,n,m}$'s are the Zernike3D coefficients. The previous coefficients determine the contribution of each component of the basis to the deformation field.

The parameters N and L determine the maximum degrees of the Zernike polynomials and spherical harmonics. Therefore, they will determine the accuracy of the deformation fields: higher values will result in sharper and more accurate deformation fields, at the expense of increased execution times. By default, the two previous parameters are set to $N=3$ and $L=2$, which should be enough to avoid overfitting and get meaningful and accurate deformation fields in most cases. Nevertheless, the parameters can be manually set by the user in case higher accuracy is desired.

The maximization of Eq. 4 is achieved through a Powell's conjugate direction method starting from an initial guess of $\alpha_{l,n,m} = 0$ for all indices l,n,m and directions x,y,z (that is, no deformation). Thanks to the optimization method and the procedure described in Fig. 8, it is

possible to find the different component contributions $\alpha_{l,n,m}$ such that the deformation field to be applied to the reference map V leads to a conformational state compatible with the particle I .

In order to avoid possible overfitting during the Powell search of the Zernike3D coefficients, an extra regularization term is added to Eq. 4:

$$\max_{g_L} \rho(I, CP_{\theta}(V(r + g_L(r)))) + \lambda_1 \int |g_L(r)|^2 dr \quad (6)$$

The additional regularization term accounts for the total deformation the reference map has suffered after applying the estimated deformation field. Depending on the value of λ_1 , the optimization search will be allowed to explore minima leading to a larger or smaller deformation, so it is recommended to set it at a low value to avoid overfitting without compromising the minima search. The user can choose the value of λ_1 to be applied to a specific dataset. We recommend selecting a value belonging to the range $[0.01, 0.001]$ to avoid undesired results during the optimization process.

It is worth mentioning that the Zernike3D algorithm does not require a minimum number of particles to be executed, as the deformation fields are estimated for every particle. Therefore, it is possible to process datasets coming from a consensus or other cleaning methods¹⁹, whose parameters are more accurately estimated but have fewer particles overall.

Deformation field consistency along the projection direction

In this work, we estimate 3D deformation fields from 2D images. However, the information limitations introduced above make this procedure conceptually challenging. Indeed, if we compare the information stored in a projection and a map, it would be possible to check that we have identical information as long as we restrain the comparison to the projection plane where the image exists. In spite of that, the image has an intrinsic loss of information in the projection direction, as we are collapsing the map information stored along this direction.

Following the previous reasoning, the deformation field is well-defined across the projection plane, but it is ill-defined along the projection direction. This implies that there are infinite ways of deforming a map along the projection direction defined by particles so that the projection of the deformed volume is still consistent with the particle. Therefore, the Powell optimization proposed previously might find different solutions for each particle along the projection direction. Moreover, this inconsistency might lead to the global optimization process being more prone to get trapped in local minima,

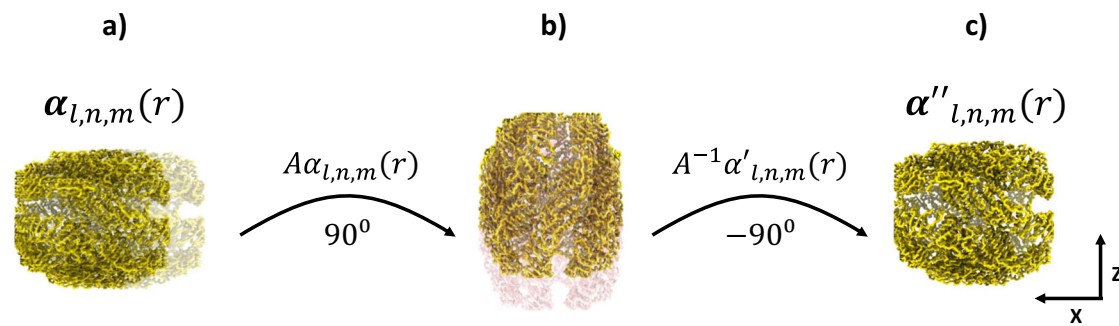


Fig. 9 | Projection direction correction workflow. Procedure to cancel excessive displacements associated with the projection direction in a deformation field defined by the coefficients $\alpha_{l,n,m}(r)$. **a** Representation of a deformation field with over-deformation along the X direction (projection direction). **b** The over-deformed volume and the associated deformation coefficients are rotated so that the projection direction matches the Z axis. At this point, it is possible to

completely cancel the Z component on the rotated coefficients, as they only contribute to the over-deformation of the deformed map. This leads to the set of coefficients $\alpha'_{l,n,m}(r)$. **c** The modified coefficients are rotated back to the original position defined by the map's grid. In this way, a new set of coefficients $\alpha''_{l,n,m}(r)$ are obtained, defining a deformation field consistent along the projection direction (X axis).

resulting in wrong estimations of the conformational landscape. An example of the undesired effect generated by not considering the missing information along the projection direction is provided in Supplementary Fig. 3.

Ideally, the best solution to the previous problem would be to drive the optima search so that all particles are not deformed along their respective projection direction. For this reason, a sensible choice would be to completely restrict the deformation along the projection direction.

One possible way to achieve this is to include different regularization terms restricting excessive deformations. Nevertheless, it would be challenging to find the weights needed for each regularization term, especially along the projection direction, a situation that introduces a new parameter quite difficult to estimate in the process.

The Zernike3D method overcomes that obstacle by taking advantage of the properties of the basis to altogether remove any deformation along the projection direction defined by a particle, either during the optimization or after it. As we showed in ref. ¹⁰, the Zernike3D basis is closed under rotations. Thus, it is possible to rotate the Zernike3D coefficients towards a different reference frame as follows:

$$Ag_L(A^{-1}r) = \sum_{n=0}^N \sum_{l=0}^L \sum_{m=-l}^l A \begin{pmatrix} \alpha_{l,n,m}^x \\ \alpha_{l,n,m}^y \\ \alpha_{l,n,m}^z \end{pmatrix} \tilde{Z}_{l,n,m}(A^{-1}r) \quad (7)$$

A being a rotation matrix. Therefore, it is possible to rotate the Zernike3D coefficients according to the angular information of the particle. For example, we can rotate the coefficients so that the Z direction of the new frame is effectively aligned with the projection direction of a particle. Then, we can cancel the rotated coefficients associated with the Z-axis in this new frame, as those only contribute to the deformation field along the projection direction.

However, it is essential to note that the previous property only holds in a continuous space. Hence, the basis is not closed under rotations due to the discretization and sampling of the space into voxels. Thus, the rotated coefficients cannot be applied to the volume, as the reference frames are entirely different. Instead, we can unrotate the modified Zernike3D coefficients, so their reference frame matches again with the reference map. Thus, it is possible to fully remove the deformation along the projection direction by combing all the previous steps, making all the deformation fields consistent, and avoiding solutions with unrealistic deformations. The whole procedure is summarized in Fig. 9.

Note that we also use a global regularization term in our optimization approach, as previously indicated, but it is global and does not differentiate among projection directions. In our experience, the

previously stated “consistency principle” introduced by mathematically clear handling of the lack of information along individual projection directions is quite important factor in our quest for estimating deformation fields, and it represents a clear advantage of the Zernike3D approach.

Zernike3D-based ART reconstruction algorithm

In general, 3D reconstruction algorithms start from the principle that we have a set of projections coming from a homogeneous set of particles. However, this assumption no longer holds for those macromolecules exhibiting large degrees of freedom. Therefore, it is not a surprise that molecular motions are a well-known source of blurring artifacts arising when reconstructing a Cryo-EM map from a set of Cryo-EM images. As a consequence, correcting the motions present in a particle image will be expected to boost the resolution and resolvability of blurry areas in Cryo-EM maps.

The per-particle estimation of the deformation fields by the Zernike3D basis can be effectively applied to correct molecular motions, aiding the reconstruction process with flexible information. To that end, we developed an ART-based (Algebraic Reconstruction Technique) reconstruction algorithm that uses the Zernike3D deformation fields to improve the final quality of motion-related blurry areas.

A detailed description of ART and its application in Cryo-EM can be found at ref. ²⁰. Here it suffices to say that ART finds the map whose projections are compatible with the experimental data through an iterative process of the form:

$$V^{(k+1)}(r) = V^{(k)}(r) + \lambda P_H^*(P_H V(r) - I_k(s)) \quad (8)$$

V being the reconstructed volume, λ the ART relaxation factor, P_H the projection operator, and P_H^* its adjoint operator, I_k the experimental image used at the k -th iteration, r a 3D coordinate, and s a 2D coordinate. The previous equation refers to the update to be applied to the reconstruction associated with a single image, although the algorithm will iterate over the whole particle dataset applying the previous correction to achieve the final reconstruction.

One advantage of ART over other reconstruction methods is that it can be easily modified to include new information to be taken into account during the iterative reconstruction process. Thus, it is possible to modify the previous equation by adding the deformation field previously estimated:

$$V(r)^{k+1} = V(r)^k + \lambda P_H^*(P_H(V(r) + g_L(r)) - I_k(s)) \quad (9)$$

$g_L(r)$ being the displacement at a given 3D position computed through Eq. 5.

By introducing the displacements g_L into the ART algorithm, we are improving the correction value that will be applied to V at each iteration, as the difference between the theoretical and experimental images is taken based on the conformational change present in the particle. Thus, the reconstruction process can generate more meaningful solutions for areas subjected to significant motions.

Moving Zernike3D coefficients through different resolutions

One of the main issues arising when working with Cryo-EM particles is the low signal-to-noise ratios that they exhibit. Although the average of a very high number of images overcomes that problem, the procedure also mixes several conformational changes at the same time. Therefore, the estimation of continuous flexibility is usually done directly on particle images, even if conditions are far from ideal.

To estimate motions more efficiently, it is common to filter the particles at a given resolution to increase the signal-to-noise ratio. Similarly, it is possible to downsample the images after the filtering process to improve the performance of the estimations.

However, the Zernike3D coefficients $\alpha_{l,m,n}$ have a strong dependency on the size of the volume under study. This implies that the coefficients computed from a downsampled map cannot be directly applied to the original volume and vice versa.

Luckily, it is possible to move a set of Zernike3D coefficients to a different resolution similar to the procedure described in previous sections. By downsampling a map, we scale its space by a given factor k . Therefore, the relation between two vectors with the same direction in the previous two spaces is:

$$r_o = kr_d \quad (10)$$

r_o and r_d being the vectors associated with the original and downsampled spaces, respectively.

