nature methods

Brief Communication

https://doi.org/10.1038/s41592-023-01874-3

New measures of anisotropy of cryo-EM maps

Received: 22 April 2022

Accepted: 5 April 2023

Published online: 29 May 2023

Check for updates

Jose-Luis Vilas [1] & Hemant D. Tagare [1,2]

We propose two new measures of resolution anisotropy for cryogenic electron microscopy maps: Fourier shell occupancy (FSO), and the Bingham test (BT). FSO varies from 1 to 0, with 1 representing perfect isotropy, and lower values indicating increasing anisotropy. The threshold FSO = 0.5 occurs at Fourier shell correlation resolution. BT is a hypothesis test that complements the FSO to ensure the existence of anisotropy. FSO and BT allow visualization of resolution anisotropy. We illustrate their use with different experimental cryogenic electron microscopy maps.

Cryogenic electron microscopy (cryo-EM) aims to elucidate the structure of proteins and macromolecular complexes. Preferred orientation of the particles on the grid, or the subsequent image analysis, can cause the resolution of a Cryo-EM structure to be anisotropic[1–4]: such structures have a directionally variable signal-to-noise ratio[5,6]. The usual measures of resolution—Fourier shell correlation (FSC)-based global resolution[7] and local resolution[8]—cannot assess anisotropy. Sphericity has been proposed as a measure of anisotropy[1], but sphericity is difficult to interpret and does not indicate the range of spatial frequencies over which the reconstruction is anisotropic.

Here, we propose two new anisotropy metrics: the FSO, and a hypothesis test of anisotropy based on the classical BT in spherical statistics[9]. A practitioner can use their properties explained later to understand the quality of cryo-EM maps. Both metrics depend on the directional FSC of a structure, which is defined as follows: Let $\mathbf{d}$ be a direction (expressed as a unit length vector along a half-infinite ray from the origin), and let $C_{\mathbf{d}}$ be a solid cone with half-angle Ω and having $\mathbf{d}$ as its axis (Fig. 1a). Set $D = S(\omega) \cap C_{\mathbf{d}}$, Then the directional FSC in $S(\omega)$, along the direction $\mathbf{d}$ is

$$dFSC(\omega, \mathbf{d}) = \frac{\int_D w(\mathbf{u})\,\mathrm{Re}\,(X^*(\mathbf{u})Y(\mathbf{u}))\,d\mathbf{u}}{\sqrt{\int_D w(\mathbf{u})X^*(\mathbf{u})X(\mathbf{u})d\mathbf{u}}\sqrt{\int_D w(\mathbf{u})Y^*(\mathbf{u})Y(\mathbf{u})d\mathbf{u}}}, \quad (1)$$

where $w(\mathbf{u})$ is a non-negative, continuous, real-valued 'weight' function satisfying $\int_D w(\mathbf{u})d\mathbf{u} = 1$. We set the weight function to the normalized Gaussian function $w(\mathbf{u}) = \frac{1}{C}e^{\frac{-(\cos\theta-1)^2}{\sigma^2}}$, where the constant C is determined by $C = \int_D w(\mathbf{u})d\mathbf{u} = 1$, and θ is the angle between $\mathbf{u}$ and the cone axis $\mathbf{d}$. The constant σ is the width of the Gaussian function and is set to $\sigma = (1 - \cos\Omega)/2$where Ω is the cone half-angle. This effectively reduces w to zero on the boundary of D. Signal-to-noise considerations suggest that a cone half-angle of 17° is optimal (Supplementary Information). This value is used in all numerical experiments below.

The FSO in shell $S(\omega)$ is defined as the fractional volume of the shell where $dFSC(\omega, \mathbf{d})$ is greater than or equal to a threshold χ, which is nominally set to $\chi = 0.143$.

$$FSO(\omega) = \frac{\int_{S(\omega)} 1(dFSC(\omega, \mathbf{u}/\parallel \mathbf{u} \parallel) \geq \chi)d\mathbf{u}}{\int_{S(\omega)} d\mathbf{u}}, \quad (2)$$

where 1(.) is the indicator function taking value 1 when $dFSC(\omega, \mathbf{d}) > \chi$, and 0 otherwise.

