The Shape of Things in Cryo-ET: Why Emojis Aren't Just for Texts

Valentin J. Maurer*,[1,2,a)] Marc Siggel*,[1,2] and Jan Kosinski[1,2,3,a)]
[1)] *European Molecular Biology Laboratory Hamburg, Notkestraße 85, 22607 Hamburg, Germany*
[2)] *Centre for Structural Systems Biology (CSSB), Notkestraße 85, 22607 Hamburg, Germany*
[3)] *Structural and Computational Biology Unit, European Molecular Biology Laboratory, Meyerhofstrasse 1, Heidelberg 69117, Germany*

(Dated: August 2023)

Detecting specific biological macromolecules in cryogenic electron tomography (cryo-ET) data is frequently approached by applying cross-correlation-based 3D template matching. Structural models or simple geometric shapes such as spheres, cylinders, or rectangles are frequently used templates. To reduce computational cost and noise, high binning is used to aggregate voxels prior to template matching. At high binning primarily low-frequency information remains and it has been shown that template matching fails if different macromolecules have overlapping low-frequency spectrums. Here, we combine these ideas and show in theory why geometric shapes can be used as templates and validate our findings using a detailed subtomogram average, a sphere, and the so-called poop emoji (💩) as templates to identify ribosomes in an annotated *S. cerevisiae* dataset. Our findings indicate that with current template-matching methods, macromolecules can only be detected with high precision if their shape and size are sufficiently different from the background, or if they are present in significantly higher abundance than other macromolecules with similar features. This implies that template matching under these conditions has by design low precision and recall. We discuss these challenges and propose potential enhancements for future template matching methodologies.

*Contributed equally

I. INTRODUCTION

Cellular cryogenic electron tomography (cryo-ET) has emerged as a key method to unravel the structural and spacial complexity of the cell. The 3D volume of a (sub)cellular region, called a tomogram, is reconstructed from 2D projection images acquired on a transmission electron microscope in many different orientations.[1,2] Macromolecular complexes can be identified in the tomogram, their spatial arrangement can be analyzed in the native environment, and their structure can potentially be determined to near-atomic resolution by averaging multiple sub-tomograms.[3–6]

However, identifying individual macromolecules in tomograms is challenging due to the missing wedge, which results from limitations on possible tilt angles of the specimen, low signal-to-noise ratio due to low electron doses used during acquisitions, and the crowded cellular environment of the cell. Because of these complications, segmentation and subsequent analysis of tomograms is still a difficult task and a major bottleneck for fully automated high-throughput analysis of large cryo-ET datasets.[1,6,7]

A commonly used approach to identify macromolecules within tomograms is so-called template matching, where a reference density of a macromolecule is used to localize the corresponding candidate positions within the tomogram.[8,9] To date, various packages have been developed to perform template matching such as PyTOM[10], STOPGAP[11], EMAN2[12], or DYNAMO[13]. All of these packages use cross-correlation based scoring metrics to identify macromolecules in tomograms (see II). Templates used for template matching range from simple shapes such as spheres, cylinders, and rectangles, which have been previously used to detect various cellular structures[14–17], to detailed maps generated from atomic structures or experimentally generated density maps.[8,18,19] A common use case is the ribosome, which is abundantly found in tomograms and can often be identified by eye due to its size. However, even for a particle as large as the ribosome, template matching has suboptimal precision.[20,21] For smaller or less abundant macromolecules the method often fares even worse and requires manual curation. These points raise the question of what the requirements and limitations are to reliably template match macromolecules.

Although tomograms are usually collected at $\leq \frac{2\,\text{Å}}{voxel}$, they are typically binned 4-8 times to a coarse voxel size for subsequent analysis to improve computational efficiency and signal-to-noise ratio.[8,14,15,21,22] However, binning removes high-frequency information from the tomogram, and makes it difficult to distinguish macromolecules if the differences in the low-frequency are not sufficiently large.[9] Therefore, template matching under such settings is prone to have low precision. The issue of low precision has been hinted at previously and one suggested solution is template matching in 2D[23,24].

Here, we investigate the pitfalls of 3D template matching with commonly used 4-8 times binned tomograms and rationalize the observed issues. First, we show in theory and practice that in binned cryo-ET data macro-

[a)] Electronic mail: valentin.maurer@embl-hamburg.de
jan.kosinski@embl.de

molecules can be detected by template matching with similar success by virtually any template with a similar overall shape as the target because the low-frequency components dominate and result in nearly indistinguishable cross-correlation scores across templates. Then we analyze in practice the precision of detecting ribosomes with a detailed electron density map, a sphere, and an emoji (💩) of varying radii. Finally, we discuss the implications of these results for the development and benchmarking of template-matching algorithms as well as requirements for data acquisition and processing.

II. THEORY

Template matching, as considered in this work, uses the cross-correlation function to determine the similarity between a target f and a template g at a given translation n:

$$(f * g)[n] = \sum_{m=-\infty}^{\infty} f[m]g[n+m] \tag{1}$$

where $*$ is the convolution operator. Cross-correlation is the sum of the element-wise product of the template and the target, subject to implementation-specific normalization procedures. In practice, this procedure is repeated for a set of rotations of the template.