We can express the earlier two vectors based on the components of the Zernike3D basis as follows:

$$\begin{aligned} \sum_{n=0}^N \sum_{l=0}^L \sum_{m=-l}^l \alpha_{l,n,m}^o \tilde{Z}_{l,n,m}(r_o) &= k \sum_{n=0}^N \sum_{l=0}^L \sum_{m=-l}^l \alpha_{l,n,m}^d \tilde{Z}_{l,n,m}(r_d) \\ &= \sum_{n=0}^N \sum_{l=0}^L \sum_{m=-l}^l k \alpha_{l,n,m}^d \tilde{Z}_{l,n,m}(k^{-1}r_o) \end{aligned} \quad (11)$$

Thanks to Eq. 11, it is possible to show that the scaling relation existing between the vectors x_o and x_d is shared by the corresponding Zernike3D coefficients:

$$\alpha_{l,n,m}^o = k \alpha_{l,n,m}^d \quad (12)$$

This leads to a very convenient and straightforward conversion to use coefficients estimated on low-resolution images in the original high-resolution maps.

Merging embeddings of different nature

Our previous work¹⁰ showed that the Zernike3D basis could effectively study the continuous heterogeneity of Cryo-EM maps and atomic structures converted to electron densities.

Similarly, we have proven in the previous sections the applicability of this same tool to a set of Cryo-EM particles. In all cases, the estimation of the deformation fields represented by the Zernike3D coefficients is comparable, meaning that we are translating the information of the three main Cryo-EM entities (maps, atomic structures, and particles) to a common framework or space defined by the coefficients $\alpha_{l,m,n}$.

Translating maps, structural models, and particles to a common framework opens interesting possibilities and advantages, such as studying and comparing discrete and continuous heterogeneity or addressing how well simulated and experimental data correlate.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The datasets analyzed with the Zernike3D algorithm and ZART are publicly available in EMPIAR under the entries: 10028 [<https://doi.org/10.6019/EMPIAR-10028>], 10514 [<https://doi.org/10.6019/EMPIAR-10514>], 10516 [<https://doi.org/10.6019/EMPIAR-10516>], and 10180 [<https://doi.org/10.6019/EMPIAR-10180>]. The phantom dataset processed in the Supplementary Material is available in GitHub in the repository Zernike3D_Phantom_Data [<https://zenodo.org/badge/latestdoi/541505536>]²¹.

Code availability

The Zernike3D algorithm has been implemented in Xmipp²² and it is available through Scipion¹² under the plugins “scipion-em-xmipp” and “scipion-em-flexutils”. The protocols corresponding to the algorithms described in this manuscript are “flexutils - angular align - Zernike3D” and “flexutils - reconstruct ZART”.

References

- Carroni, M. & Saibil, H. R. Cryo electron microscopy to determine the structure of macromolecular complexes. *Methods* **95**, 78–85 (2016).
- Serna, M. Hands on methods for high resolution cryo-electron microscopy structures of heterogeneous macromolecular complexes. *Front. Mol. Biosci.* **6**, 33 (2019).
- Gomez-Blanco, J., Kaur, S., Strauss, M. & Vargas, J. Hierarchical autotclassification of cryo-EM samples and macromolecular energy landscape determination. *Comput. Methods Prog. Biomed.* **216**, 106673 (2022).
- Jin, Q. et al. Iterative elastic 3D-to-2D alignment method using normal modes for studying structural dynamics of large macromolecular complexes. *Structure* **22**, 496–506 (2014).
- Zhong, E. D., Bepler, T., Berger, B. & Davis, J. H. CryoDRGN: reconstruction of heterogeneous cryo-EM structures using neural networks. *Nat. Methods* **18**, 176–185 (2021).
- Ludtke, S. J. & Muyuan, C. Deep learning-based mixed-dimensional Gaussian mixture model for characterizing variability in cryo-EM. *Nat. Methods* **18**, 930–936 (2021).
- Frank, J. & Abbas, O. Continuous changes in structure mapped by manifold embedding of single-particle data in cryo-EM. *Methods* **100**, 61–67 (2016).
- A. Punjani, A. & Fleet, D. J. 3D flexible refinement: structure and motion of flexible proteins from Cryo-EM. *bioRxiv*, <https://www.biorxiv.org/content/10.1101/2021.04.22.440893v1> (2021).
- Lederman, R. R., Anden, J. & Singer, A. Hyper-molecules: on the representation and recovery of dynamical structures for applications in flexible macro-molecules in cryo-EM. *arXiv*, <https://arxiv.org/abs/1907.01589> (2020).
- Herreros, D. et al. Approximating deformation fields for the analysis of continuous heterogeneity of biological macromolecules by 3D Zernike polynomials. *IUCrJ* **8**, 992–1005 (2021).
- Wong, W. et al. Cryo-EM structure of the 80S ribosome bound to the anti-protozoan drug emetine. *eLife* **3**, e03080 (2014).
- de la Rosa-Trevín, J. M. et al. Scipion: a software framework toward integration, reproducibility and validation in 3D electron microscopy. *J. Struct. Biol.* **195**, 93–99 (2016).
- McInnes, L., Healy, J., Saul, N. & Großberger, L. UMAP: uniform manifold approximation and projection. *J. Open Source Softw.* **3**, 861 (2018).
- Pettersen, E. F. et al. UCSF ChimeraX: structure visualization for researchers, educators, and developers. *Protein Sci.* **30**, 70–82 (2021).

15. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017).
16. Plaschka, C., Lin, P. C. & Nagai, K. Structure of a pre-catalytic spliceosome. *Nature* **564**, 617–621 (2017).
17. Melero, R. et al. Continuous flexibility analysis of SARS-CoV-2 spike prefusion structures. *IUCr* **7**, 1059–1069 (2020).
18. Jolliffe, I. & Cadima, J. Principal component analysis: a review and recent developments. *Philos. Trans. A Math. Phys. Eng. Sci.* **374**, 20150202 (2016).
19. Sorzano, C. O. S. et al. On bias, variance, overfitting, gold standard and consensus in single-particle analysis by cryo-electron microscopy. *Acta Crystallogr. Sect. D.* **78**, 410–423 (2022).
20. Sorzano, C. O. S. et al. A survey of the use of iterative reconstruction algorithms in electron microscopy. *BioMed. Res. Int.* **2017**, 1–17 (2017).
21. Herreros, D. Estimating conformational landscapes from Cryo-EM particles by 3D Zernike polynomials <https://doi.org/10.5281/zenodo.7334391>, (2022).
22. de la Rosa-Trevín, J. M. et al. Xmipp 3.0: an improved software suite for image processing in electron microscopy. *J. Struct. Biol.* **184**, 321–328 (2013).
23. Heymann, J. B. Guidelines for using Bsoft for high resolution reconstruction and validation of biomolecular structures from electron micrographs. *Protein Sci.* **27**, 159–171 (2018).

Acknowledgements

Funding is acknowledged from the Ministry of Science and Innovation through grants: Grant PID2019-104757RB-I00 funded by MCIN/AEI/10.13039/501100011033/ and “ERDF A way of making Europe”, by the “European Union; the ‘Comunidad Autonoma de Madrid’ through grant S2017/BMD-3817; and the European Union (EU) and Horizon 2020 through grants EnLaCES (H2020-MSCA-IF-2020, Proposal: 101024130 to J.M.K.), HighResCells (ERC-2018-SyG, Proposal: 810057) and iNEXT-Discovery (Proposal: 871037). This work has also been supported by the NIH/NIGMS (No. 1R01GM136780-01 to RRL) and AFSOR FA9550-21-1-0317.

Author contributions

D.H. developed and tested the Zernike3D and ZART algorithms presented in the manuscript. R.R.L. helped with the understanding and demonstration of the mathematical properties of the Zernike3D basis.

J.M.K. and M.M. helped with the understating of the results and the preparation of the data. A.J.M. helped with the implementation of the Zernike3D algorithm in Scipion. D.M., D.S., and J.F. helped with the optimization of the codes. C.O.S.S. and J.M.C. jointly supervised this work.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-023-35791-y>.

Correspondence and requests for materials should be addressed to D. Herreros.

Peer review information *Nature Communications* thanks Yoel Shkolnisky and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

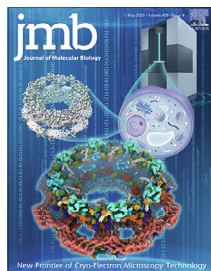
Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023

ZART: A Novel Multiresolution Reconstruction Algorithm with Motion-blur Correction for Single Particle Analysis

D. Herreros, J. Kiska, E. Ramírez-Aportela, J. Filipovic, et al. ZART: A Novel Multiresolution Reconstruction Algorithm with Motion-blur Correction for Single Particle Analysis. *Journal of Molecular Biology* 435, 168088 (2023). <https://doi.org/10.1016/j.jmb.2023.168088>



ZART: A Novel Multiresolution Reconstruction Algorithm with Motion-blur Correction for Single Particle Analysis

D. Herreros^{1*}, J. Kiska², E. Ramírez-Aportela¹, J. Filipovic², J. M. Carazo^{1*} and C. O. S. Sorzano^{1*}

1 - Centro Nacional de Biotecnología-CSIC, C/ Darwin, 3, 28049, Cantoblanco, Madrid, Spain

2 - Institute of Computer Science, Masaryk University, Botanická 68a, 60200 Brno, Czech Republic

Correspondence to D. Herreros, J.M. Carazo and C.O.S. Sorzano: dherreros@cnb.csic.es (D. Herreros), carazo@cnb.csic.es (J.M. Carazo), coss@cnb.csic.es (C.O.S. Sorzano) @HerosCGM, @InstructI2PC (D. Herreros)

<https://doi.org/10.1016/j.jmb.2023.168088>

Edited by Fei Sun

Abstract

One of the main purposes of CryoEM Single Particle Analysis is to reconstruct the three-dimensional structure of a macromolecule thanks to the acquisition of many particle images representing different poses of the sample. By estimating the orientation of each projected particle, it is possible to recover the underlying 3D volume by multiple 3D reconstruction methods, usually working either in Fourier or in real space. However, the reconstruction from the projected images works under the assumption that all particles in the dataset correspond to the same conformation of the macromolecule. Although this requisite holds for some macromolecules, it is not true for flexible specimens, leading to motion-induced artefacts in the reconstructed CryoEM maps. In this work, we introduce a new Algebraic Reconstruction Technique called ZART, which is able to include continuous flexibility information during the reconstruction process to improve local resolution and reduce motion blurring. The conformational changes are modelled through Zernike3D polynomials. Our implementation allows for a multiresolution description of the macromolecule adapting itself to the local resolution of the reconstructed map. In addition, ZART has also proven to be a useful algorithm in cases where flexibility is not so dominant, as it improves the overall aspect of the reconstructed maps by improving their local and global resolution.