The FSO has a number of interesting properties (Supplementary Information contains calculations and evaluations of these claims): FSO as a function of ω is approximately 1 for low frequency shells, and drops to 0 at high frequency shells (similar to the familiar FSC function), see Fig. 1b. In the range $0.9 < FSO(\omega) \leq 1$, most directional FSCs are above the threshold χ, and the structure may be considered isotropic. In the range $0.1 \leq FSO(\omega) \leq 0.9$ the fraction of directional FSC's above the threshold χ begin to fall, and the reconstruction becomes increasingly anisotropic. This is the anisotropy transition zone. In addition to these properties, it turns out that FSO takes a value of 0.5 in the global resolution shell (see Supplementary Information for an explanation). Thus, the FSO provides an efficient summary of the reconstruction quality – it shows where the reconstruction is isotropic, where the reconstruction becomes anisotropic, and it also shows the global resolution, all in a single plot. When the structure is isotropic, FSO exhibits a step function-like behavior, which is the step occurring in the global resolution shell (see Supplementary Information for an explanation).

The FSO is a visual display of anisotropy. The BT complements this by providing a statistical measure (P value). In all our experiments FSO and BT are in agreement.

[1]Department of Radiology and Biomedical Imaging, Yale University, New Haven, CT, USA. [2]Department of Biomedical Engineering, Yale University, New Haven, CT, USA. e-mail: jlvilas@cnb.csic.es; hemant.tagare@yale.edu