The computational complexity of the cross-correlation operation on two identical cubes with edge length N is $\mathcal{O}(N^6)$, but in practice, template matching tools reduce the complexity to $\mathcal{O}(N^3 \log(N^3))$. This is achieved by expressing the cross-correlation in the spatial domain as multiplication in the Fourier domain through the cross-correlation theorem:

$$(f * g)(t) = \mathcal{F}^{-1}(\mathcal{F}(f) \cdot \mathcal{F}(g)^*) \tag{2}$$

where $\mathcal{F}$ and $\mathcal{F}^{-1}$ denote the forward and inverse Fourier transform and * the complex conjugate. To build some intuition, lets consider the case where g differs from f only by a translation n:

$$\begin{aligned} \mathcal{F}(g(t-n)) &= \int_{-\infty}^{\infty} g(t-n)e^{-i2\pi k\cdot t}dt \\ &= \int_{-\infty}^{\infty} g(t')e^{-i2\pi k\cdot(t'+n)}dt' \\ &= F(g)e^{-i2\pi k\cdot n} \end{aligned} \tag{3}$$

where k is the wave number in Fourier and t the position vector in the real domain. From this it becomes apparent that a shift in the spatial domain corresponds to frequency-dependent phase shift in the Fourier domain. Furthermore, we note that since:

$$\begin{aligned} e^{-i2\pi k\cdot n} &= \cos(2\pi kn) - i\sin(2\pi kn) \\ |e^{-i2\pi k\cdot n}| &= \sqrt{\cos^2(2\pi kn) + \sin^2(2\pi kn)} \\ &= 1 \end{aligned} \tag{4}$$

the magnitude of the Fourier transform is independent of the phase shift. Therefore, we can determine the cross-correlation in the real domain by inverse Fourier transform of the element-wise product of amplitudes and the sum of phases of $\mathcal{F}(f)$ and $\mathcal{F}(g)^*$. The magnitude of the cross-correlation score depends on the element-wise product of amplitudes, while the sum of phases contains the mapping between translation and score, which is maximal at translation n.

Above we considered the ideal case where the template is a shifted version of the target. In practice this rarely holds and the template rather approximates the amplitude and shifted phase spectrum of the target sufficiently well. Therefore, previous research has seen use of geometric shapes for template matching, such as spheres for localization of ribosomes or RuBisCO[14], cylinders for nucleosomes[15] or proteasomes[16], and rectangles for membranes[17]. Intuitively geometric shapes can be used for template matching if they approximate the structure of interest sufficiently well in the given data. However, why that is the case has not been shown explicitly. We aim to do so in the following and start by deriving the Fourier transforms of the aforementioned geometric shapes.

A sphere of radius R centered around the origin can be defined in real space as:

$$f(r) = \begin{cases} 1, & \text{if } r \leq R \\ 0, & \text{otherwise} \end{cases} \tag{5}$$

here r represents the magnitude of the position vector, i.e., the Euclidean distance from the origin. All points with a distance less than or equal to R are occupied by the sphere. Since the sphere is real symmetric function, its Fourier transform is also real and follows as[25]:

$$\begin{aligned} \mathcal{F}(k) &= \int_{-\infty}^{\infty} d^3r e^{-i2\pi k\cdot r} f(r) \\ &= \int_0^{2\pi} d\phi \int_{-1}^{1} d\cos\theta \int_0^R drr^2 e^{-i2\pi kr\cos\theta} \\ &= \int_0^R drr^2 \frac{2\pi}{i2\pi kr}\left(e^{-i2\pi kr} - e^{i2\pi kr}\right) \\ &= \frac{2\pi}{\pi k}\int_0^R drr\sin(2\pi kr) \\ &= \frac{2\pi}{(2\pi)^2\pi k}\left(\frac{\sin(2\pi kr)}{k^2} - \frac{2\pi R\cos(2\pi kR)}{k}\right) \\ &= 4\pi R^3\left(\frac{sin(2\pi kr) - 2\pi kR\cos(2\pi kR)}{(2\pi kR)^3}\right) \\ &= 4\pi R^3 \frac{j_1(2\pi kR)}{2\pi kR} \end{aligned} \tag{6}$$

where $j_1(x)$ is the spherical Bessel function of first kind and order defined as:

$$j_1(x) = \frac{\sin(x) - x\cos(x)}{x^2} \tag{7}$$

A one-dimensional rectangle, i.e., a box-function can be defined in real space as:

$$g(r) = \begin{cases} 1, & \text{if } -\frac{w}{2} \leq r \leq \frac{w}{2} \\ 0, & \text{otherwise} \end{cases} \tag{8}$$

where w is the width of the box function. The Fourier transform of the one-dimensional box function g(r) is:

$$\begin{aligned} \mathcal{F}(k) &= \int_{-\infty}^{\infty} g(r) e^{-i2\pi k \cdot r} dr \\ &= \int_{-w/2}^{w/2} e^{-i2\pi k r}\, dr \\ &= \frac{1}{\pi k} \cdot \frac{e^{-\pi i k w} - e^{\pi i k w}}{-2i} \\ &= w \cdot \frac{\sin(\pi k w)}{\pi k w} \\ &= w \cdot \mathrm{sinc}(\pi k w) \end{aligned} \tag{9}$$

The definition of the box-function Fourier transform can be used to synthesize the Fourier transform of three-dimensional rectangles with width a, b, and c as:

$$\mathcal{F}(k) = a \cdot \mathrm{sinc}(\pi a k_x) \cdot b \cdot \mathrm{sinc}(\pi b k_y) \cdot c \cdot \mathrm{sinc}(\pi c k_z) \tag{10}$$

where k_x, k_y and k_z are the wave numbers corresponding to the spatial dimensions x, y, and z, respectively.

The cylinder is essentially a combination of a circle and a box-function and can be defined as follows:

$$f(r, z) = \begin{cases} 1 & \text{for } r \leq R \text{ and } |z| \leq \frac{h}{2}, \\ 0 & \text{otherwise.} \end{cases} \tag{11}$$

where R is the radius of the circle and h the width of the box-function. We can make use of the cylindrical symmetry and the separability of the Fourier transform to derive the closed form of the cylinder Fourier transform as follows:

$$\begin{aligned} \mathcal{F}(k_r, k_z) &= \int_0^{2\pi} \int_0^{R} \int_{-\frac{h}{2}}^{\frac{h}{2}} e^{-i2\pi(k_r r \cos\theta + k_z z)} dr\, r\, dz\, d\theta \\ &= \int_0^{2\pi} \int_0^{R} e^{-i2\pi(k_r r \cos\theta)} dr\, r\, d\theta \int_{-\frac{h}{2}}^{\frac{h}{2}} e^{-i2\pi k_z z} dz \\ &= 2\pi \int_0^{R} j_0(2\pi k_r r) dr\, r \int_{-\frac{h}{2}}^{\frac{h}{2}} e^{-i2\pi k_z z}\, dz \\ &= \frac{R}{k_r} j_1(2\pi k_r R)\ h\, \mathrm{sinc}(\pi k_z h) \end{aligned} \tag{12}$$

The Fourier transforms of the sphere, rectangle or cylinder either contain a Bessel or a sinc function or a combination thereof. From this is becomes apparent that these geometric shape concentrate the majority of their Fourier energy in low-frequency components and dampen with a shape-specific rate toward higher frequencies. Böhm *et al.*[9] already hinted at the fact that low-frequencies are essential for particle identification and discusses the detection limits related to low resolution.

To use geometric shapes in template matching the macromolecule of interest would also need to concentrate the majority of Fourier space energy in low-frequency components in a similar manner to yield a high cross-correlation score. Since low-frequency components generally recapitulate the shape and size of the analyzed object in real space, macromolecules have been template matched by geometric shapes with similar size.[14–17] Tomograms are binned four to eight times prior to template matching to increase computational efficiency and reduce noise. The test data used in this study was binned to $\frac{13.481\,\text{Å}}{voxel}$[21], therefore, no features smaller than ≈27 Å can be represented without artifacts according to the Shannon-Nyquist sampling theorem.[26] 27 Å is in excess of most detailed structural features in a macromolecule and consequently the majority of Fourier space energy is also concentrated in low-frequency components, analogous to the geometric shapes we derived above.

These findings have three important implications for template matching: (i) Since template matching at high binning is primarily about matching object size, macromolecules of similar size to the biomolecule of interest will be identified as false positives; (ii) Small macromolecules would be mainly represented through high-frequencies, which overlaps with noise in the data. This relation makes template matching small macromolecules at high binning near-impossible; (iii) More accurate templates are unlikely to improve template matching performance because high-resolution information cannot be accurately represented at the typically used binning. More accurate templates might help to better capture rotational orientations. However, orientations are typically recomputed in subsequent reconstruction steps anyway.

III. METHODS

To test the above-noted theoretical considerations, template matching was performed using PyTOM[10] on an annotated tomogram (EMPIAR-10988, TS_037, ref. 27) reported by de Teresa-Trueba *et al.*[21]. For this sample tomogram, 1,646 ribosomal and 22 fatty acid synthase (FAS) positions were previously identified using PyTOM[10] and subsequently manually refined by an expert user. The template matching was performed for three different template classes that were provided in varying sizes as shown in Fig. 1. First, a previously reported map of the *S. cervisiae* 80S ribosome (EMD-3228; ref 28) was used as a baseline reference (Fig. 1A) with an approximate radius of ten voxels.[21] The structure was scaled to match radii between 1 and 19 voxels using linear spline interpolation. Second, we used a

sphere with varying radius r $1 \leq r < 20$ and homogenous density (Fig. 1B). The third template was the emoji 💩 from the Apple Color Emoji font. The 2D bitmap was converted to a volume with homogenous density by axial symmetrization sampling 360 equidistant angles and subsequently blurred using a Gaussian filter (scipy.ndimage.gaussian_filter, version 1.11.1) with $\sigma = 1$ (Fig. 1C). The 💩 emoji was scaled from the initial 160 x 160 bitmap to appropriate radii using linear spline interpolation. All templates were placed in the center of a cubical volume with edge length 51 and voxel size 13.481 Å, which corresponds to the voxel size of the used tomogram. As mask, we used a sphere whose radius was two voxels larger than the template radius, which is in good concordance with the 337 Å diameter used by de Teresa-Trueba *et al.*[21]. Wedge correction was performed using 40° angles, which corresponds to the -50° to 50° tilt acquisition angles.[21] Templates and masks were prepared with a custom in-house Python script. PyTom was used to sample translations exhaustively at voxel size resolution. For the spheres, 90° (two angles), and for the other templates, 11° (15,192 angles) as well as 25.25° (980 angles) were sampled, which corresponds to the angle lists angles_90_2, angles_11_15192, and angles_25.25_980 that are shipped with PyTom.[10] The rotational sampling rate of 11° is in excess of what is typically used for template matching in previous work.[10,21,29]. In addition to spherical masks, we also tested cubical masks that contained only identical positive values, i.e. had no masking effect for all templates (Fig. S1). PyTom was run assuming a spherical mask, used a bandpass filter with a low and high-frequency cutoff of 3 and 15, split the tomogram into 3 parts along each axis, and performed no further binning. 4,000 peaks were called on the scores obtained from PyTom using skimage.feature.peak_local_max (version 0.21.0) with a minimal allowed Euclidean distance separating peaks of 10 and a 15 voxel exclusion volume around the boundaries of the tomogram. Subsequently, peaks were ordered by their score in decreasing order. We analyzed precision and recall at decreasing score thresholds up to 4,000 top-scoring peaks.

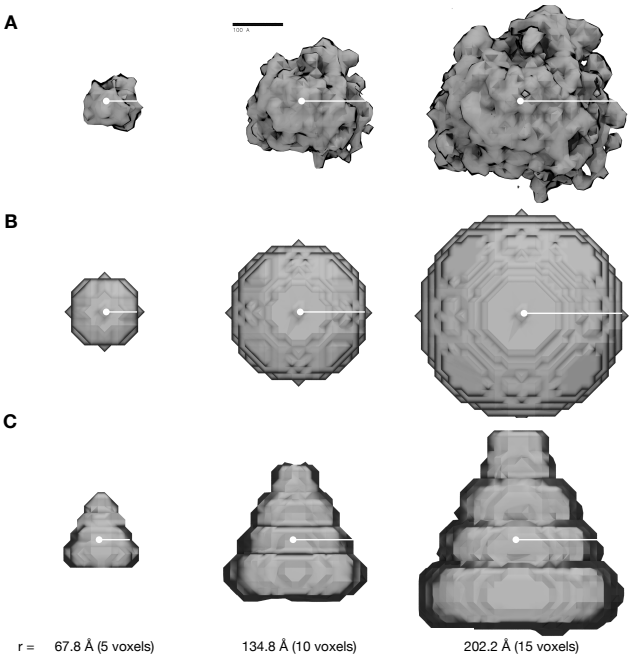

FIG. 1. Template variations used to match the ribosome in the previously annotated tomogram from de Teresa-Trueba *et al.*[21]. Different shapes with different radii sampled at a 13.48 Å voxel size, to match the voxel size of the tomogram, were used as templates for template matching with pyTOM. Specifically, the map of the 80S ribosome (**A**, EMD-3228), a sphere (**B**), and a rotationally symmetrized emoji, the so-called poop emoji (💩), (**C**) were used. The used radii range from 1 to 19 voxels in one voxel increments. Only three representative structures are shown.

IV. RESULTS AND DISCUSSION

To assess our theoretical considerations (sec. II), we used PyTOM[10] to perform template matching with three different templates on a tomogram provided by de Teresa-Trueba *et al.*[21]. The used tomogram contained 1,646 ribosomes and 22 FAS particles which were annotated by the authors using template matching and manual curation. We considered their annotation as ground truth.

As reported by de Teresa-Trueba *et al.*[21], we used a 80S ribosome[28] map as initial template and scaled its radius to see how this affects template matching performance. Based on our theoretical considerations, any template with a similar Fourier spectrum should perform similarly well as the 80S ribosome template, which is why we also tested spheres, and an irregularly shaped emoji (💩) at various scales. Spheres have already been successfully used in practice for identifying RuBisCO[14] and other basic shapes such as cylinders for nucleosomes[15] or the proteasome[16] and rectangles for membranes.[17] We subsequently compared the template matching results for the different templates and scaled radii with respect to common metrics such as precision ($precision = \frac{TP}{TP+FP}$) and recall ($recall = \frac{TP}{TP+FN}$), where TP, FP, and FN correspond to the number of true positives, false positives, and false negatives, respectively. The picked particles were chosen from a sorted list of all scores in descending order. All results are summarized in Fig. 2.

First, we compared the recall between the varying radii and across the different templates with respect to the number of picked particles (Fig 2A). Overall the performance up to 4,000 top-scoring picks across templates was comparable, with a maximum recall of around 55-60%. We conclude that at the voxel size of 13.481 Å used here, a real *C. cervisiae* ribosome map (EMD-3228; ref. 28) performs no better than a sphere or an emoji of similar size. Furthermore, given the shape of the curves, it appears unlikely that the remaining ribosomes could be

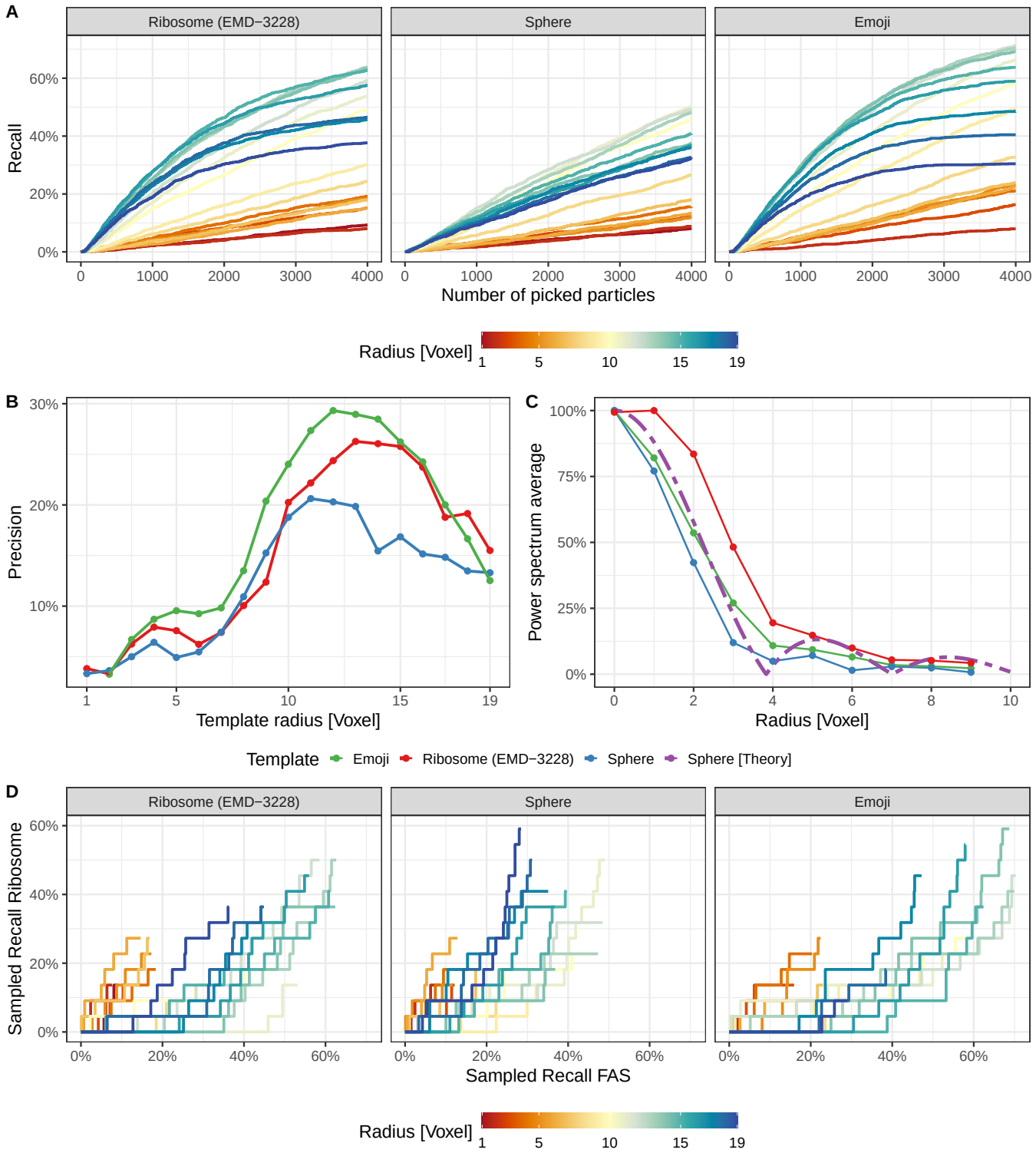

FIG. 2. Template matching performance using three distinct templates (see Fig. 1). **A** Ribosome picking recall by number of picked particles. We used linear sum assignment to achieve an optimal one-to-one mapping between ground truth and picked particles. A particle is considered correctly picked if it is within a 20-voxel distance from its assigned ground truth particle. Consequently, all particles without assignment to ground truth particles were considered false positives. All 4,000 picked particles were considered in the following figures. **B** Proportion of true positives out of all picked particles ("Precision") by radius of templates. **C** Fourier power spectrum averages by radius and template. "Radius" was computed as Euclidean distance from the zero-frequency component and rounded to the nearest integer. Shown are templates used for template matching at a radius of ten (Fig. 1). "Sphere [Theory]" refers to the theoretical derivations made in the theory section with $R = 10$.[25] Power spectrum averages were linearly scaled to the interval $[0, 1]$ to facilitate curve shape comparison. **D** Picked particles were one-to-one assigned to the union of ground truth FAS and ribosome coordinates using linear-sum-assignment. Each particle is assigned to no more than one class and considered to correctly pick that class if it is within a 20-voxel distance from its assigned ground truth particle.

recovered at reasonable precision. Although we picked 4,000 particles, no template at any size could fully recapitulate the ground truth particle set beyond a recall of 60% and a precision of $\approx$ 25%. Therefore, it becomes questionable whether using improved experimental or predicted structures as templates will be sufficient to identify small or lowly abundant proteins in cryo-ET data at low-resolution especially since manual curation was always necessary.[21] In accordance with our theoretical considerations, the experimental results show that more accurate templates most likely are not sufficient to improve template matching performance in highly binned tomograms.

As displayed in Fig. 2B, template matching precision at 4000 picks is maximal around eight and 12 voxels. We compared the recall of template matching at rotational sampling densities of 11° and 25.25° to ensure this did not affect the result and found a negligible difference between the rotational sampling rates (Fig. S1). In both cases, the emoji has maximal precision over scales followed by the ribosome and the sphere template. Templates smaller than a radius of ten voxels performed considerably worse. This is likely due to the presence of noise or additional macromolecules that are smaller than the ribosomes but are composed of comparable density. This is in line with the fact that all current works test template matching with large macromolecules including, but not limited to, ribosomes[10,21,29], proteasomes[8,16], thermosomes[8], RuBisCO[14].

Notably, none of the templates had maximal precision around the anticipated ribosome radius of ten voxels. This can be explained by the behavior of PyTom's default fast local correlation (FLC) score. FLC normalizes cross-correlation scores based on the local variance of the tomogram computed with respect to the provided template mask.[10] Since the score depends on both the template itself and the mask for normalization, the properties of both can influence the overall performance of template matching. The mask should be sufficiently tight, approximate the template, and not overlap with other high-intensity regions. A mask that is too loose or too tight can lead to a higher local variance, which in turn leads to a lower score. This is reflected by a shift of the maximal precision to larger template radii (Fig. 2).

To confirm this we tested both the spherical mask (Fig. 2) and compared it to an excessively large cubical mask (Fig. S1). When using an excessively large cubical mask, the ribosome and emoji template with maximal precision is closed to the expected radius ten voxel radius of the ribosome (Fig. S1 C). Notably, the precision of ribosome and emoji is lower when using a cubical mask, potentially due to crowdedness in the tomogram whereas they perform better when using a spherical mask. This indicates further intricacies which should be explored in the future.

To compare the tested templates with the theoretical results, we computed the radially averaged Fourier power spectrum for three templates at radius of ten voxels and compared them to the theoretical curve of a sphere (Fig. 2C). All templates concentrate their Fourier domain energy in low-frequency components. This matches the theoretical assumption that template matching using geometric shapes is possible if the majority of the Fourier space energy is concentrated in a similar manner.

Finally, we assessed whether using the three templates at different scales, we could pick the FAS protein complex, whose shape differs substantially from the ribosome's shape but has a similar size. Although the number of annotated FAS in the particular tomogram is only 22, FAS particles are among the 4,000 highest-scoring picks across templates and scales (Fig. 2D). Although the low number of annotated FAS impedes quantitative claims, the trends are clear. For a sphere as many as 60% of FAS are recovered and even for the ribosome as a template more than 60% of annotated FAS instances are recovered. This finding further highlights that molecular details play a minor part in template matching at high binning. In general, this finding indicates that proteins that are low in abundance cannot be practically identified with sufficient precision if many other macromolecules of similar size exist. In this example with the 13.481 Å pixel size used here, FAS and ribosomes are similarly sized, resulting in fairly similar scores and thus poor differentiation between the two.

V. CONCLUSIONS AND OUTLOOK

In this article, we discussed using theoretical considerations and practical examples why simple geometric shapes such as spheres, cylinders, or rectangles of appropriate size can be used to find macromolecules equally well as detailed structural models. Based on our theoretical considerations this is because highly binned tomograms primarily consist of low-frequency information. Thus, the internal structure has less influence and the shape and size of the template dominate the cross-correlation, as seen for the reported simple geometric shapes. In practice, we showed this by using a subtomogram averaged reference, a sphere, and the so-called poop emoji (💩) as a template to identify ribosomes in an annotated *S. cerevisiae* dataset[21] resulted in similar performance. As expected from our theory, macromolecules of similar size will often be identified as false positives over the macromolecule of interest. Importantly, this issue will be more pronounced for small molecules where the high frequencies will overlap with noise, and template quality will not improve the performance either.

Based on these considerations we suggest the following moving forward. First, cross-correlation based scoring methods appear to be a suboptimal measure of similarity in tomograms. This is particularly apparent for high binnings. Therefore, different perhaps non-linear similarity metrics such as those used in machine learning can enhance template matching performance.[21,30,31] However, for small macromolecules generating adequate training data sets could be highly challenging, as man-

ual curation would be limited by noise levels and visibility. Secondly, analyzing lower binned tomograms can potentially improve cross-correlation based template matching. However, this has only been performed in 2D[23,24] and its extension to 3D would require solving additional challenges such as noise, specimen motion resulting from problems with tilt alignment and sample deformation, errors in CTF correction, and issues with computational feasibility.[23,32] Lastly, we suggest broadening benchmark entities beyond large, and highly abundant globular structures like the ribosome when evaluating new template-matching algorithms.

CONFLICT OF INTEREST

The authors declare no conflicts of interest.

AUTHOR CONTRIBUTIONS

VM and MS conceived the study, performed research, wrote the initial draft, and edited the manuscript. JK provided feedback and edited the manuscript.

ACKNOWLEDGMENTS

We thank Lukas Grunwald (Max Planck Institute for the Structure and Dynamics of Matter, Center for Free-Electron Laser Science (CFEL), Hamburg, Germany) for critical reading and editing of the theory section and Julia Mahamid (EMBL Heidelberg, Germany) for critical reading and feedback of the manuscript. We thank the EMBL IT and HPC resources for providing essential computational infrastructure. VM and JK acknowledge funding from the CSSB flagship project Plasmofraction. MS acknowledges support from a research fellowship from the EMBL Interdisciplinary Postdoc (EIPOD) Programme under Marie Curie Cofund Actions MSCA-COFUND-FP (grant agreement number: 847543).

REFERENCES

[1] E. Pyle and G. Zanetti, Biochem. J. **478**, 1827 (2021).
[2] N. Volkmann, in *Methods Enzymol.*, Vol. 483 (Elsevier Inc., 2010) 1st ed., pp. 31–46.
[3] J. Mahamid, S. Pfeffer, M. Schaffer, E. Villa, R. Danev, L. Kuhn Cuellar, F. Forster, A. A. Hyman, J. M. Plitzko, and W. Baumeister, Science (80-.). **351**, 969 (2016).
[4] S. Pfeffer, J. Dudek, M. Schaffer, B. G. Ng, S. Albert, J. M. Plitzko, W. Baumeister, R. Zimmermann, H. H. Freeze, B. D. Engel, and F. Förster, Nat. Commun. **8**, 14516 (2017).
[5] F. Wilfling, C.-W. Lee, P. S. Erdmann, Y. Zheng, D. Sherpa, S. Jentsch, B. Pfander, B. A. Schulman, and W. Baumeister, Mol. Cell **80**, 764 (2020).
[6] V. Lučić, A. Rigort, and W. Baumeister, J. Cell Biol. **202**, 407 (2013).
[7] X. Wu, X. Zeng, Z. Zhu, X. Gao, and M. Xu, in *Comput. Biol.* (Codon Publications, 2019) pp. 175–186.
[8] A. S. Frangakis, J. Böhm, F. Förster, S. Nickell, D. Nicastro, D. Typke, R. Hegerl, and W. Baumeister, Proc. Natl. Acad. Sci. U. S. A. (2002), 10.1073/pnas.172520299.
[9] J. Böhm, A. S. Frangakis, R. Hegerl, S. Nickell, D. Typke, and W. Baumeister, Proc. Natl. Acad. Sci. U. S. A. **97**, 14245 (2000).
[10] T. Hrabe, Y. Chen, S. Pfeffer, L. Kuhn Cuellar, A.-V. Mangold, and F. Förster, J. Struct. Biol. **178**, 177 (2012).
[11] W. Wan, S. Khavnekar, J. Wagner, P. Erdmann, and W. Baumeister, Microsc. Microanal. **26**, 2516 (2020).
[12] G. Tang, L. Peng, P. R. Baldwin, D. S. Mann, W. Jiang, I. Rees, and S. J. Ludtke, J. Struct. Biol. **157**, 38 (2007).
[13] D. Castaño-Díez, M. Kudryashev, M. Arheit, and H. Stahlberg, J. Struct. Biol. **178**, 139 (2012).
[14] B. D. Engel, M. Schaffer, L. Kuhn Cuellar, E. Villa, J. M. Plitzko, and W. Baumeister, Elife **4**, 1 (2015).
[15] S. Cai, C. Chen, Z. Y. Tan, Y. Huang, J. Shi, and L. Gan, Proc. Natl. Acad. Sci. **115**, 10977 (2018).
[16] S. Nickell, O. Mihalache, F. Beck, R. Hegerl, A. Korinek, and W. Baumeister, Biochem. Biophys. Res. Commun. **353**, 115 (2007).
[17] M. N. Lebbink, W. J. Geerts, T. P. van der Krift, M. Bouwhuis, L. O. Hertzberger, A. J. Verkleij, and A. J. Koster, J. Struct. Biol. **158**, 327 (2007).
[18] M. Beck, J. A. Malmström, V. Lange, A. Schmidt, E. W. Deutsch, and R. Aebersold, Nat. Methods **6**, 817 (2009).
[19] S. Kühner, V. van Noort, M. J. Betts, A. Leo-Macias, C. Batisse, M. Rode, T. Yamada, T. Maier, S. Bader, P. Beltran-Alvarez, D. Castaño-Diez, W.-H. Chen, D. Devos, M. Güell, T. Norambuena, I. Racke, V. Rybin, A. Schmidt, E. Yus, R. Aebersold, R. Herrmann, B. Böttcher, A. S. Frangakis, R. B. Russell, L. Serrano, P. Bork, and A.-C. Gavin, Science (80-.). **326**, 1235 (2009).
[20] K. Zhang, B. Lucas, and N. Grigorieff, Microsc. Microanal. **29**, 931 (2023).
[21] I. de Teresa-Trueba, S. K. Goetz, A. Mattausch, F. Stojanovska, C. E. Zimmerli, M. Toro-Nahuelpan, D. W. C. Cheng, F. Tollervey, C. Pape, M. Beck, A. Diz-Muñoz, A. Kreshuk, J. Mahamid, and J. B. Zaugg, Nat. Methods **20**, 284 (2023).
[22] L. Xue, S. Lenz, M. Zimmermann-Kogadeeva, D. Tegunov, P. Cramer, P. Bork, J. Rappsilber, and J. Mahamid, Nature **610**, 205 (2022).
[23] B. A. Lucas, B. A. Himes, L. Xue, T. Grant, J. Mahamid, and N. Grigorieff, Elife **10**, 1 (2021).
[24] J. P. Rickgauer, N. Grigorieff, and W. Denk, Elife **6**, 1 (2017).
[25] J. Friedman, Protein engineering **10**, 851 (1997).
[26] C. Shannon, Proceedings of the IRE **37**, 10 (1949).
[27] A. Iudin, P. K. Korir, S. Somasundharam, S. Weyand, C. Cattavitello, N. Fonseca, O. Salih, G. J. Kleywegt, and A. Patwardhan, Nucleic Acids Research **51**, D1503 (2022), https://academic.oup.com/nar/article-pdf/51/D1/D1503/49645431/gkac1062.pdf.
[28] T. A. M. Bharat and S. H. W. Scheres, Nat. Protoc. **11**, 2054 (2016).
[29] S. Pfeffer, F. Brandt, T. Hrabe, S. Lang, M. Eibauer, R. Zimmermann, and F. Förster, Structure **20**, 1508 (2012).
[30] E. Moebel, A. Martinez-Sanchez, L. Lamm, R. D. Righetto, W. Wietrzynski, S. Albert, D. Larivière, E. Fourmentin, S. Pfeffer, J. Ortiz, W. Baumeister, T. Peng, B. D. Engel, and C. Kervrann, Nat. Methods **18**, 1386 (2021).
[31] G. Rice, T. Wagner, M. Stabrin, O. Sitsel, D. Prumbaum, and S. Raunser, Nat. Methods **20**, 871 (2023).
[32] L. M. Voortman, M. Vulović, M. Maletta, A. Voigt, E. M. Franken, A. Simonetti, P. J. Peters, L. J. van Vliet, and B. Rieger, J. Struct. Biol. **187**, 103 (2014).

SUPPORTING INFORMATION

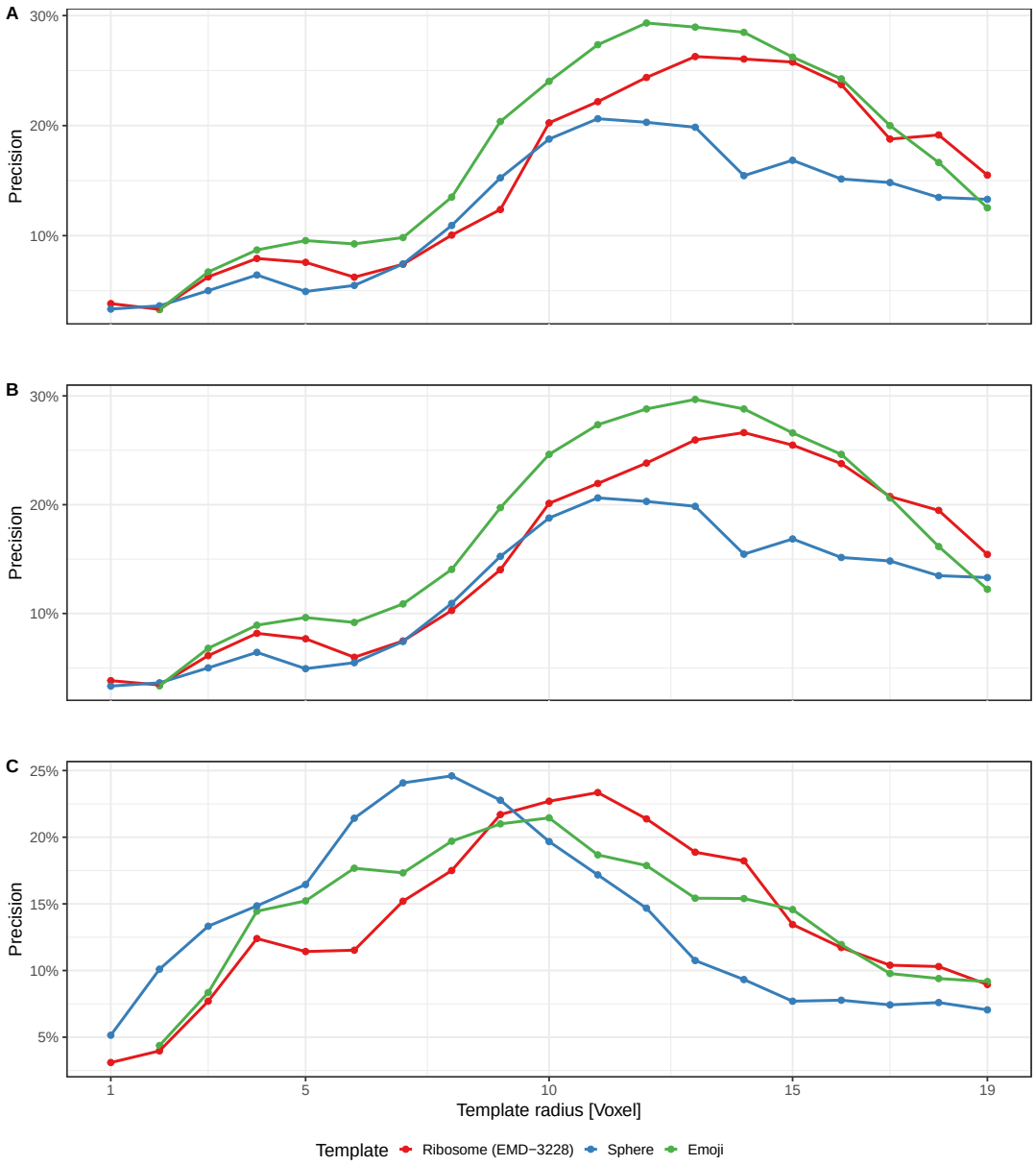

FIG. S1. Proportion of true positives out of all picked particles ("Precision") by radius of templates. Due to the symmetry of the sphere template, only two angles were used which is the smallest angle list shipped with PyTom. **A** Using 15,192 angles and spherical masks, **B** using 980 angles and spherical masks, and **C** using 980 angles and cubical masks.