© 2023 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Introduction

CryoEM SPA¹ has proven to be one of the most successful techniques to recover the structure of a macromolecule at near-atomic resolution. In addition, its ability to capture different macromolecular states is driving the CryoEM field towards a new way of analyzing and understanding macromolecular flexibility.

However, the reconstruction of structures at high resolution is compromised when the sample is heterogeneous, as the reconstruction algorithms

assume that all particles involved in the reconstruction process are compatible with a single underlying 3D map. To overcome the current challenges arising with heterogeneous datasets, it is possible to take advantage of 3D classifications² to isolate the particles belonging to a given conformation to improve the resolution of the reconstruction. Nevertheless, the classification process will decrease the number of particles that are available to reconstruct a given conformation, compromising again the maximum resolution achievable. In addition, new methods specifically

developed to estimate richer conformational landscapes are also appearing in the field.^{3–8} However, these algorithms focus on the estimation of new conformational states, but they do not take advantage of the estimated heterogeneity information to further improve the resolution in those areas exhibiting larger degrees of flexibility.

In this work, we propose a novel reconstruction algorithm called ZART (Zernike3D-based Algebraic Reconstruction Technique), an ART-based^{1,9} multiresolution reconstruction method able to take advantage of the estimated per-particle continuous heterogeneity computed by the Zernike3D algorithm to correct particle inhomogeneities.¹⁰ Thanks to the previous correction, it is possible to improve the local resolution of the reconstructions. Additionally, ZART has been designed to improve the features of the final maps even if motion corrections are not considered, helping with the tracing of the map to achieve better structures. Finally, a multiresolution approach has been adopted to adapt the local complexity of the reconstructed map to the estimated local resolution.

Compared to the first implementation of ZART presented in,¹⁰ this work proposes a major change in the algorithm, allowing to reconstruct heterogeneity corrected maps even if the Zernike3D information was estimated in particles with a larger sampling rate. Also, in our previous implementation, the reconstruction sampling rate was restricted to the particles' sampling rate used during the landscape estimation, thus compromising the maximum resolution possible. Now, we have dropped this constraint as we are able to adapt the Zernike3D coefficients to any pixel size of interest.

In addition, the multiresolution reconstruction mode of ZART is also introduced, which provides a new approach to further improve the quality of the reconstructions compared to the previous implementation.

Results

SARS-CoV-2 spike phantom reconstruction

The main objective of the following results is to analyze the characteristics of ZART through the reconstruction of a synthetic dataset. To that end, we simulated a SARS-CoV-2 Coulomb potential map using the EASFs^{2,11} from the PDB 6VSB.¹² The map was then projected to generate a gallery of 18,309 images with a sampling rate 2.40 Å/px. The motivation behind downsampling the images is to get a dataset where the sharpening effect of ZART can be more easily assessed. It is worth mentioning that the images were not further processed to include noise or the CTF³ of the microscope. Our aim at this moment is to produce an ideal dataset that allows

characterizing ZART capabilities in optimal conditions. Later in the article, we will use experimental data in which noise and CTF are naturally present.

We compared the reconstructions by ZART and Fourier gridding¹³ as implemented in Xmipp.¹⁴ The resulting maps are shown in Figure 1(a) and Figure 1(b). As can be seen from the 3D volumes and the slices, ZART yields a reconstruction with finer details, simplifying the interpretation of the map features compared to a standard Fourier reconstruction.

Since the simulated particles constitute an ideal homogeneous dataset, no motion correction was applied during the reconstruction of the ZART map. Therefore, the main reason behind the improvement in the features of the map comes from the deconvolution with a Gaussian implied by our method. This idea is further discussed in Section 'Zernike3D Algebraic Reconstruction Technique'.

In addition to the evaluation of the quality of the reconstructions, Figure 1(c) also includes the convergence curve extracted from the reconstruction error computed during the first iteration of the ZART algorithm for every image. The plot shows that a meaningful reconstruction (in terms of convergence) can be achieved after processing 5000 images from the original phantom dataset. However, it is important to note that the convergence speed has a strong dependence on the relaxation factor chosen to reconstruct a given dataset. A more detailed discussion of the importance of the relaxation factor in the reconstruction process is available in Section 'Zernike3D Algebraic Reconstruction Technique'.

Figure 2 shows a more detailed comparison of the reconstructed Fourier and ZART maps against the structure used to generate the phantom data. Thanks to the intrinsic sharpening of ZART, the reconstructed map is able to define better the local features present in the spike atomic structure.

Arabinofuranosyltransferase reconstruction

In order to assess the performance of our ZART reconstruction algorithm on a more realistic scenario, the dataset EMPIAR-10391 was used for this task.¹⁵ This dataset was processed inside Scipion following a complete workflow to extract a complete particle set of 35 k images with a sampling rate 1.06 Å/px. The alignment and CTF information associated with the particles was estimated with Relion.¹⁶

The particles were afterwards reconstructed with Relion and ZART and in all cases sharpened with 3DEMhancer.¹⁸ In the case of ZART, we followed three different reconstruction approximations: a standard reconstruction and two multiresolution reconstructions with $M = 3$ and $M = 6$ multiresolution levels respectively. Heterogeneity correction was not applied in this cases in order to compare

1 Algebraic Reconstruction Technique.

2 Electron Atomic Scattering Factors.

3 Contrast Transfer Function.

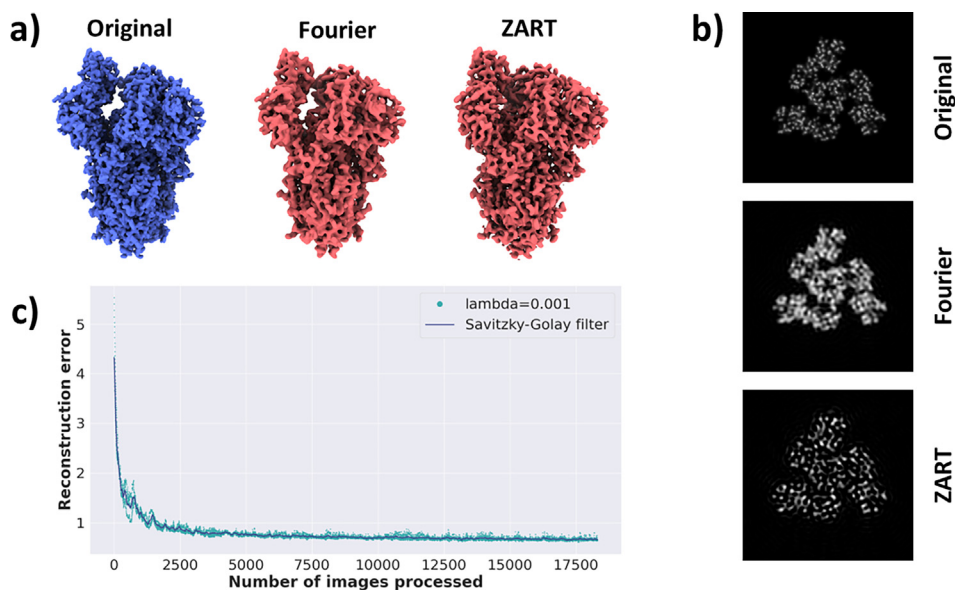


Figure 1. Results obtained for the comparison of ZART and Fourier-based reconstruction for the SARS-CoV-2 spike phantom images. (a) shows the original phantom simulated from PDB 6VSB, and the Fourier and ZART reconstruction from the phantom images. The map comparison shows that ZART provides a sharper representation of the different features of the spike. (b) shows some slices of the maps represented in a). Similarly to the conclusion drawn from the maps, ZART provides a sharper representation of the features compared to Fourier, providing a more meaningful representation. (c) shows the convergence curve of ZART during the first iteration of the algorithm. As can be seen from the curve, around 5000 images are needed to be processed to get a meaningful reconstruction in terms of convergence. Together with the original scatter data, the Savitzky–Golay filtered curve is also provided to aid in the visualization of the plot.

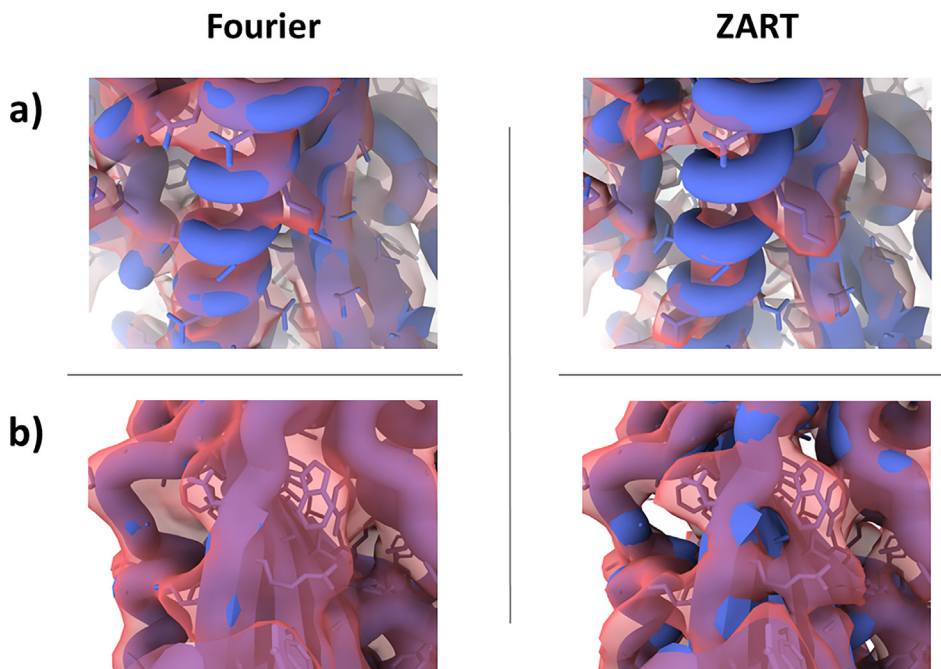


Figure 2. Comparison of Fourier and ZART reconstructions against the atomic structure 6VSB used to simulate the phantom particles. Thanks to the intrinsic sharpening characteristics of ZART, the reconstruction is able to define more accurately different areas of the map. The panels were generated with ChimeraX software,¹⁷ and histogram thresholds were set to 0.388 and 0.209 for Fourier and ZART respectively.

more effectively the standard and multiresolution reconstruction modes. The resulting reconstructed maps are shown in Figure 3(a). The colormap represents the local resolution estimation in Angstroms for each voxel, computed with MonoRes.¹⁹ As can be seen from the maps and their corresponding local resolutions, ZART provides a reconstruction with improved local resolution and better features compared to the Relion reconstruction.