https://doi.org/10.1038/s41592-023-01874-3

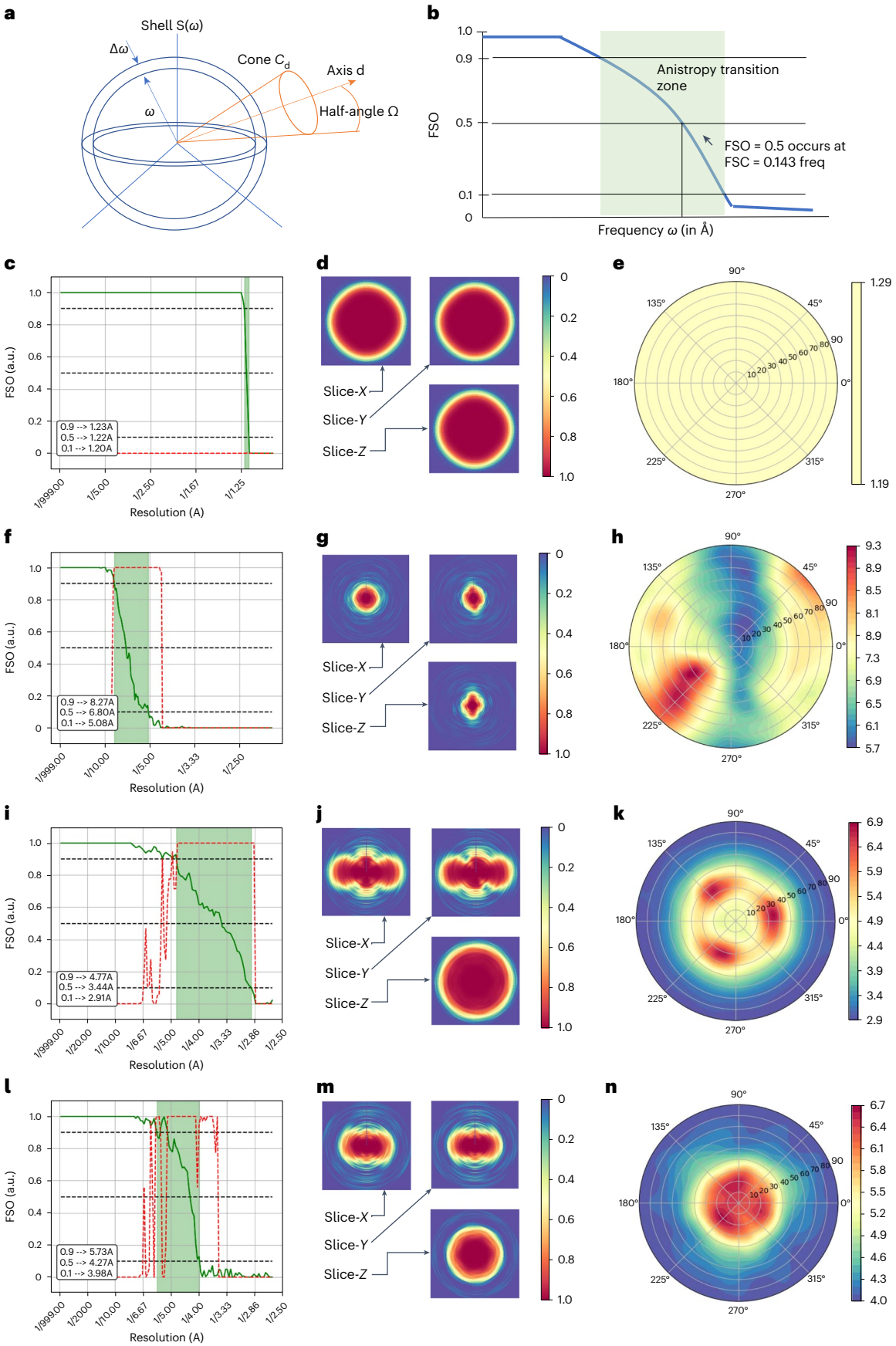

Fig. 1 | Anisotropy analysis. a, Scheme of directional FSC measurement. **b**, Description of the FSO curve. **c**–**e**, Apoferritin. **f**–**h**, Human CMG bound to AND-1. **i**–**k**, Influenza hemagglutinin trimer untilted map. **l**–**n**, Influenza hemagglutinin trimer tilted map. **c**,**f**,**i**,**l**, FSO (in continuous) and *P* value for the BT (dashed). **d**,**g**,**j**,**m**, Central slices along the *X*, *Y* and *Z* directions of the 3DFSC (FSC threshold of 0.143). **e**,**h**,**k**,**n**, DR as a function of the zenith and azimuth angles of the direction. a.u., arbitrary units.

https://doi.org/10.1038/s41592-023-01874-3

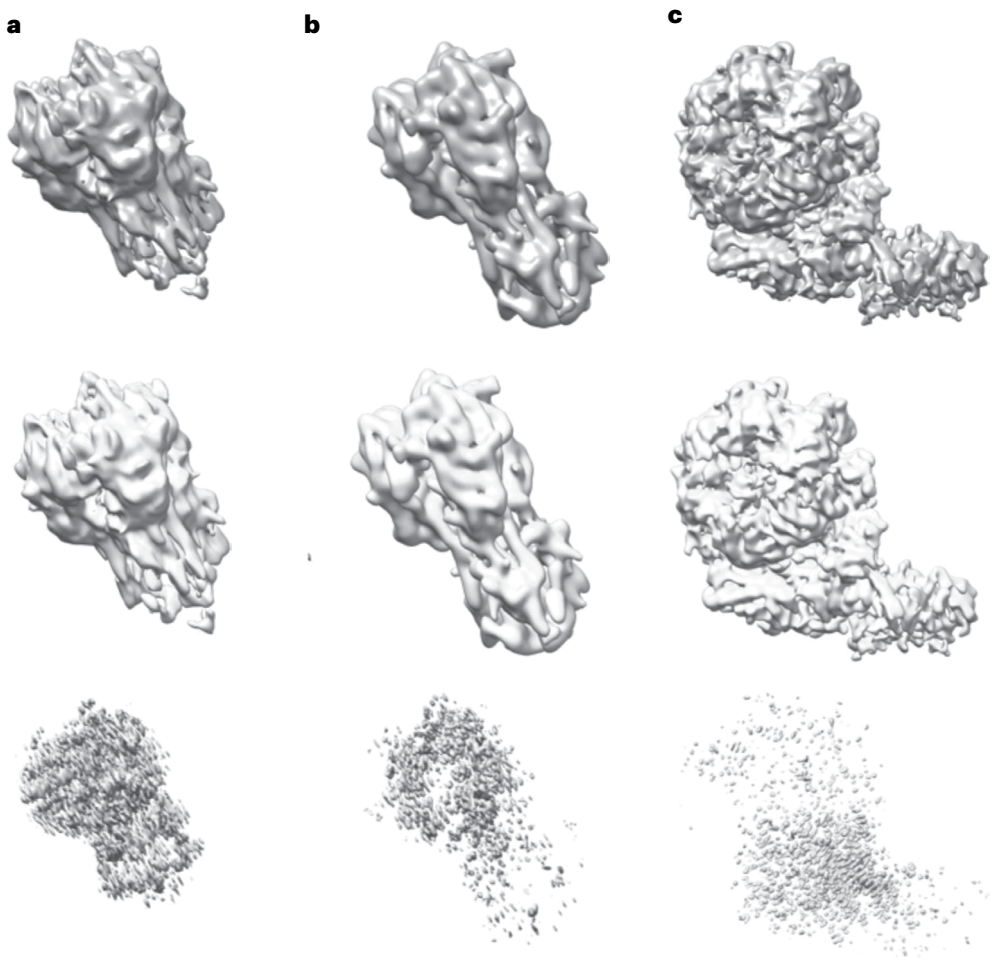

Fig. 2 | Effect of anisotropy on cryo-EM reconstruction. a–**c**, Original reconstruction (top), isotropic reconstruction obtained by low pass filtering at the resolution given by the threshold FSO = 0.9 (middle), absolute value of the difference of the two (down). **a**, Untilted reconstruction of the influenza hemagglutinin trimer. **b**, Tilted reconstruction of the influenza hemagglutinin trimer. **c**, Human CMG bound to AND-1.

The (BT) is a precise hypothesis test of (an-)isotropy that we adopt from spherical statistics[9] The classic BT for uniform distribution of points on a sphere in $k-$ dimensions is as follows: The data are points (column vectors) $\mathbf{x}_i$, $i = 1, \cdots, N$ in a sphere in $k-$ dimensions (that is $\mathbf{x}_i$ is a $k \times 1$ vector satisfying $\mathbf{x}_i^T \mathbf{x}_i = 1$ for $i = 1, \cdots, N$). From this data, we calculate the mean of the outer product $T = \frac{1}{N} \sum_{i=1}^{N} \mathbf{x}_i \mathbf{x}_i^T$, which is a $k \times k$ matrix, and the statistic $B = \frac{k(k+2)}{2} N(\text{ tr } T^2 - \frac{1}{k})$, where tr denotes the trace of a matrix. The theory of the BT[9] shows that S has a $\chi^2_{(k-1)(k+2)/2}$ distribution, if the points $\mathbf{x}_i$ are distributed uniformly on the sphere. Thus, evaluating the P value of B using a table of χ^2 distribution, and comparing it to 0.05 evaluates whether the points are uniformly distributed on the sphere. If the P value is greater than the significance value (0.05), then the uniformity hypothesis is rejected.

We adapt the BT as follows: For every direction $\mathbf{d}$ in the Fourier shell $S(\omega)$ we calculate whether $dFSC(\omega, \mathbf{d})$ is greater than 0.143. All directions $\mathbf{d}$ for this is true are indicated as points $\mathbf{x}_i$ on a unit sphere, the points being located at the intersection of the half-ray in direction $\mathbf{d}$ and the sphere. The B statistic is calculated for these $\mathbf{x}_i$ using $k = 3$. If the $p-$ value of the statistic is plotted as function of ω, then the range of ω where the $p-$ value exceeds 0.05 is the range in which the reconstruction is anisotropic. Below, we show empirically that this range of ω's is similar to the FSO transition zone.

Next, we demonstrate the use of *FSO*, and BT, to characterize Cryo-EM structures using gold-standard half-maps deposited in the EMDB. We also display the directional resolution (DR)[1,3] and the three-dimensional FSC (3DFSC)[1]. These provide additional visualization of the anisotropy.

The first structure apoferritin EMDB-11638 (ref. 10) has octahedral symmetry, O, and a resolution of 1.22 Å (FSC at 0.143). This volume is highly isotropic, as the FSO and BT confirm (Fig. 1c). The sphericity of the map is 1. The 3DFSC for the map is uniform Fig. 1d. The resolution of the shell at which $FSO = 0.5$ is the global FSC resolution of the map. The FSO and the BT required 7 s of central processing unit (CPU) computation. Sphericity took 1 h 14 min for the CPU implementation and 3.81 min for the graphics processing unit (GPU) implementation.