In order to provide a more quantitative comparison of the maps, the local resolution histograms computed with MonoRes were also compared. Figure 3(b) shows the comparison of the resolution histogram of Relion and ZART multiresolution map with $M = 6$ resolution levels. As can be seen from the histograms, ZART pushes a larger number of voxels towards the

high-resolution regime in the range of 2.5 Å–3.2 Å. Thanks to this push, the average of the local resolutions improves from 7.0 Å to 4.0 Å.

We also computed the FSC curves of all the reconstructions analyzed during this section. The resulting curves are shown in Figure 3(c). It can be noted that ZART also improves the global resolution of the map from 3.3 Å to 3 Å. This implies an improvement of 0.3 Å due to ZART and of 0.5 Å compared to the reported resolution of the originally published maps, thanks in this latter case of the quality of the complete workflow used in this work.

Another effect that can be observed when comparing the standard and multiresolution reconstructions in ZART is that multiresolution can further improve the local resolution histograms of

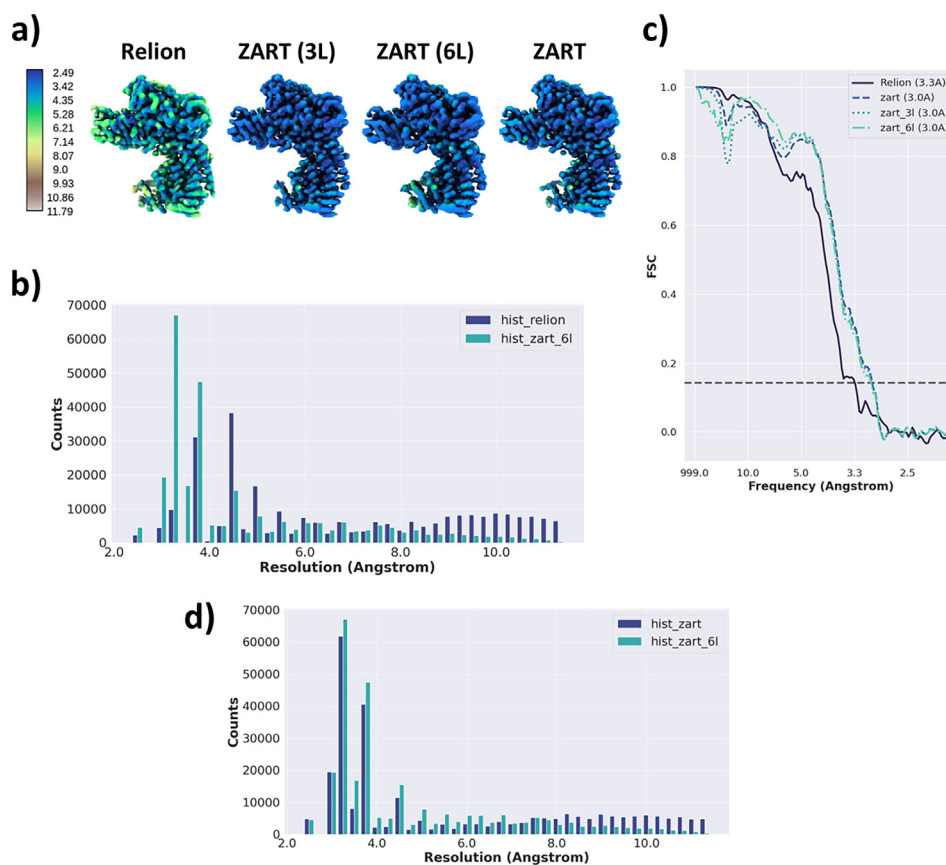


Figure 3. Reconstructions of the Arabinofuranosyltransferase from the EMPIAR-10391 dataset. (a) shows a comparison of the Relion and ZART reconstructions sharpened by DeepEMhancer¹⁸ software. In the case of ZART, we perform two multiresolution reconstructions with 3 and 6 levels respectively and a standard reconstruction without multiresolution. The map comparison suggests that ZART provides a sharper representation with improved local resolution. However, this is difficult to appreciate directly on these small representations and for that, we are making a quantitative assessment in other panels of this Figure. Panel (b) shows the local resolution histograms of the maps shown in (a) computed with MonoRes.¹⁹ The histograms show that ZART pushes a large number of voxels towards the high-resolution regime. (c) shows the FSC curves of the maps shown in (a). As shown in the previous panels, our new reconstruction method pushes the overall resolution of the map around 0.3 Å. (d) Comparison of the local resolution between the ZART reconstructions both in standard and in multiresolution mode. The comparison shows that multiresolution supposes a little improvement in the local resolution values thanks to the consideration of multiresolution grids with variable Gaussian widths.

the reconstructed volumes, although the difference is small. Figure 3(d) shows the comparison of the local resolution histograms between the final reconstruction of ZART (in standard mode) and the ZART in multiresolution mode (with 6 multiresolution levels). The two histograms show a similar type of distribution of the local resolutions, although multiresolution pushes a larger amount of values towards the high-resolution regime. In addition, multiresolution implementation has little impact on the performance of every reconstruction iteration, allowing to achieve maps in a similar time compared to the standard reconstruction mode.

Figure 4 shows a more detailed comparison of the Relion and ZART reconstructions against the original structure published with the dataset. In order to make the comparison more reliable, the structural model has not been further refined against any of the two previous maps, and no sharpening was applied to the volumes. The result illustrates how the sharpening applied by ZART leads to a reconstruction representing more

accurately the features of the arabinofuranosyltransferase.

Additionally, a comparison of ZART with CryoSPARC non-uniform refinement was performed, in order to better assess the performance against other algorithms applying a de-blur to the resulting map. The results are provided in Figure 5. Overall, both CryoSPARC and ZART maps show similar features, although for some regions ZART showed a better definition of the molecular structure. We also provide a comparison of the local resolution histograms associated with the two previous reconstructions, together with some measurements drawn from them in Supplementary Figure 1.

P. falciparum 80S ribosome reconstruction

The *P. falciparum* 80S ribosome of EMPIAR-10028²⁰ has become a quite standard dataset to evaluate the performance and accuracy of continuous heterogeneity algorithms, due to the presence of significant conformational changes that can be extracted directly at particle level.

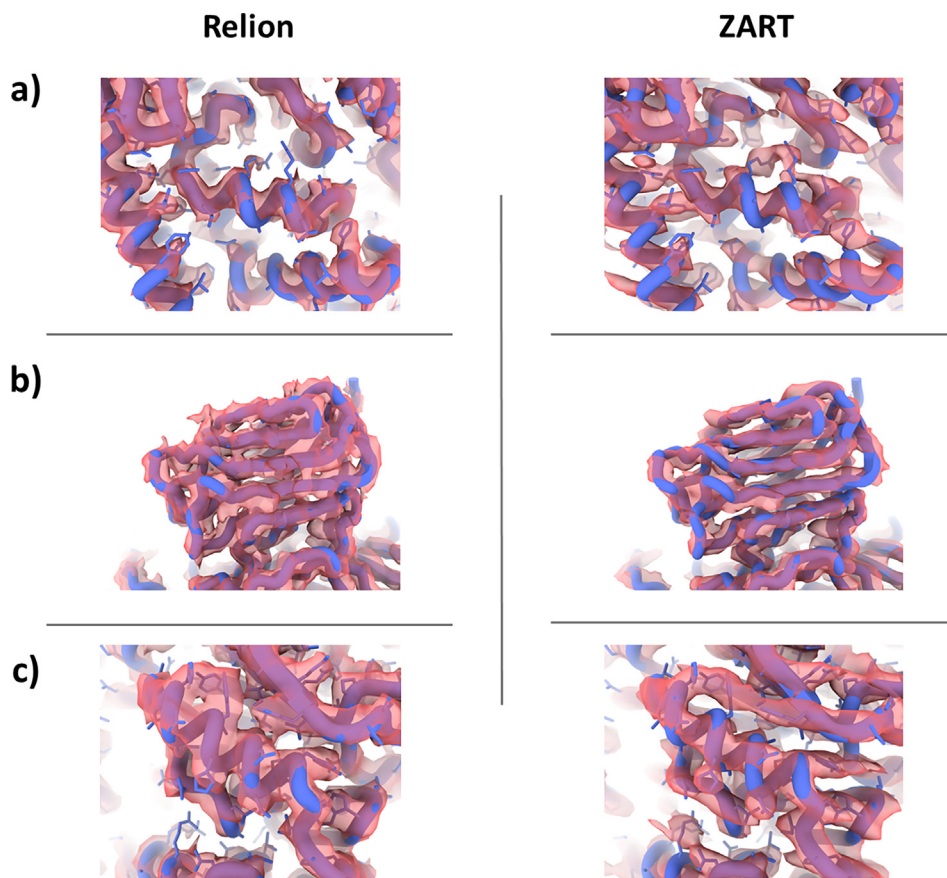


Figure 4. Comparison of Relion and ZART reconstructions (without sharpening) against the original atomic structure published with the EMPIAR-10391 dataset (PDB entry 6WBX). Thanks to the intrinsic sharpening characteristics of ZART, the reconstruction is able to better define different map regions at similar map thresholds. The panels were generated with ChimeraX software,¹⁷ and histogram thresholds were set to 0.012 and 0.255 for Relion and ZART respectively.

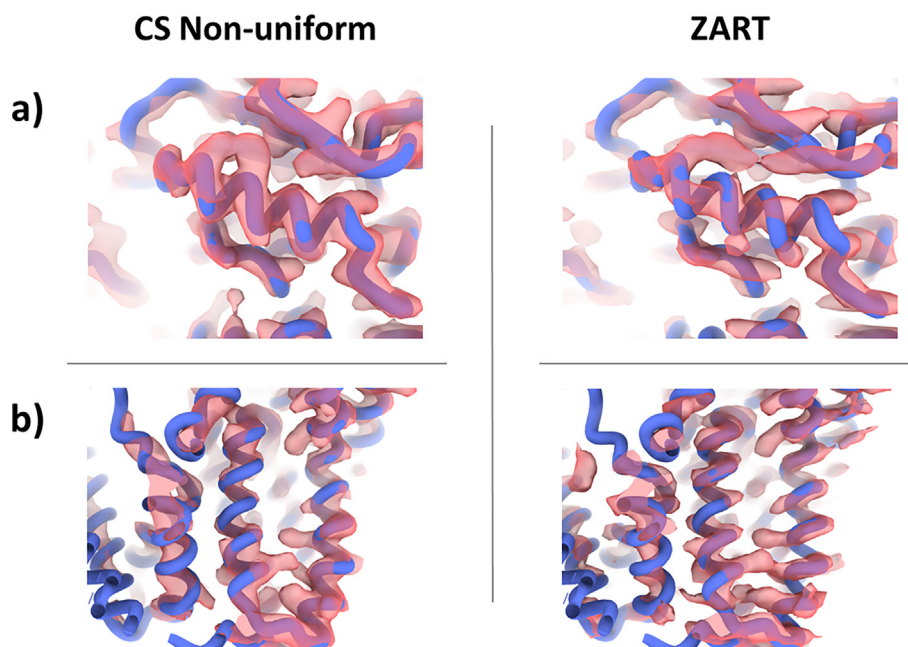


Figure 5. Comparison of CryoSPARC non-uniform refinement and ZART reconstructions (without sharpening) against the original atomic structure published with the EMPIAR-10391 dataset (PDB entry 6WBX). The comparison shows that the de-blurring applied by the two algorithms leads to similar results, although ZART is able to define better some structural features. The panels were generated with ChimeraX software,¹⁷ and histogram thresholds were set to 0.85 and 0.255 for CryoSPARC and ZART respectively.

Therefore, we decided to evaluate the ability of ZART to revert conformational changes and improve motion-blurred areas with the EMPIAR-10028 dataset. To that end, we processed the submitted data inside Scipion to get a set 50 k particles with CTF and alignment information. Both, the alignment and CTF were estimated with CryoSPARC²¹ in two independent runs, followed by a consensus step²² to improve the accuracy of the measurements. The sampling rate of the images fed to the reconstruction algorithms was 1.34 Å/px.

We estimated the per-particle conformational changes with the Zernike3D algorithm.¹⁰ This method relies on a mathematical basis able to express a deformation field that can be used to approximate any particle conformation in 3D at the level of CryoEM maps or atomic structures.

In this example, the estimated deformation fields are considered during the reconstruction process to reduce the structural differences among the particles and reduce motion blur artefacts. The application of the deformation field compensation can be done either using the standard ZART reconstruction algorithm or its multiresolution version. However, in this case, and with the aim of reducing the number of parametric choices and concentrating only on the differences due to motion correction, the reconstruction of the ribosome was done using standard ZART.

The comparison of the map reconstructed with CryoSPARC²¹ and ZART with motion correction is provided in Figure 6(a). The maps were coloured according to their local resolution value estimated with MonoRes. As can be seen from the results, ZART provides better features (mostly on moving regions, such as the small subunit of the ribosome) thanks to the per-particle structural corrections.

Figure 6(b) shows the comparison of the local resolution histograms computed with MonoRes. Similarly to the results offered in the previous section, the application of ZART with motion correction increases the resolution of a larger number of voxels, mainly in the range from 3.0 Å to 4.5 Å. The average of the local resolution is also improved from 5.5 Å to 5.2 Å. The comparison of the FSC curves is also provided in Figure 6(c). In this case, ZART pushes the FSC around 0.1 Å.

Figure 7 shows a more detailed comparison of the differences between the CryoSPARC and ZART reconstructions in regions of high flexibility. The figure shows the comparison of the CryoSPARC and ZART reconstructions against the original structural model published together with the dataset. The structural model has not been further refined considering any of the two maps to simplify the comparison. Thanks to the correction of the ribosome motions, ZART is able to define better different features in flexible regions such as

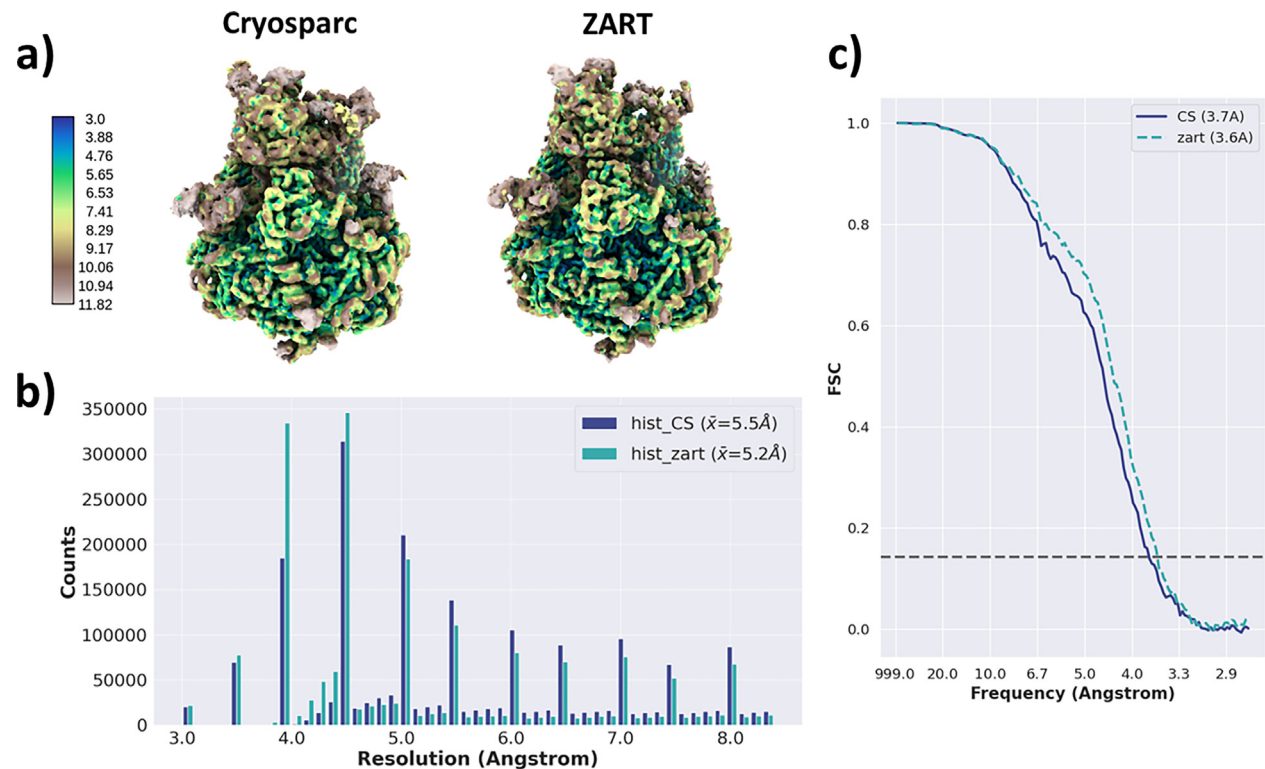


Figure 6. Comparison of ZART and CryoSPARC reconstructions for the *P. falciparum* 80S ribosome from EMPIAR-10028 dataset. (a) shows a comparison of the CryoSPARC and ZART reconstructions sharpened by DeepEMhancer¹⁸ software. In the case of ZART, we perform a standard reconstruction with no multiresolution, but correcting the motion blur artefacts thanks to the Zernike3D deformation fields computed for each particle in the dataset. The map comparison shows that ZART provides a sharper representation with improved local resolution. (b) shows the local resolution histograms of the maps in (a) computed with MonoRes.¹⁹ The histograms show that ZART pushes a large number of voxels towards the high-resolution regime, being most of the voxels at the range 3.0 Å–4.5 Å. (c) shows the FSC curves of the maps shown in (a). For this case, ZART improves the overall resolution of the map by around 0.1 Å.

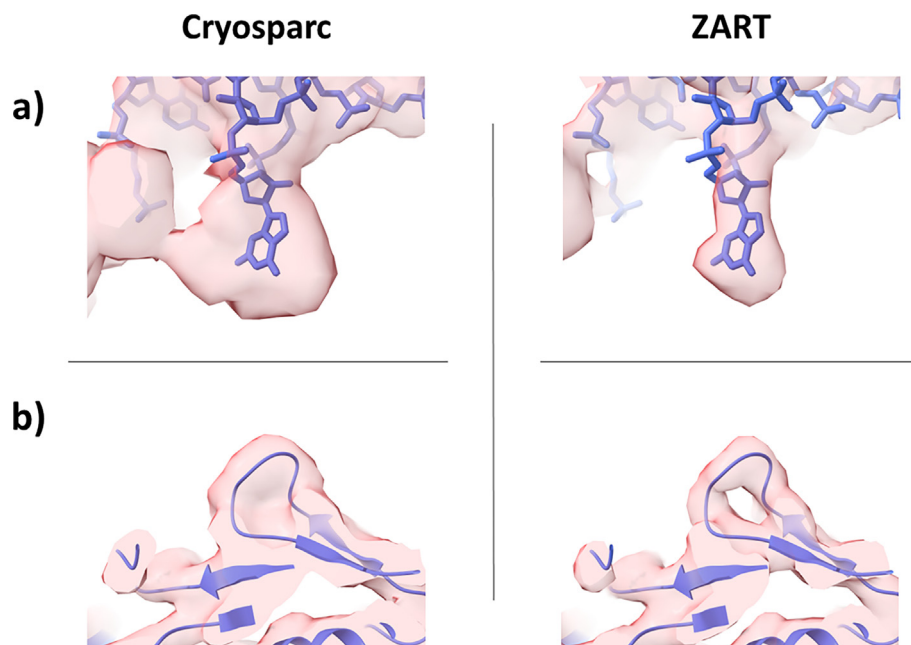


Figure 7. Comparison of CryoSPARC and ZART reconstructions (with motion blur correction) against the original atomic structure published with the EMPIAR-10028 dataset (PDB entries 3J79 and 3J7A). Thanks to the correction of each particle conformation during the reconstruction, ZART is able to define better the features usually hidden due to motion blur artefacts. The panels were generated with ChimeraX software.¹⁷

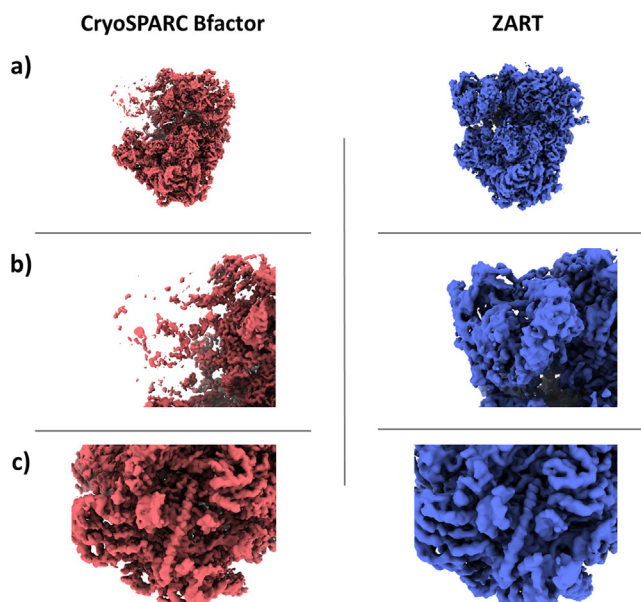


Figure 8. Comparison of Bfactor corrected CryoSPARC reconstruction and ZART reconstructions (with motion blur correction). Both maps were post-process with DeepEMHancer¹⁸ software in order to decrease the noise of the maps and improve the interpretability of the results. The Bfactor correction applied to the Relion reconstruction enhances the features located mostly on the core of the ribosome as observed in (c) compared to ZART. However, flexible regions are not properly recovered, being lost due to their low SNR as shown in (a) and (b). In contrast, ZART motion correction helps reducing the motion blur induced by the molecular motions, thus yielding a more complete map even after applying its intrinsic sharpening and DeepEMHancer. The panels were generated with ChimeraX software.¹⁷

Table 1 Execution times for the ZART algorithm (note that the standard ZART version has been the one used for this reporting). N and M are the parameters defining the size of the Zernike3D base. We show the performance for $N = 3$ and $M = 2$, which corresponds to our most common choice of parameters when estimating deformation fields, and also the one for $N=-$ and $M=-$, which refers to the case when no motion correction is considered during the reconstruction. We observe that for one million particles of size 300x300 pixels the time goes from three-quarters of an hour to an hour and a half, depending on whether we consider or not the deformation field correction (on a powerful CPU server, the GPU version is still under development).

Performance metrics for ZART algorithm			
Image size	N	M	Iteration time 10^6 particles (hours - 150 threads)
300	3	2	1.407
300	-	-	0.746

loops and key residues, as shown in the figure, producing information that can now be used to further refine the traced structural model starting from the ZART map, leading to better results.

We also provide a more in-depth comparison of the signal of different map slices in [Supplementary Figure 2](#). As can be seen from the slice comparison, ZART provides a better representation of the features in the map, including the small subunit of the ribosome thanks to the heterogeneity correction coming from the Zernike3D deformation fields. Negative values have been excluded from the slices of CryoSPARC and ZART to enhance the visualization of the protein signal.

In addition, ZART reconstruction was compared against the previous CryoSPARC map after

applying a Bfactor to better assess the intrinsic sharpening capabilities of ZART against other sharpening methods. In addition, both ZART and CryoSPARC Bfactor-corrected maps were post-process with DeepEMHancer using wide target mode, in order to decrease the noise of the reconstruction while keeping as reliably as possible the original information of the volumes.

The comparison is provided in [Figure 8](#). The application of the Bfactor improves the quality of the ribosome core features, as this region is more rigid. However, the low SNR of the high-flexibility regions prevents them from being recovered appropriately after the Bfactor correction.

In contrast, the motion correction applied by ZART provides a better definition of the moving regions in the ribosome, which are properly

recovered even after applying the intrinsic sharpening of ZART and DeepEMHancer.

Table 1 provides some metrics about the computational performance of ZART.

Conclusions

We have introduced a new ART-based reconstruction method called ZART in this manuscript. Thanks to the modification in the interpolation scheme and volume recovery process, ZART is able to improve the local resolution of the reconstructed maps, without the need to further modify or refine the alignment and CTF information of the input particles.

Moreover, ZART can also create a map where the estimated Zernike3D deformation field per particle can be compensated for, or computationally “reverted”, thus reducing motion blurring during the reconstruction and increasing the resolution of very flexible areas in the macromolecule. Compared to the implementation introduced in,¹⁰ ZART reconstructions are no longer limited to the sampling rates used during the estimation of conformational landscapes. Therefore, is now possible to correct the heterogeneity-induced blurring of maps without compromising resolution, allowing ZART to reach high-resolution heterogeneity-corrected maps.

In addition, in this work we described the ZART multiresolution reconstruction approach, which takes advantage of the local resolution estimation of the maps reconstructed every two iterations to determine the resolution that should be used to further update different regions of the map. Multiresolution has proven to be a useful approach, as it leads to high-quality results without the need for so many ZART iterations, reducing the execution times compared to the initial implementation introduced in.¹⁰ Finally, a multiresolution approach has intrinsically the potential to be less prone to be trapped in local minima.²³ Although we have not observed this effect in the cases presented in this work, these characteristics remain to be further explored in other data sets.

Methods

The Algebraic Reconstruction Technique

We express the reconstructed map in a series expansion with a basis function $b(\mathbf{r})$:

$$V(\mathbf{r}) = \sum_j x_j b\left(\frac{\mathbf{r} - \mathbf{r}_j}{\sigma}\right) \quad (1)$$

where \mathbf{r} is a coordinate in the 3D space, \mathbf{r}_j represents the location of the j -th basis function and σ its scale. The basis functions chosen are Gaussian functions. If all sigmas are equal, then the expression above can be expressed as a convolution.

$$V(\mathbf{r}) = \left(\sum_j x_j \delta(\mathbf{r} - \mathbf{r}_j) \right) \star b\left(\frac{\mathbf{r}}{\sigma}\right) \quad (2)$$

These Gaussians are distributed in a regular, rectangular grid with a step of σ pixels. Our motivation for this choice is that Gaussians can approximate a partition of unity with a very low error²⁴ when the spacing between two Gaussians is equal to their standard deviation. Partition of unity is an important property of basis functions that guarantees that they can reproduce any Sobolev function of the space W^2 (that is, sufficiently smooth, square-integrable functions). Additionally, Gaussians are spherically symmetric and their projection does not depend on the projection direction, making them very computationally efficient.

The projection of this volume onto an image is modelled as the line integral of this map. The projection direction and the in-plane shift are given by an Euler matrix, E , that acts on the spatial coordinates.²⁵ We define the auxiliary matrix H_t that will help to transform the 2D coordinate of the image into a 3D coordinate of the volume and will also help to perform the line integral as

$$\tilde{H}_t^T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & t \end{pmatrix} \quad (3)$$

Then, the projection is given by

$$\begin{aligned} I(\mathbf{s}) &= \int V(E^{-1} \tilde{H}_t^T \mathbf{s}) dt \\ &= \int \left(\sum_j x_j b\left(\frac{E^{-1} \tilde{H}_t^T \mathbf{s} - \mathbf{r}_j}{\sigma}\right) \right) dt \\ &= \sum_j x_j \int \left(b\left(\frac{E^{-1} \tilde{H}_t^T \mathbf{s} - \mathbf{r}_j}{\sigma}\right) \right) dt \end{aligned} \quad (4)$$

Tildes above represent homogeneous coordinates.

The projection of a 3D Gaussian is a 2D Gaussian of the same σ . Let us refer to the 2D Gaussian as G . Then, the projection above can be simplified to

$$I(\mathbf{s}) = \sum_j x_j G\left(\frac{\mathbf{s} - E\mathbf{r}_j}{\sigma}\right) \quad (5)$$

If there is an in-plane shift, \mathbf{s}_0 , then the model above has to be modified to

$$I(\mathbf{s}) = \sum_j x_j G\left(\frac{(\mathbf{s} - \mathbf{s}_0) - E\mathbf{r}_j}{\sigma}\right) \quad (6)$$

For a particular pixel \mathbf{s}_i we will have

$$I(\mathbf{s}_i) = \sum_j x_j G\left(\frac{(\mathbf{s}_i - \mathbf{s}_0) - E\mathbf{r}_j}{\sigma}\right) = \sum_j x_j a_{E, \mathbf{s}_0, ij} \quad (7)$$

where the coefficient $a_{E, \mathbf{s}_0, ij}$ represents the projection of the j -th basis of the volume onto the i -th pixel of the image whose alignment parameters are given by the matrix \tilde{E} and the in-plane shift \mathbf{s}_0 .

We may collect all experimental images into a single vector, \mathbf{I} , and construct the corresponding

matrix, A , with all the matrices for the individual projections. Then, the reconstruction problem boils down to solve a linear equation system of the form

$$A\mathbf{x} = \mathbf{I} \quad (8)$$

Among the possible algorithms to solve this equation system, we have chosen Block-ART as it provides a good balance between computational complexity and convergence speed.⁹

The 3D reconstruction process starts with an initial solution, $\mathbf{x}^{(0)}$, typically an empty vector of zeroes. Then, it iterates over the different experimental images and updates the k -th estimate of the 3D reconstruction according to

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \lambda_k \sum_i \frac{I_i - \mathbf{a}_i \cdot \mathbf{x}^{(k)}}{\|\mathbf{a}_i\|^2} \mathbf{a}_i \quad (9)$$

where \mathbf{a}_i is a vector with all the contributions of all the j basis functions onto the pixel i , and i goes over the selected image (different for each k). λ_k is a relaxation factor between 0 and 2.

The closer to 1 the faster the algorithm will converge, although this is only feasible if the experimental measurements are of good quality. In CryoEM, the Signal-to-Noise Ratio of the projections is in the order of 0.01. For this reason, λ is typically kept low.

Our output is the vector \mathbf{x}_j , that is, the Gaussian coefficients. By doing this, we are deconvolving the map with a Gaussian at the same time that we reconstruct it. In the multiresolution section below, we will show that the width of this Gaussian is locally defined and that we are, thus, doing a local deconvolution based on the local resolution of the map.

Zernike3D Algebraic Reconstruction Technique

One of the main reasons to choose ART as the basis for the ZART algorithm is the freedom it gives to supply prior knowledge during the reconstruction process. In the case of ZART, this information is related to the structural information of each particle, which is computed by the Zernike3D algorithm.

The Zernike3D method²⁶ is a novel approach able to describe per-particle conformational changes based on the estimation of a deformation field \mathbf{g}_L , so that the relationship between the deformed and the undeformed volume is

$$V_{\text{deformed}}(\mathbf{r}) = V_{\text{undeformed}}(\mathbf{r} + \mathbf{g}_L(\mathbf{r})) \quad (10)$$

where the deformation field is expressed in terms of the Zernike3D basis $Z_{l,n,m}$ and a series of Zernike3D coefficients $\alpha_{l,n,m}$ as:

$$\mathbf{g}_L(\mathbf{r}) = \sum_{l=0}^L \sum_{n=0}^N \sum_{m=-l}^l \begin{pmatrix} \alpha_{l,n,m}^x \\ \alpha_{l,n,m}^y \\ \alpha_{l,n,m}^z \end{pmatrix} Z_{l,n,m}(\mathbf{r}) \quad (11)$$

The estimation of the deformation field amounts to estimating the α coefficients in the equation

above. Once they are determined, we may use incorporate them in the series expansion to have a deformation, projection model (see Eq. 6):

$$I_{\text{deformed}}(\mathbf{s}) = \sum_j x_j G\left(\frac{(\mathbf{s} - \mathbf{s}_0) - E(\mathbf{r}_j + \mathbf{g}_L(\mathbf{r}_j))}{\sigma}\right) = \sum_j x_j a_{E, \mathbf{s}_0, \mathbf{g}_L, j} \quad (12)$$

That is, in the deformed projection, the coefficients of the undeformed map, x_j , have to be projected to a point given by the projection direction and the deformation field. The term $a_{E, \mathbf{s}_0, \mathbf{g}_L, j}$ now encodes the projection, in-plane shift, and deformation.

The main advantage of including the heterogeneity information in the ART reconstruction process through the Zernike3D deformation fields is the possibility to properly correct the non-rigid alignments associated with macromolecular motions. In a normal reconstruction process, the reconstruction volume is assumed to be the conformational state represented by all the particles in the dataset. However, the previous assumption does not hold when the macromolecule exhibits large degrees of flexibility, leading to a motion-induced blurring in the resulting map. The per-particle deformation fields introduced a non-rigid alignment able to correct for the heterogeneity inconsistencies of the particles, making the dataset more consistent during the reconstruction process and thus reducing motion blur artefacts.

The unknowns of the linear equation system in Eq. 8, x_j , refer to the undeformed map. However, the A matrix contains projections of the deformed images. In this way, we obtain the remarkable result that ZART can reconstruct an undeformed map with information coming from the deformed particles together with the deformation field. Still, we should warn that since the calculation of the deformation field itself will never be perfect due to the image noise, errors in this estimation will translate into the “estimated” undeformed map; in other words, the quality of the reconstructed undeformed map depends on the capacity of the Zernikes to faithfully reproduce the observed deformations.

In order to improve the aspect of the final maps, ZART only uses the Gaussian convolution for the computation of the volume projections that will be compared to the experimental image. Once the correction image has been computed, ZART will look at the Gaussian coefficients to update the reconstruction. Since the Gaussian coefficients are sharper than the Gaussian itself, the reconstructed map will be sharper, making it easier to choose an appropriate threshold to visualize the reconstruction.

Regarding the CTF, particles should be previously corrected (for example, by means of a Wiener2D filter) before inputting them in ZART. The current criterion to stop iterations in ZART is

by reaching a maximum number of iterations specified by the user (by default, 10). Still, the user can decide whether to save or not the partial reconstructions for every iteration, in order to analyze how the reconstruction evolves and determine which iteration is yielding better results.

Multiresolution reconstruction

In addition, we have exploited the almost partition of unity property of Gaussians to propose a multiresolution reconstruction scheme. The resolution of the reconstructed map is limited by the standard deviation of the basis of the series expansion. Large standard deviations will lead to lower-resolution maps as we are introducing a stronger low-frequency component during the reconstruction. But, because the separation between bases is equal to the standard deviation, we will also require fewer coefficients and they will be less affected by noise.

Our multiresolution implementation discretizes the possible σ 's to a finite set of possible values determined by the local resolution. Let us assume that we have M multiresolution levels, each one with a σ_m . Then, we have M superposing grids (at locations \mathbf{r}_j^m), and the series expansion above becomes:

$$V(\mathbf{r}) = \sum_m \left(\left(\sum_j x_j^m \delta(\mathbf{r} - \mathbf{r}_j^m) \right) \star b\left(\frac{\mathbf{r}}{\sigma_m}\right) \right) \quad (13)$$

Our multiresolution reconstruction algorithm follows the next steps:

1. Initially, ZART performs two ART iterations to compute an initial even/odd reconstruction with the particles. At this moment, a large and fixed Gaussian spacing and standard deviation are used.
2. The two half maps are used to measure the current local resolution of the reconstruction with MonoRes.¹⁹ We calculate the histogram of the local resolution and divide it into M equally populated bins. Then, we approximate the local resolution at the location \mathbf{r}_j by the centre of its local resolution bin, let us refer to it as $\hat{R}(\mathbf{r}_j) = \hat{R}_m$. We associate a Gaussian standard deviation to each one of these bin centres:

$$\sigma_m = \frac{\hat{R}_m}{2T_s}$$

being T_s the sampling rate of the particles involved in the reconstruction. In this way, the Nyquist resolution corresponds to a value $\sigma = 1.0$, which defines the size of the Gaussian fulfilling the partition of unity.

3. Once the local standard deviations have been computed, ZART performs another two ART iterations in multiresolution mode. In our implementation, for each of the M grids, only the x_j^m coefficients of the series expansion whose local resolution is associated with σ_m are updated, while all the rest remain as 0.
4. Steps 2 and 3 are then repeated according to the total number of ART iterations specified by the user.

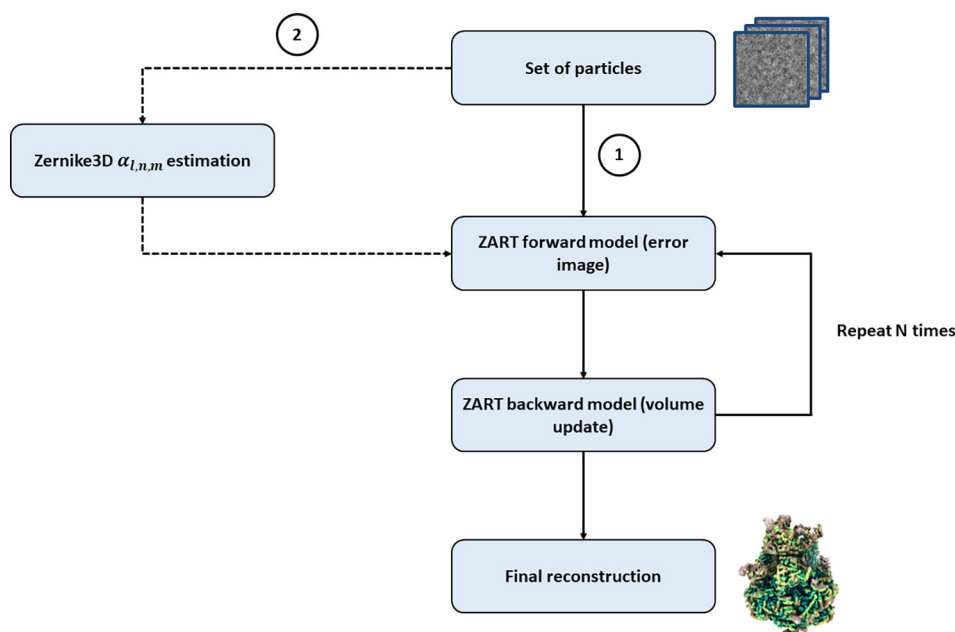


Figure 9. Diagram of the ZART reconstruction process in standard mode. (1) represents the steps followed when heterogeneity information is not available. (2) shows the motion blur correction workflow thanks to the estimation of the Zernike3D deformation fields for every particle in the dataset.

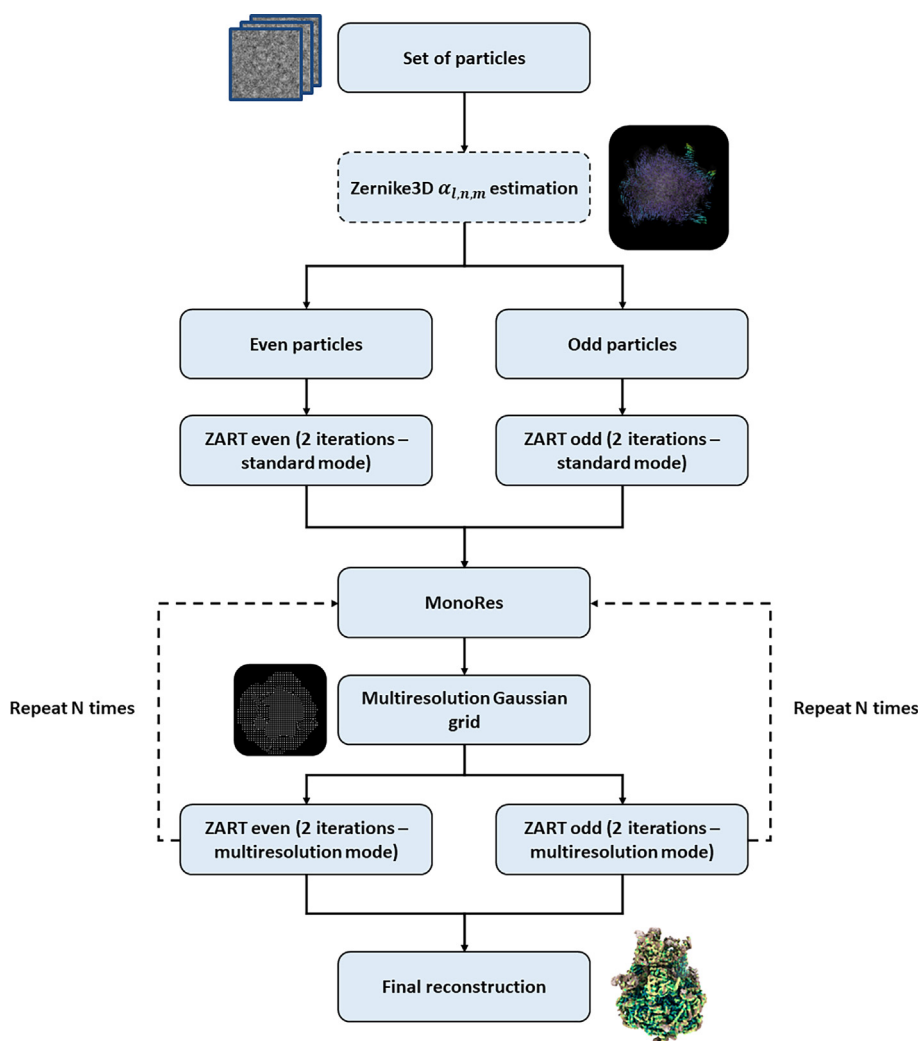


Figure 10. Diagram of the ZART reconstruction process in multiresolution mode. The estimation of the Zernike3D deformation fields is dotted, as it represents an optional step: if the estimation is computed, the multiresolution mode will reconstruct a motion blur-corrected version of the macromolecule.