The next structure is the human CMG bound to AND-1 (ref. 11) EMD-10621 (resolution 6.77 Å). The FSO ($\chi = 0.143$; Fig. 1f) has an anisotropy transition region of (8.27, 5.08) Å. BT shows anisotropy in the range of (8.51, 4.37) Å. Further, $FSO(\omega) = 0.5$ occurs at 6.8 Å, the global FSC resolution. Figure 1g,h shows the 3DFSC and the DR, the latter as a function of zenith and azimuth angles of the directions. Note the low resolutions around an azimuth angle of 225°, and the relatively higher resolutions around 90° and 270°. The sphericity of this map is 0.84 (ref. 1). The FSO, BT and resolution distribution took 10 s to compute on the CPU. Sphericity took 22.5 min for the CPU implementation and 99 s for the GPU implementation (https://github.com/nysbc/anisotropy/).

The influenza hemagglutinin trimer[1] has preferred directions on the grid, which induce anisotropy. To reduce anisotropy, the original publication tilted the sample. As a consequence, there are two

datasets: One with the sample untilted (EMPIAR-10097, resolution of 3.4 Å) and another with the sample tilted by 40° (EMPIAR-10097, resolution of 4.4 Å). The FSOs for both maps are in Fig. 1i (untilted) and Fig. 1l (tilted)). The transition zones are (4.77, 2.91) Å (untilted) and (5.73, 3.98] Å (tilted). BT shows anisotropy in (5.58, 2.84] Å (untilted) and (5.87, 3.98) Å (tilted). The sphericities are 0.84 (untilted) and 0.91 (tilted). All results show that tilting reduces anisotropy. The DR plots show the anisotropy, and its reduction with tilting (Fig. 1k,n). The resolutions of the shells at which FSO = 0.5 are 3.44 Å (untilted) and 4.27 Å, closely matching the global FSC. The computational times were 2 s for FSO and BT, and 47 min (untilted, CPU), 16.81 min (tilted, CPU), 2.45 min (untilted, GPU) and 1.03 min (tilted, GPU).

Finally, we visualized the difference between the original map and the anisotropy-eliminated map to understand how anisotropy manifested in the original map. All of the above maps were filtered at the anisotropy onset frequency, subtracted from the original map, and the absolute value of the result visualized (Fig. 2). Each column of Fig. 2 shows results for one map. The original maps are in the first row. The second row contains the filtered maps, and the third row contains the absolute value of the difference of the two. Directionally elongated blobs in the difference maps show the anisotropy in the original maps. This effect is especially strong in the untilted reconstruction of the influenza hemagglutinin (Fig. 2a). Higher magnification images of the difference maps are available in the Supplementary Information.

An experimentalist can use FSO to understand how anisotropy affects the reconstruction at different resolutions. Changes in the anisotropy transition zone are especially useful to assess the effect of stage-tilting: Stage-tilting is not likely to be useful if most of the anisotropy reduction occurs in the low-resolution part.

Difference maps, such as those in Fig. 2, can be used to visualize anisotropic parts of reconstruction, and to assess whether a docked model (especially side chains) has directional support. In principle, it is also possible to compare a map with a model using FSO (Supplementary Information).

It is unclear whether anisotropy can be fixed computationally. Anisotropic filters, like a 3DFSC filter, cannot recover the missing directional information, but may be able to reduce anisotropic noise. Nonlinear sharpening algorithms, like DeepEMhancer[12], may be able to improve or restore isotropy.

In conclusion, FSO and BT provide useful metrics for evaluating reconstruction anisotropy. We suggest using the endpoints of the anisotropy transition zone as numbers that summarize the map anisotropy.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41592-023-01874-3.