When the multiresolution mode is selected, the user can determine the number of Gaussians (referred to as “levels” in the program) that will be used during the reconstruction (i.e. if $M = 3$ levels are chosen, the reconstruction will be able to use up to three Gaussians of different spacing and standard deviation).

A visual representation of the reconstruction process of both, standard and multiresolution modes is provided in [Figure 9](#) and [Figure 10](#).

Data Availability

The data used to test the ZART algorithm are publicly available in EMPIAR under the entries 10028 and 10391.

Code Availability

The ZART algorithm has been implemented in Xmipp¹⁴ and it is available through Scipion 3.0²⁷ under the plugin `scipion-em-flexutils`.

CRedit authorship contribution statement

D. Herreros: Conceptualization, Methodology, Software, Validation, Writing – original draft. **J. Kiska:** Software. **E. Ramirez:** Resources. **J. Filipovic:** Software, Supervision. **J.M. Carazo:** Supervision, Writing – review & editing. **C.O.S. Sorzano:** Supervision, Writing – review & editing.

Competing Interests

The authors declare no competing interests.

Funding

Funding is acknowledged from the Ministry of Science and Innovation through grants: Grant PID2019-104757RB-I00 funded by MCIN/AEI/10.13039/501100011033/ and “ERDF A way of making Europe”, by the “European Union; the ‘Comunidad Autónoma de Madrid’ through grant S2017/BMD-3817; and the European Union (EU) and Horizon 2020 through grants EnLaCES (H2020-MSCA-IF-2020, Proposal: 101024130 to JMK), HighResCells (ERC-2018-SyG, Proposal: 810057) and iNEXT-Discovery (Proposal: 871037). This work has also been supported by the NIH/NIGMS (No. 1R01GM136780-01 to RRL) and AFSOR A9550-21-1-0317.

Appendix A. Supplementary Data

Supplementary data associated with this article can be found, in the online version, at <https://doi.org/10.1016/j.jmb.2023.168088>.

Received 26 October 2022;

Accepted 1 April 2023;

Available online 7 April 2023

Keywords:

Cryo-Electron Microscopy (CryoEM);
spherical harmonics;
Zernike polynomials;
map reconstruction;
Zernike3D-based Algebraic Reconstruction Technique
(ZART)

References

- Carroni, M., Saibil, H.R., (2016). Cryo electron microscopy to determine the structure of macromolecular complexes. *Methods* **95**, 78–85.
- Scheres, S.H.W., Nuñez-Ramírez, R., Gomez-Yorente, Y., San Martín, C., Eggermont, P.P.B., Carazo, J.M., (2007). Modeling experimental image formation for likelihood-based classification of electron microscopy data. *Structure* **15** (10), 1167–1177.
- Jin, Q., Sorzano, C.O.S., de la Rosa-Trevín, J.M., Bilbao-Castro, J.R., Núñez-Ramírez, R., Llorca, O., Tama, F., Jonić, S., (2014). Iterative elastic 3D-to-2D alignment method using normal modes for studying structural dynamics of large macromolecular complexes. *Structure* **22**, 496–506.
- Zhong, E.D., Bepler, T., Berger, B., Davis, J.H., (2021). Cryodrgn: reconstruction of heterogeneous cryo-em structures using neural networks. *Nat. Methods* **18**, 176–185.
- Ludtke, S.J., Muyuan, C., (2021). Deep learning-based mixed-dimensional gaussian mixture model for characterizing variability in cryo-em. *Nat. Methods* **18**, 930–936.
- Frank, J., Abbas, O., (2016). Continuous changes in structure mapped by manifold embedding of single-particle data in cryo-EM. *Methods* **100**, 61–67.
- Punjani, A., Fleet, D.J., (2021). 3D flexible refinement: Structure and motion of flexible proteins from cryo-EM. *bioRxiv* **36**
- Lederman, R.R., Anden, J., Singer, A., (2020). Hyper-molecules: on the representation and recovery of dynamical structures for applications in flexible macromolecules in cryo-em. *Inverse Prob.* **36**
- Sorzano, C.O.S., Vargas, J., Otón, J., Vilas, J.L., Kazemi, M., Melero, R., del Caño, L., Cuenca, J., et al., (2017). A survey of the use of iterative reconstruction algorithms in electron microscopy. *BioMed Res. Int.* **1–17**, 2017.
- Herreros, D., Lederman, R.R., Krieger, J., Jiménez-Moreno, A., Martínez, M., Myška, D., Strelak, D., Filipovic, J., et al., (2023). Estimating conformational landscapes from Cryo-EM particles by 3D Zernike polynomials. *Nat. Commun.* **14**, 154.
- Sorzano, C.O.S., Vargas, J., Otón, J., Abrishami, V., de la Rosa-Trevín, J.M., Fernández-Alderete, A., Martínez-Rey, C., Marabini, R., et al., (2015). Fast and accurate conversion of atomic models into electron density maps. *AIMS Biophys.* **2** (1), 8–20.
- Wrapp, D., Wang, N., Corbett, K.S., Goldsmith, J.A., Hsieh, C., Abiona, O., Graham, B.S., McLellan, J.S., (2020). Cryo-em structure of the 2019-ncov spike in the prefusion conformation. *Science* **367** (6483), 1260–1263.
- Abrishami, V., Bilbao-Castro, J.R., Vargas, J., Marabini, R., Carazo, J.M., Sorzano, C.O.S., (2015). A fast iterative convolution weighting approach for gridding-based direct Fourier three-dimensional reconstruction with correction for the contrast transfer function. *Ultramicroscopy* **157**, 79–87.
- de la Rosa-Trevín, J.M., Otón, J., Marabini, R., Zaldívar, A., Vargas, J., Carazo, J.M., Sorzano, C.O.S., (2013). Xmipp 3.0: An improved software suite for image processing in electron microscopy. *J. Struct. Biol.* **184**, 321–328.
- Tan, Y.Z., Zhang, L., Rodrigues, J., Zheng, R.B., Giacometti, S.I., Rosário, A.L., Kloss, B., Dandey, V.P., et al., (2020). Cryo-em structures and regulation of arabinofuranosyltransferase aftd from mycobacteria. *Molecular Cell* **78** (4), 683–699. al et..
- Scheres, S.H.W., (2012). Relion: Implementation of a bayesian approach to cryo-em structure determination. *J. Struct. Biol.* **180**, 519–530.
- Pettersen, E.F., Goddard, T.D., Huang, C.C., Meng, E.C., Couch, G.S., Croll, T.I., Morris, J.H., Ferrin, T.E., (2021). Ucsf chimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci.* **30** (1), 70–82.
- Sanchez-Garcia, R., Gomez-Blanco, J., Cuervo, A., Carazo, J.M., Sorzano, C.O.S., (2021). Deepenhancer: a deep learning solution for cryo-em volume post-processing. *Commun. Biol.* **4**, 874.
- Vilas, J.L., Gomez-Blanco, J., Conesa, P., Melero, R., de la Rosa-Trevin, J.M., Oton, J., Cuenca, J., Marabini, R., et al., (2018). Monores: Automatic and accurate estimation of local resolution for electron microscopy maps. *Structure* **26**, 337–344.

20. Wong, W., Bai, X., Brown, A., Fernandez, I.S., Hanssen, E., Condrón, M., Tan, Y.H., Baum, J., et al., (2014). Cryo-em structure of the Plasmodium falciparum 80s ribosome bound to the anti-protozoan drug emetine. *eLife* **3**, e03080.
21. Punjani, A., Rubinstein, J.L., Fleet, D.J., Brubaker, M.A., (2017). cryoSPARC: algorithms for rapid unsupervised cryo-em structure determination. *Nat. Methods* **14**, 290–296.
22. Sorzano, C.O.S., Jiménez-Moreno, A., Maluenda, D., Martínez, M., Ramírez-Aportela, E., Krieger, J., Melero, R., Cuervo, A., et al., (2022). On bias, variance, overfitting, gold standard and consensus in single-particle analysis by cryo-electron microscopy. *Acta Crystall. Section D* **78** (4), 410–423.
23. Thevenaz, P., Ruttimann, U.E., Unser, M., (1998). A pyramid approach to subpixel registration based on intensity. *IEEE Trans. Image Process.* **7** (1), 27–41.
24. Bale, R., Grossman, J., Margrave, G., Lamoureux, M., (2002). Multidimensional partitions of unity and gaussian terrains. *CREWES Res. Report* **14**
25. Sorzano, C.O.S., Marabini, R., Vargas, J., Otón, J., Cuenca-Alba, J., Quintana, A., de la Rosa-Trevín, J.M., Carazo, J.M., (2014). Computational Methods for Three-Dimensional Microscopy Reconstruction, chapter Interchanging geometry information in electron microscopy single particle analysis: mathematical context for the development of a standard. Springer.
26. Herreros, D., Lederman, R.R., Krieger, J., Jiménez-Moreno, A., Martínez, M., Myška, D., Strelak, D., Filipovic, J., et al., (2021). Approximating deformation fields for the analysis of continuous heterogeneity of biological macromolecules by 3D Zernike polynomials. *IUCrJ* **8** (6), 992–1005.
27. de la Rosa-Trevín, J.M., Quintana, A., del Cano, L., Zaldívar, A., Foche, I., Gutiérrez, J., Gómez-Blanco, J., Burguet-Castell, J., et al., (2016). Scipion: A software framework toward integration, reproducibility and validation in 3D electron microscopy. *J. Struct. Biol.* **195** (1), 93–99.