References

1. Tan, Y. Z. et al. Addressing preferred specimen orientation in single-particle cryo-EM through tilting. *Nat. Methods* **14**, 793–796 (2017).
2. Vilas, J. L., Tagare, H. D., Vargas, J., Carazo, J. M. & Sorzano, C. O. S. Measuring local-directional resolution and local anisotropy in cryo-EM maps. *Nat. Commun.* **11**, 55 (2020).
3. Aiyer, S., Zhang, C., Baldwin, P. R. & Lyumkis, D. Evaluating local and directional resolution of Cryo-EM density maps. *Methods Mol. Biol.* **2215**, 161–187 (2021).
4. Sorzano, C. O. S. et al. Algorithmic robustness to preferred orientations in single particle analysis by Cryo-EM. *J. Struct. Biol.* **213**, 107695 (2021).
5. Penczek, P. A. Three-dimensional spectral signal-to-noise ratio for a class of reconstruction algorithms. *J. Struct. Biol.* **128**, 34–46 (2002).
6. Penczek, P. A. Chapter three—resolution measures in molecular electron microscopy. *Methods Enzymol.* **482**, 73–100 (2010).
7. Sorzano, C. O. S. et al. A review of resolution measures and related aspects in 3D electron microscopy. *Prog. Biophys. Mol. Biol.* **124**, 1–30 (2017).
8. Vilas, J. L. et al. Local resolution estimates of cryo-EM reconstructions. *Curr. Opin. Struct. Biol.* **64**, 74–78 (2020).
9. Mardia, K. V. & Jupp, P. E. *Directional Statistics* (Wiley, 1999).
10. Nakane, T. et al. Single-particle cryo-EM at atomic resolution. *Nature* **587**, 152–156 (2020).
11. Rzechorzek, N. J., Hardwick, S. W., Jatikusumo, V. A., Chirgadze, D. Y. & Pellegrini, L. Cryo-EM structures of human CMG-ATPγS-DNA and CMG-AND-1 complexes. *Nucleic Acids Res.* **48, D1**, 6980–6995 (2020).
12. Sanchez-Garcia, R., Gomez-Blanco, J., Cuervo, A., Sorzano, C. O. S. & Vargas, J. DeepEMhancer: a deep learning solution for cryo-EM volume post-processing. *Commun. Biol.* **4**, 874 (2021).
13. 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).
14. de la Rosa-Trevin, J. M. et al. Scipion: a software framework toward integration, reproducibility, and validation in 3D electron microscopy. *J. Struct. Biol.* **195**, 93–99 (2016).
15. Sorzano, C.O.S. et al. Image processing tools for the validation of Cryo-EM. *Faraday Discuss.* **240**, 210–227 (2022).

Reporting summary

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

Data availability

The maps used in this work were taken from the Electron Microscopy Data Bank under accession codes EMD-11638 for the apoferritin; EMD-10621 for the human CMG bound to AND-1, and Electron Microscopy Public Image Archive under accession codes EMPIAR-10096 and EMPIAR-10097 for the influenza hemagglutinin trimer maps. Data used in the Supplementary Information are EMD-10659, EMD-2984, EMD-10691, EMD-22854, EMD-10525, EMD-22949, EMD-11220, EMD-22777 and EMD-22963.

Code availability

Our software can be found in GitHub (https://github.com/Vilax/FSO/), Xmipp[13] (https://github.com/I2PC/xmipp/), Scipion 3.0 (ref. 14) (https://scipion.i2pc.es/) and the validation server[15] (https://biocomp.cnb.csic.es/EMValidationService/).

Acknowledgements

This research was supported by the National Institutes of Health grant R01GM125769 (H.D.T.). We thank F. Sigworth for discussions. We dedicate this paper to the memory of M. C. Vilas-Diaz and D. M. Tagare, and M. D. Tagare who passed away during the coronavirus disease 2019 pandemic.

Author contributions

The authors contributed equally to theory and experiments.

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/s41592-023-01874-3.

Correspondence and requests for materials should be addressed to Jose-Luis Vilas or Hemant D. Tagare.

Peer review information *Nature Methods* thanks Philip Baldwin and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editor: Arunima Singh, in collaboration with the *Nature Methods* team.

nature research

Corresponding author(s): JL Vilas

Last updated by author(s): Feb 8, 2023

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☒	☐ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☒	☐ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☒	☐ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐ A description of all covariates tested
☒	☐ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☒	☐ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	NA
Data analysis	Scipion-> https://github.com/I2PC/scipion, FSO -> https://github.com/Vilax/FSO

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

EMD-10659, EMD-0012, EMD-2984, EMD-10691, EMD-22854, EMD-10525, EMD-22949, EMD-11220, EMD-22777, EMD-22963, EMDB-10621, EMDB-11638, EMPIAR-10097, EMPIAR-10096

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	NA
Data exclusions	NA
Replication	NA
Randomization	NA
Blinding	NA

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging