Reviewed Preprint
v1 • January 19, 2026
Not revised

Reviewed Preprint
v2 • June 17, 2026
Revised by authors

For correspondence:
kexin.zhang1@umassmed.edu
niko@grigorieff.org

Competing interests: No competing interests declared

Funding: See page 32

Reviewing editor: Sjors HW Scheres, MRC Laboratory of Molecular Biology, United Kingdom

Improved cryo-EM reconstruction of sub-50 kDa complexes using 2D template matching

Kexin Zhang[a,b], **Timothy Grant**[c,d], **Nikolaus Grigorieff**[a,b]

[a]RNA Therapeutics Institute, University of Massachusetts Chan Medical School, Worcester, United States • [b]Howard Hughes Medical Institute, Worcester, United States • [c]Department of Biochemistry, University of Wisconsin-Madison, Madison, United States • [d]Morgridge Institute for Research, Madison, United States

eLife Assessment

This **important** study builds on previous work from the same authors to present a conceptually distinct workflow for cryo-EM reconstruction that uses 2D template matching to enable high-resolution structure determination of small (sub-50 kDa) protein targets. The paper describes how density for small-molecule ligands bound to such targets can be reconstructed without these ligands being present in the template. However, the evidence described for the claim that this technique improves the alignment of the reconstruction of small complexes compared to standard techniques is **incomplete**. The authors could better evaluate the effects of model bias on the reconstructed densities, as suggested by reviewer #1.

https://doi.org/10.7554/eLife.109790.2.sa3

Abstract

Visualizing the structures of small proteins and complexes has been a longstanding challenge in single-particle cryo-EM. Some of these targets have been successfully resolved by binding to antibody fragments (Fabs) or fusing with external scaffolds to increase their size. Recent advances in conventional single-particle techniques have enabled the determination of an increasing number of structures smaller than 100 kDa, achieving resolutions relevant to drug research. Compared to X-ray crystallography, cryo-EM preserves the near-native states of biomolecules, can resolve structural heterogeneity, and has the potential to apply to a wide range of targets. In this work, we demonstrate that the alignment and reconstruction of small macromolecular complexes can be improved using high-resolution structures as priors combined with 2D template matching. Using this method, we reconstructed a previously intractable ~ 43 kDa protein kinase and improved the density of its ligand-binding site. Our theoretical analysis predicts that this method can further extend single-particle cryo-EM to important drug-binding complexes well below 50 kDa.

Introduction

Since the “resolution revolution” in 2014 (1), single-particle cryo-electron microscopy (cryo-EM) has become widely used, emerging as a powerful alternative, and in some cases surpassing X-ray crystallography for structurally challenging samples. In particular, the 1.2 Å reconstruction of apoferritin, where individual hydrogen atoms are resolved, demonstrates that single-particle cryo-EM has now reached true atomic resolution for well-defined systems (2, 3).

Despite recent advances in single-particle cryo-EM, structure determination for sub-50 kDa complexes remains challenging. Among EM structures deposited in the Protein Data Bank (PDB) since 2015 with resolutions better than 4 Å, 97% correspond to macromolecules larger than 50 kDa (4). Small complexes contain fewer atoms and scatter fewer electrons, leading to images with

lower signal-to-noise ratios (SNRs). Particle alignment, typically based on calculating the cross-correlation coefficients between the particle and reference (5), depends on the amount of phase contrast (signal) in the image. The weak contrast of small complexes against background noise makes it difficult to align particles accurately to calculate a high-resolution reconstruction. Moreover, single-particle experiments typically apply a total exposure of 30-50 electrons per $Å^2$, leading to accumulating radiation damage. The diffusion of small particles due to beam-induced motion may further reduce the alignable high-resolution signal (6).

Extending single-particle cryo-EM to sub-50 kDa targets will open new avenues for studying critical drug-binding interactions and advance structure-based drug discovery. Approximately 75% of proteins in the human proteome are below 50 kDa (7). While X-ray crystallography is not inherently limited by small molecular weight, it requires well-ordered crystals, which are often difficult to obtain. Solution and solid-state NMR spectroscopy, on the other hand, provide atomic-level information derived from observables such as chemical shifts. However, these methods are generally limited to smaller particles, require highly concentrated samples, and data analysis is laborious, making them less suitable for high-throughput studies (8–10). Single-particle cryo-EM offers unique benefits compared to these two techniques, producing rich structural data in the form of images and enabling the visualization of biomolecules in near-native states at high resolution without the need for crystallization. To overcome the lower size barrier in cryo-EM, many strategies focus on increasing the apparent size and rigidity of small particles by binding them to external scaffolds or antibodies (11–14). However, finding good scaffolds or antibodies is difficult, and they sometimes create structural artifacts (15). Thus, there remains an incentive to develop approaches that enable high-resolution structure determination of isolated sub-50 kDa complexes in their near-native conformations.

Theoretical estimates of the lower molecular weight limit, dating back to 1995, suggested that a 3D reconstruction at 3 Å resolution is possible for particles as small as 38 kDa, given that ~ 12,600 images are averaged (16). This calculation assumed perfect images, a perfect reference to align the particle images, and a total exposure limited to 5 electrons $Å^{-2}$. Later, using the Rose criterion, a more optimistic prediction of 17 kDa was calculated, requiring only one-ninth as many images to be averaged (17). Since those early calculations, single-particle cryo-EM has seen significant technical advancements. The introduction of direct electron detectors (DEDs), combined with exposure weighting, now allows much higher electron exposures by recording data as movies and improving image resolution through motion correction (18, 19). More recently, the development of a laser phase plate offers the potential to further improve image contrast by using a high-intensity laser beam to introduce a stable and tunable phase shift to enhance low-resolution features in samples with weak contrast such as small complexes (20–23). Additionally, cooling specimens at liquid-helium temperatures can delay radiation damage and reduce information loss during imaging (24). Recent work has shown that using gold specimen support with 100 nm diameter holes at liquid-helium temperatures allowed imaging with better quality compared to liquid-nitrogen temperatures (25).

2D template matching (2DTM) was previously developed to identify particles in cellular cryo-EM images using a high-resolution template with high accuracy (26–28). In a 2DTM search, cross-correlation coefficients are calculated between the image and the template to determine the targets' location and orientation. Previous work from our lab has shown that molecular features not in the template can be reconstructed from 2DTM-derived targets with high resolution (29). Building upon this, in this paper, we show that the alignment of sub-50 kDa complexes can be improved using 2DTM and stringent particle selection. We apply this method to a previously published dataset of the ~43 kDa catalytic domain of protein kinase A (30) and demonstrate improved reconstruction of its ligand-binding sites. We develop a theoretical framework showing that the lower molecular weight limit can be further reduced to approximately 7.1 kDa with an exhaustive five-dimensional search, and to 5.7 kDa with a constrained search, assuming phase plate and liquid-helium cooling are used. Our findings highlight the potential of 2DTM to expand

the applicability of cryo-EM to a broader range of biologically and pharmaceutically important targets, paving the way for structural studies of small complexes that remain difficult for standard workflows.

Results

Unbiased reconstruction of omitted densities in a 43 kDa protein kinase

We evaluated the ability of 2DTM to recover ligand densities without model bias by performing a template matching search and reconstruction with specific features omitted from the template. We used the published dataset of a 42.8 kDa protein kinase (EMPIAR-10252) (30). Particle stacks were generated from single-particle data using coordinates output from 2DTM searches with templates missing certain structural components. A subset of particles was then selected based on 2DTM statistics and image quality measurement from fits of the contrast transfer function (CTF) of the micrographs these particles came from. The selected particle stack was used for 3D reconstruction. This "omit-template" strategy follows the *baited reconstruction* approach (29) and was designed to detect template bias in reconstructions calculated from targets detected by 2DTM. Here, we use it to show that we can reconstruct the density of small ligands that are not included in the template but are bound to the imaged molecules in the cryo-EM dataset.

We explored a range of template deletion scenarios: from omitting ATP and residues on an alpha-helix, to removing the binding pockets around ATP at various radii. Below, we detail how each deletion strategy affected the final density map. Consistently, we found that ATP density was robustly recovered even when ATP was deleted from the 2DTM search template. These results demonstrate that 2DTM can recover ligand densities, providing biological insights into ligand binding and flexibility. The density is free of template bias where corresponding density in the template was omitted (29).

Deleting ATP, inhibitor, and an alpha-helical turn

We used the 2.2 Å X-ray model (PDBID: 1ATP, (31)) to generate a high-resolution 3D template in which the ATP, two Mn^{2+} ions, and six alpha-helix residues (residues 222–227) were deleted from the model. The final template, with a molecular weight of 38.0 kDa (non-hydrogen atoms only), was simulated using the `simulate` program in *cis*TEM with a uniform B-factor of 30 $Å^2$ and a pixel size of 1.117 Å/pixel (32). Using the previously developed 2DTM p-value for particle picking and angular alignment (33), we performed 2DTM searches at Nyquist resolution and generated an initial stack of candidate particles, which were then subjected to further selection using 2DTM-derived statistics and CTF fitting parameters. A final stack of 7,353 particles was used for 3D reconstruction in *cis*TEM. Figure 1(a) compares the cryo-EM maps from the single-particle reconstruction (30), the 2DTM template, and the 2DTM-based reconstruction. The angular distribution of the 2DTM-derived particle stack and the Fourier Shell Correlation (FSC) curve are shown in Figure 1b. Because particles were aligned to a single high-resolution template during 2DTM, the reported 2.6 Å resolution is **not a gold-standard** FSC and does not reliably reflect true map quality. Figure 1c provides a close-up view of the ATP-binding site and the deleted residues, with the X-ray model overlaid for comparison. The average Q-score between the 2DTM reconstruction and the X-ray model was calculated using the `MapQ` command line tool (34, 35). Q-scores approaching 1 suggest atomic resolution, where individual atoms are resolved. Values near 0.5 indicate visible side chains while scores around 0.2 reflect unresolved side chains but resolved secondary structure (35). The Q-score for ATP was 0.60, indicating good agreement between the 2DTM reconstruction and the X-ray model, while the two Mn^{2+} ions had Q-scores of 0.48 and 0.73. The Q-scores of deleted residues 222–227 (Trp, Ala, Leu, Gly, Val, Leu) were 0.63, 0.51, 0.57, 0.54, 0.61, and 0.53, respectively. Additionally, the shape of the alpha-helix turn is clearly resolved, despite being absent in the template. Since these features were not included in the search template, they did not originate from template bias. To further validate that the recovered density is genuine, we performed Phenix real-space grouped occupancy refinement of the full 1ATP model

against the omit reconstruction. Omitted residues 222–227 refined to occupancies of 0.55–0.80 (mean 0.72) and ATP to 0.61, while template-included control residues 150–155 remained near 1.0 (mean 0.96), confirming partial, unbiased recovery of omitted-region density (Figure 1—source data 1).

Robust reconstruction of the ATP binding pocket

To test the robustness of 3D reconstruction as more atoms are excluded from the template, we performed a series of 2DTM searches and reconstructions using templates with residues deleted within specified radii of the bound ATP. In Figure 2, we show the results of three different templates. We tested spherical deletion radii of 3.0 Å (Figure 2a and b) and 5.5 Å (Figure 2c) centered on the ATP position in the X-ray model. In addition to the ATP binding pocket, we also deleted Mn^{2+} in all templates and furthermore deleted IP20, an inhibitory pseudo-peptide substrate, in Figure 2b. The molecular weights of templates and numbers of particles in the final stacks are shown in the figure.

In all three cases, the reconstructions showed well-defined ATP and Mn^{2+} densities in the binding pocket. The average Q-scores of ATP in three cases were 0.61, 0.57, and 0.61, respectively. Although some discontinuity was observed at the chosen contour level (σ = 5) in Figure 2b, the ATP densities in all reconstructions closely matched the ligand shape in the X-ray model. These results suggest that ATP and Mn^{2+} densities can be robustly recovered from the data, even when the search template lacks these ligands and nearby residues. We also found that nearby protein residues were more affected by template deletion than ATP. When residues within 3 Å of ATP were deleted (Figure 2a), most omitted residues were still recovered with backbone Q-scores above 0.50, except for residues 53 (Ser) and 127 (Glu), which had backbone Q-scores of 0.32 and 0.42, respectively (red circles in Figure 2a). However, when the deletion radius increased to 5.5 Å (Figure 2c), many more omitted residues fell below 0.50, and the densities corresponding to those residues began to show visible discontinuities. This likely was caused by the overall reduction of signal in the template that contributes to the cross-correlation signal. In contrast, ATP density was consistently recovered more robustly than nearby residues. This may be because small angular misalignments disproportionately blur peripheral residues as their displacement scales with distance from the alignment center, whereas the centrally buried ATP density remains relatively unaffected.

Finally, by comparing Figure 2a and b, we observed that deleting IP20 strongly reduced signal at several residues: Phe54, Gly55, Lys72, Glu127, Glu170, and Asp184 all dropped below a backbone Q-score of 0.5 when IP20 was additionally removed, compared with only Ser53 and Glu127 when IP20 was kept. IP20 not only contributes to the overall mass (2.1 kDa) of the template but also sits at the edge of the protein, where it may generate distinct low-resolution features that can facilitate alignment.

Altogether, our experiments demonstrate that it is safe to delete a ligand and residues in the nearby regions to avoid template bias without destroying too much signal in the reconstruction.

2DTM provides more accurate alignments than RELION refinement in omitted regions

To evaluate the quality of particles and poses obtained from 2DTM, we imported the stack of 7,353 particles described in Figure 1 into RELION and performed 3D classification without angular refinement using five classes. As shown in Figure 3a, we combined Class 1-4 into a new stack of 7,197 particles and performed both 3D reconstruction without angular refinement (using the 2DTM-derived orientations directly) and regular (alignment-enabled) 3D auto-refinement in RELION. Each resulting map was then post-processed and low-pass filtered for comparison. As shown in Figure 3b, in both maps, densities for the deleted residues were recovered. Specifically, the densities at ATP and residues 222–227 obtained with the directly imported 2DTM Euler angles were sharper and more continuous than the same region produced with RELION angular refinement. The reconstruction using 2DTM orientations reached 3.1 Å at the FSC=0.143 threshold, compared with 3.7 Å for the RELION auto-refined reconstruction (Figure 3b). We note that neither FSC is a true gold-standard estimate, as both use orientations ultimately derived from

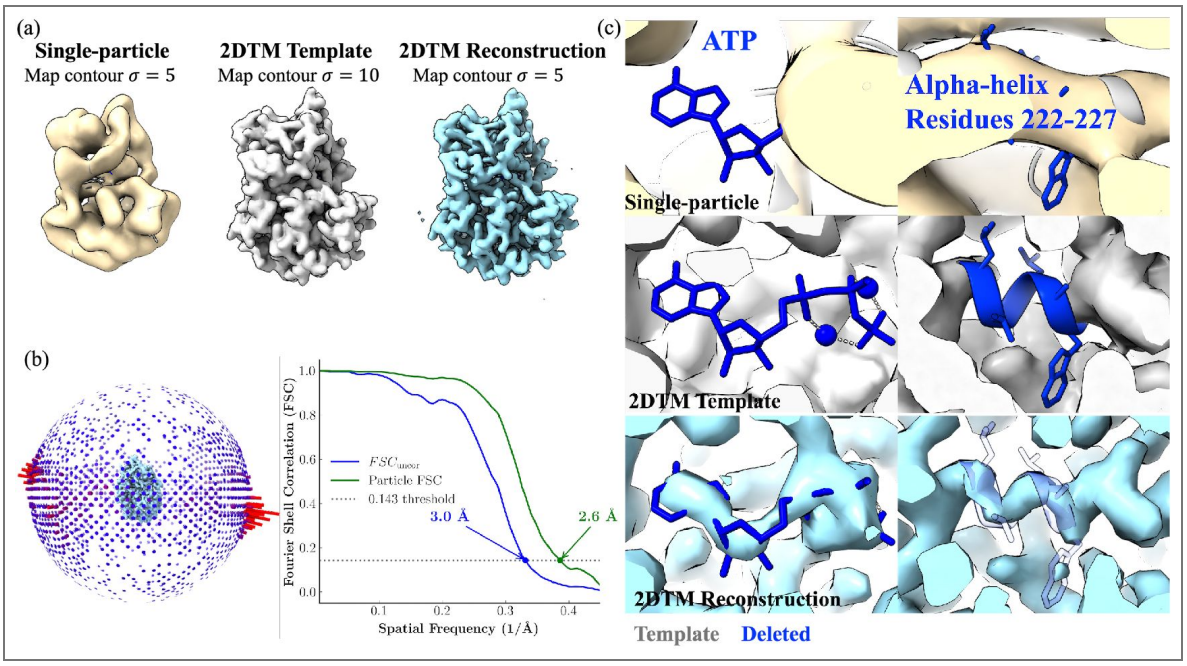


Figure 1. 3D reconstruction of protein kinase using 2DTM-derived particle stack.

(a) From left to right: the single-particle reconstruction (EMDB:0409) (30), the 2DTM template, and the 2DTM-derived reconstruction of the protein kinase. **(b)** Angular distribution plot and FSC curve calculated using *cis*TEM. Note that these FSC curves are not gold-standard FSCs, as the reconstruction uses orientations determined by 2D template matching rather than independent half-set refinement. FSC_{uncor} denotes the uncorrected FSC computed within a generous mask, which crosses the 0.143 threshold at 3.0 Å, and "Particle FSC" denotes the solvent-corrected FSC obtained using the mask-volume correction factor as described in the *cis*TEM/Frealign framework (63). **(c)** Densities at the ATP-binding site and deleted residues 222–227. The X-ray structural model (PDBID: 1ATP) (31) is shown with atoms retained in the template colored grey and deleted atoms colored blue. The alpha-helical model for residues 222–227 is rendered transparent to better show the recovery of the helical turn in the reconstruction. The average Q-scores between the 2DTM reconstruction and the X-ray model were calculated using the MapQ command line tool (34, 35). Q-scores for ATP, Mn^{2+}, and deleted residues 222–227 were 0.60, 0.48/0.73, 0.63, 0.51, 0.57, 0.54, 0.61, and 0.53, respectively.

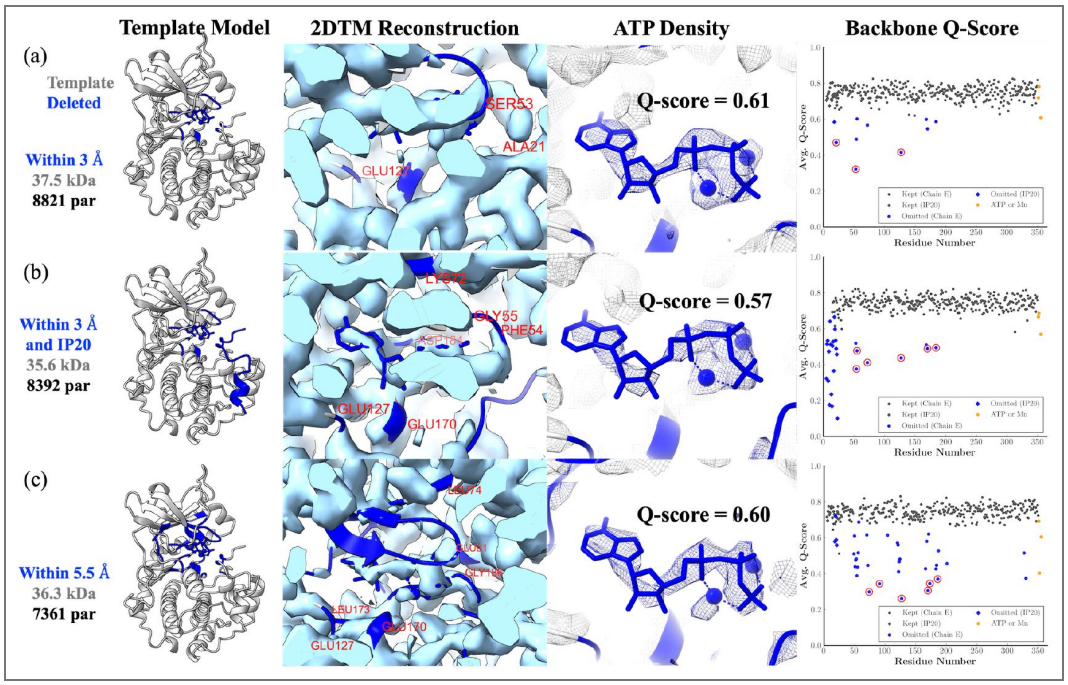


Figure 2. 2DTM-based reconstruction of the ATP-binding pocket.

Each row shows, from left to right: (1) the atomic model used to generate the template, with residues within a specified radius from ATP deleted (highlighted in blue); (2) the 2DTM reconstruction fitted with the full model, with labeled residues corresponding to those circled in red in the Q-score plot; (3) the recovered ATP density fitted with the full model; and (4) backbone Q-scores of all residues. Omitted residues are shown in blue (circles for chain E, diamonds for IP20); kept residues in gray. ATP and Mn^{2+} are shown in orange. Red circles indicate the omitted residues with the lowest backbone Q-scores, corresponding to the labeled residues in column 2. Q-scores were calculated using the MapQ command line tool (34, 35). **(a)** Residues within 3 Å of ATP were deleted. **(b)** IP20 and residues within 3 Å of ATP were deleted. **(c)** Residues within 5.5 Å of ATP were deleted. Map contour $\sigma = 5$.

the template; however, the comparison between the two processing strategies remains informative. These results suggest that the 2DTM-derived orientations are more accurate than those determined by RELION. To test whether RELION's performance was limited by its initial low-pass filter, we repeated the auto-refinement with different initial low-pass filter values (Figure 3c). The final resolution remained between 3.7 and 4.0 Å, indicating that RELION's gold-standard regularization, not the starting reference resolution, is the limiting factor. We also confirmed that including all five classes yielded 3.7 Å, identical to the reconstruction with Class 5 removed (Figure 3c). A more quantitative assessment of the alignment accuracies attained by 2DTM and RELION requires further work.

Effective selection strategy enables reconstruction of ATP binding site from around 8k particles

An important difference between our 2DTM approach and the traditional single-particle workflow is the way particles were included in the final 3D reconstruction. In the original publication (30), a final stack of 74,413 particles was used to obtain a reconstruction at ~ 4.3 Å resolution by gold-standard FSC. The resulting density, however, lacked the expected features at 4 Å, but rather appeared more consistent with a ~ 6-7 Å map. In particular, the ATP-binding pocket was not resolved in the map (Figure 1a). This particle stack included both untilted and 30°-tilted data to reduce preferred orientation. The paper noted that multiple optimizations were attempted, including iterative particle selection based on RELION metadata metrics, but none led to significant improvement.

In our analysis, we implemented a different strategy focused on more stringent particle selection. At the image level, CTF fitting scores from untilted images were computed using CTFFIND5 (36, 37), and only images with scores between 0.05 and 0.2, corresponding to well-fit CTFs, were retained for subsequent 2DTM searches (Figure 1—figure supplement 1a). Examples of micrographs excluded from 2DTM searches are shown in Figure 1—figure supplement 2. At the particle level, we applied two main selection criteria prior to 3D reconstruction:

1. 2DTM-derived statistics

Due to the small molecular weight of the target, the 2DTM z-score threshold, calculated from the cross-correlations, led to the rejection of most particles and did not yield meaningful detections. Instead, we used the newly developed 2DTM p-value approach to extract particles during the initial processing step (33). Specifically, instead of using the standard first-quadrant p-values, we calculated a three-quadrant p-value to retain particles with low 2DTM z-scores but high 2DTM SNRs (see Methods). We found that a p-value threshold of 8.0 consistently gave us the best reconstruction. Following the initial extraction, as described in the Methods section, we applied additional selection steps based on several 2DTM-derived metrics, including 2DTM SNR, and the pixel-level average and standard deviation of cross-correlations from the angular search.

2. CTF-based ice thickness

We excluded images of thicker samples with reduced high-resolution signal, as indicated by their estimated sample thickness from CTF fitting (37). The distribution of mean defocus and thickness of untilted images are shown in Figure 1—figure supplement 1b and c. The mean defocus was 9900 Å. Given that the largest dimension of the protein kinase is ~ 65 Å, we selected images with estimated thickness between 100 and 800 Å. We found that images with thickness at 300-400 Å contained the most particles based on our criteria (Figure 1—figure supplement 1d and e).

Applying these selection criteria drastically reduced the final stack size. For experiments in Figure 1 and Figure 2, around ~8,000 particles were used to generate the reconstruction, only approximately 10% of what was used in the original single-particle pipeline. Despite the order-of-magnitude reduction in data, the resulting 3D map showed a significant improvement at the ATP and IP20 binding sites. However, we note that the global FSC in Figure 1b is not reliable due to

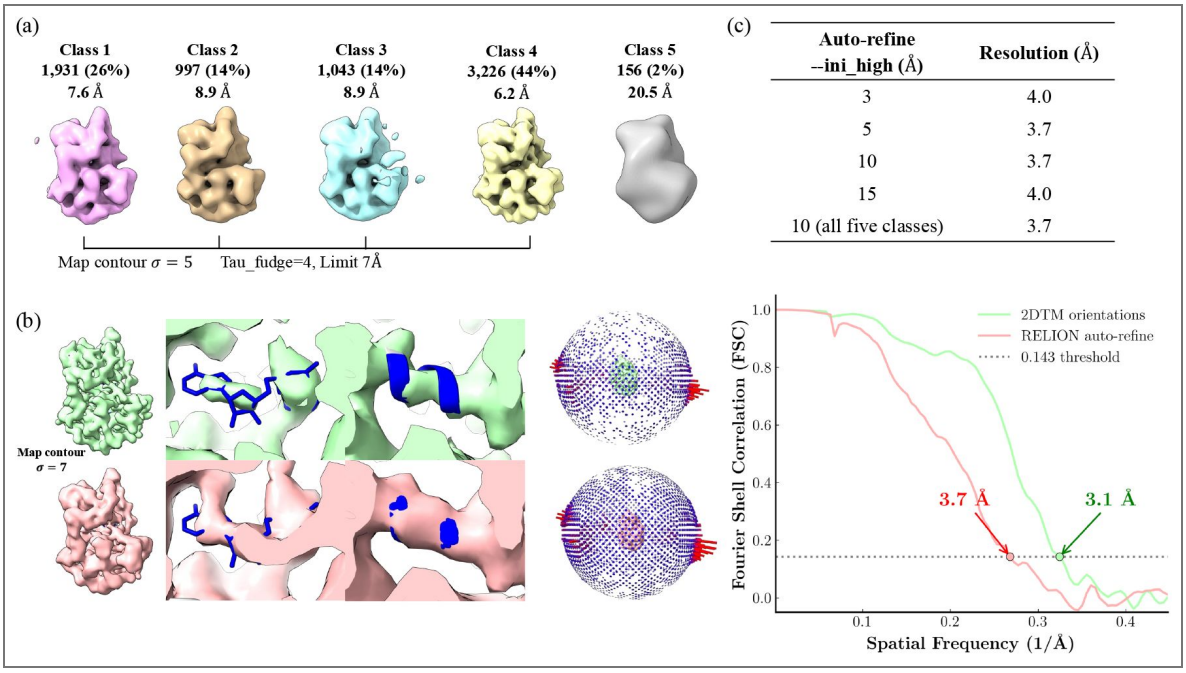


Figure 3. RELION processing of 2DTM-derived particle stack.

(a) RELION 3D classification without angular refinement of the 2DTM-derived particle stack from Figure 1, using five classes (Tau_fudge=4, angular limit 7 Å). Classes are colored individually and labeled with particle count, percentage of the dataset, and estimated resolution. Classes 1–4 were merged for 3D reconstruction; Class 5 (156 particles, ~2%, ~20.5 Å) was excluded. Map contour σ = 5. (b) Comparison of 3D reconstructions using 2DTM orientations directly (top row, green) and RELION auto-refine (bottom row, red). From left to right: full map, zoomed view of the ATP binding pocket and deleted residues 222–227 with the atomic model (blue) overlaid (map contour σ = 7), orientation distribution, and Fourier shell correlation (FSC) curves. The 2DTM-orientation reconstruction reaches 3.1 Å and the RELION auto-refined reconstruction 3.7 Å at the FSC=0.143 threshold. In both maps, densities for ATP and the backbone of the deleted residues were recovered, but the 2DTM-derived orientations yielded sharper and more continuous density in the omitted regions. (c) RELION auto-refinement resolution as a function of the initial low-pass filter (`--ini_high`). The final row includes all five classes (i.e., keeping Class 5), yielding 3.7 Å, identical to the reconstruction without Class 5.

the use of the high-resolution template and the resulting template bias (29). Nevertheless, the recovery of density not included in the template confirms that the alignment of the selected particles was sufficiently accurate to generate clear density for the binding sites.

Our experiments underline the importance of selecting good particles, rather than maximizing the number of particles selected. Previous studies have pointed out that many particles in the final stack are unnecessary and removing them can improve reconstruction (30, 38). In our experiments, we found that including a larger number of low-quality or misaligned particles, or false positives, may boost the global FSC but blur out weak features such as ligand densities. To test this, we generated a particle stack using a lowered p-value threshold of 7.0. This larger stack (13,669 particles) led to the degradation of signal in the deleted regions, likely due to increased false positives interfering with the true signal (Figure 4). Similarly, applying a 2DTM SNR threshold of 7.5 produced a particle stack of comparable size (8,456 particles) but lower quality, as demonstrated by poor reconstructions in the deleted regions (Figure 4). We also observed preferred orientation as shown by the angular distribution plot of the final stack in Figure 1b. Since we applied very stringent selection, only 2,551 particles were extracted from the tilted images. However, adding these particles did not improve the density at the deleted regions (Figure 4). Although particles from tilted images provide additional angular views, the images often have thicker ice, reducing high-resolution information for accurate alignment. Thus, careful particle curation can enable high-resolution reconstruction of sub-50 kDa complexes with an order-of-magnitude fewer particles.

Global density recovery assessed by composite omit maps

To quantify the effect of template bias across different particle selection conditions, we computed a bias metric Ω adapted from (29) (Figure 4). Across the tested conditions, Ω ranged from 46% to 53%, confirming that template bias is present in the reconstruction for features included in the template. To assess whether we can fully eliminate template bias, we performed a composite omit map experiment. We generated 36 omit templates from the X-ray model, each deleting ~ 10 non-overlapping residues distributed across the protein, including peripheral and surface-exposed regions. For each template, an independent 2DTM search and reconstruction was performed using orientations determined from the corresponding omit-template search. For each reconstruction, density surrounding the omitted residues was extracted and assembled into a composite map (Figure 5), such that each voxel is contributed by a single omit reconstruction (see Methods). This design ensures that the density at each location is derived from a reconstruction in which the corresponding residues were not present in the alignment template. The resulting composite map shows that density can be recovered at distributed locations across the structure, including regions far from the alignment center. Recovery is variable across sites, with some regions exhibiting weaker or fragmented density. This variability is consistent with local differences in structural heterogeneity and residual alignment error.

Using predicted structures as templates

It is possible that experimental structures are unavailable for the target of interest. We examine whether predicted structures can be used as templates to validate the predictions, or identify novel structures or interactions. We generated a predicted structure for the protein kinase using AlphaFold3 (39). The atomic model includes IP20, ATP, and Mn^{2+}. A high-resolution template was simulated using the same parameters as above using the `simulate` program (32). We show the comparison between the X-ray model and AlphaFold3 model in Figure 6a. Overall, the structures show good agreement, with an RMSD of 0.45 Å across 336 aligned residue pairs. The differences between the AlphaFold3 model and the X-ray structure of protein kinase A are mainly found in the flexible loop regions (e.g., residues 53–55), the surface-exposed side chains, and IP20. In the experiment in Figure 2a, residues within 3 Å of ATP were removed from the X-ray-derived template. Here, the same residues were deleted from the AlphaFold3-derived template, resulting in a remaining molecular weight of 37.3 kDa. Shown in Figure 6b, the densities of ATP in the AlphaFold3-derived reconstruction was slightly worse than that obtained using the X-ray-

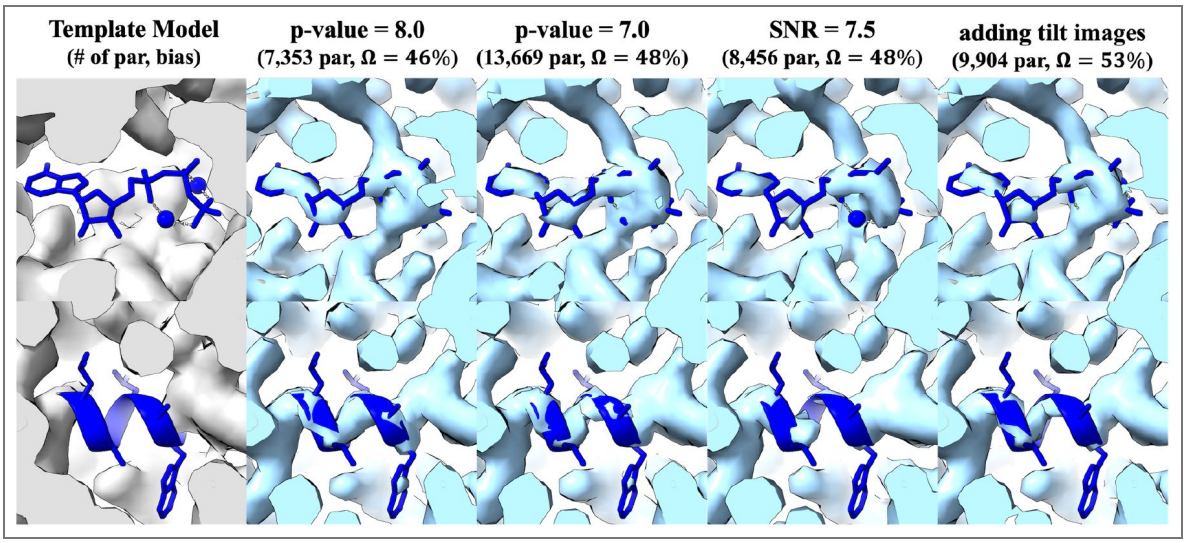


Figure 4. 2DTM reconstructions with varying particle selection parameters.

The first column shows the template model with the X-ray structure overlaid (deleted atoms in blue). The remaining columns show reconstructions using different selection thresholds (map contour $\sigma = 5$). The top and bottom rows show zoomed views of the ATP-binding pocket and deleted residues 222–227, respectively. Each column header lists the number of particles and the template bias Ω, defined as $\Omega = (\sum_{mask} V_{full} - \sum_{mask} V_{omit}) / \sum_{mask} V_{full}$, where V_{full} and V_{omit} are reconstructions using orientations and particles derived from independent 2DTM searches with the full and omit templates, respectively, and the sum is restricted to the omission mask derived from the difference between the two templates. Ω = 0 means the omit reconstruction recovers the same density (no bias from the template), while Ω = 1 means all density disappears without the template. Q-scores for ATP and residues 222–227 are reported in Figure 4—source data 1. Template bias was calculated using the `2DTM_postprocess_tool` Python package, adapted from (29). The omission mask used to compute Ω is shown in Figure 4—figure supplement 1.

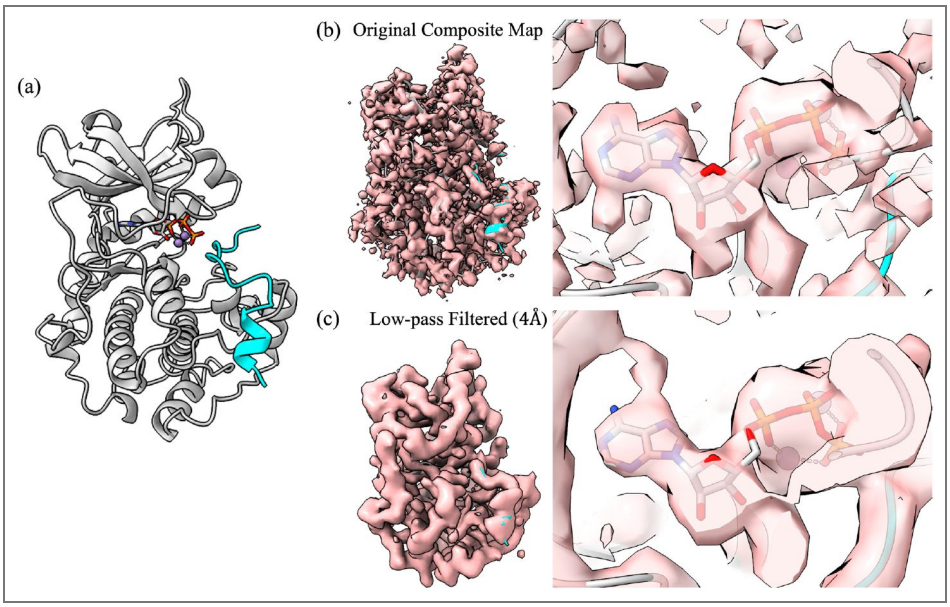


Figure 5. Site-specific composite omit map assembled from 36 partial-deletion reconstructions.

Each of the 36 omit templates deletes ~10 non-overlapping residues. For each template, local density was extracted within 3 Å of the omitted atoms, with neighboring residues (i±1) excluded using a 2 Å cutoff on backbone Cα atoms. Local density patches were assigned uniquely and merged such that each voxel is contributed only by the reconstruction in which the corresponding region was omitted from the template. (a) Atomic model of the protein kinase (PDBID: 1ATP). (b) Left: composite omit map. Right: zoomed view of the ATP-binding pocket. (c) Left: composite omit map low-pass filtered to 4 Å. Right: zoomed view of the ATP-binding pocket. The composite demonstrates that density can be recovered at distributed locations across the protein, including peripheral and surface-exposed regions, although the quality of recovery varies across sites. Map contour σ = 4.

derived template, highlighting the importance of template accuracy in enabling efficient detection and reconstruction of small complexes. We also observed that both reconstructions resemble closely the search templates, including differences in the side chain densities (Figure 6a). This suggests potential template bias and false positives remained in the particle stack despite the use of stringent selection criteria. Further improvements could involve systematically deleting residues from the predicted model and assessing the impact on the reconstruction, as demonstrated in the composite omit map experiment.

How small a particle can we study by single-particle cryo-EM

The cross-correlation between the particle image and the reference needs to be larger than the expected cross-correlation between noise and the reference to be alignable (17). This lower molecular weight limit of single-particle cryo-EM was estimated as 38 kDa assuming a total exposure (N_e) of 5 $e^-/Å^2$ (16). Later, the estimated limit was lowered to 17 kDa (17). The primary difference between these two predictions lies in the statistical criteria used to assess image visibility: Henderson in (16) required that the intensity of the average Fourier component should be three times the standard deviation of the shot noise, while Glaeser (17) applied the Rose criterion. More recent work has shown that it is now possible to reconstruct proteins below 50 kDa and even smaller nucleic acids, with predictions that the lower molecular weight limit could be extended below 20 kDa (40). Specifically, a high-resolution 3D reconstruction of the 14 kDa hen egg white lysozyme (HEWL) was obtained from a simulated dataset generated with an ideal phase plate (41). More recently, the 40 kDa Aca2–RNA complex was reconstructed to 2.5 Å resolution using Blush regularization, a data-driven denoising prior that improves image alignment for small particles with low signal-to-noise ratios (42). Here, following the rationale of 2DTM, we sought to calculate the lower molecular weight limit for hydrated biological samples for single-particle cryo-EM that takes into account advancements in instrumentation made over the past decades.

During 2DTM, cross-correlation coefficients are calculated between the particle image and 2D projections of the template. *Peaks* in the cross-correlation map indicate regions of high similarities between the image and the template. A signal-to-noise ratio (SNR) can be defined as the maximum correlation observed when aligning an image against a reference, divided by the standard deviation of the correlations from the background. We consider two scenarios: (1) the image contains pure random noise and the corresponding SNR is SNR_n; (2) the image contains real phase contrast from the target and the corresponding SNR is SNR_s. In both scenarios, the SNR is expected to be larger than zero because even in the presence of pure noise, a positive correlation will be obtained after aligning the noise image to a reference. Determining the lower molecular weight limit is equivalent to identifying the intersection at which SNR_n and SNR_s are equal. When SNR_s is smaller than SNR_n, it will be impossible to distinguish signal from noise.

For the first case, cross-correlation coefficients between two pure Gaussian noise images with N_p pixels can be approximated with a Gaussian distribution with zero mean and a variance of $1/N_p$, where N_p is the number of pixels in the image. In cryo-EM particle alignment, a five-dimensional search is performed for each particle, including two translational parameters (x and y) and three orientational parameters (Φ, θ, ψ). For each particle, N_s correlation coefficients are calculated to find the correct alignment. N_s is dependent on the size of the particle and the resolution limit of the alignment. A larger particle and higher resolution limit require a more finely sampled search space. The parameter set, $\Theta_0 = \{x_0, y_0, \Phi_0, \theta_0, \psi_0\}$, corresponding to the maximum value among N_s correlations, is then used to register the particle's alignment. We define SNR_n using the maximum and standard deviation of N_s correlation coefficients. Based on Supporting Information A,

$$\mathrm{SNR}_n = \sqrt{2 \log N_s}. \tag{1}$$

This means the more search locations are evaluated, the higher the likelihood of observing a large correlation value purely by chance.

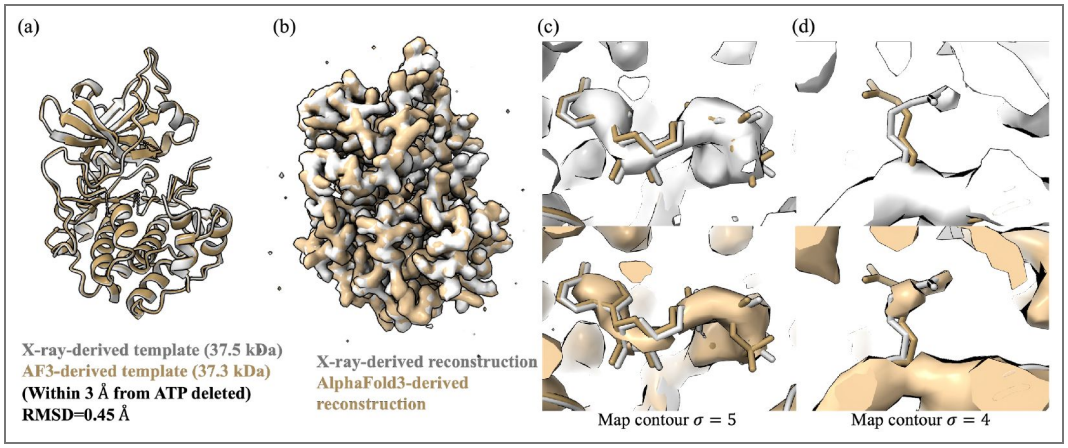


Figure 6. 2DTM reconstruction using the AlphaFold model as the search template.

(a) Structural comparison between the X-ray model (PDB ID: 1ATP, gray) and the AlphaFold3 predicted model (tan) (39). Residues within 3 Å of ATP were deleted from both templates and not shown. RMSD between the undeleted structures is 0.45 Å. (b) 2DTM-derived maps using the AlphaFold3 template (tan) and the X-ray template (gray). (c) Reconstruction at the ATP-binding site using the X-ray model (top) and the AlphaFold3 model (bottom) as the template (map contour $\sigma = 5$). (c) Reconstruction at residue 18 (ARG) on IP20 using the X-ray model (top) and the AlphaFold3 model (bottom) as the template (map contour $\sigma = 4$).

For the second case, we follow Henderson's assumption that the particle is roughly spherical and consists of only randomly positioned carbon atoms (16). Cross-correlations are calculated between a M -frame summed image and 2D projections from the perfect reference. Similarly, SNR_s is defined as the ratio between the maximum cross-correlation and the standard deviation of background correlations. Based on Supporting Information B,

$$\mathrm{SNR}_s = \sqrt{4fN_0D^2 \int_k \mathrm{CTF}(k)^2 \left(\sum_{i=1}^{M} Q\left(k, N_i\right) \right)^2 dk} \tag{2}$$

where $Q(k, N_i)$ is the normalized exposure-weighting transfer function at spatial frequency k for cumulative frame exposure N_i (19), N_0 is the exposure per frame, D is the particle diameter, and f is the fraction of electrons being elastically scattered. Assuming a total exposure of 5 $e^-/\text{Å}^2$ and a single-frame acquisition (16), we derived a simplified form of Eq. 2 by assuming a realistic CTF with multiple oscillations, such that the integral $\int_k \mathrm{CTF}(k)^2 dk \approx 0.5$. This leads to

$$\mathrm{SNR}_s = \sqrt{2fN_e}D. \tag{3}$$

Under these conditions, we estimate a minimum detectable molecular weight of 38.0 kDa. While this result is numerically similar to Henderson's original estimate (16), the underlying assumptions differ somewhat. (16) assumed ideal imaging conditions without CTF oscillations. In contrast, we assume standard cryo-EM conditions with a realistic CTF and compute the signal and noise terms using different models. In particular, our calculation of the number of correlations calculated (N_s) is different. Nevertheless, the agreement in molecular weight limit suggests that this value is a reasonable estimate given the main experimental assumptions (single frame acquisition, 5 $e^-/\text{Å}^2$).

For movies with multiple frames, Eq. 2 can be numerically integrated over spatial frequency range [$k_{\min}$,$k_{\max}$] to calculate SNR_s. Eqs. 1 and 2 are then plotted across a range of molecular weights to identify their intersection point, shown in Figure 7. The parameters used for these calculations, along with their corresponding values, are listed in Table 1. Using conventional single-particle analysis conditions with a resolution limit of 2 Å, which can be achieved by collecting images using a pixel size of 1 Å/pixel, assuming a perfect beam (i.e., no envelope function), and using a total exposure of 45 $e^-/\text{Å}^2$, particles as small as 14.8 kDa can be accurately aligned through a full search. If considering the inelastic scattering from ice with a thickness of 30 nm, particles needs to be at least 16.3 kDa to be detected, which is closely consistent with the prediction of 17 kDa in (17). Ideally, if thin ice can be obtained, which is just thick enough to embed the particle, considering the defocus variation across the particle along its diameter, and assuming 10% amplitude contrast, the smallest alignable particle is 14.8 kDa. In practice, such thin ice may be difficult to achieve, leading to an increase in the weight limit. If the low-resolution contrast of the particle allows it to be roughly centered in the x,y plane, we may not need to search the entire area covered by the particle. By constraining the translational search to a 5-by-5 pixel region, the molecular weight limit can be reduced to 11.8 kDa under previously mentioned conditions. Further incorporating a 90° phase plate and using zero defocus lower the limit to 7.4 kDa under constrained search conditions. Previous work has shown that electron diffraction spots fade more slowly at liquid-helium temperatures by a factor between 1.2 and 1.8, compared to those at liquid-nitrogen temperatures (24). Assuming an additional cryo-protection factor of 1.8, particles as small as 5.7 kDa are theoretically alignable by 2DTM.

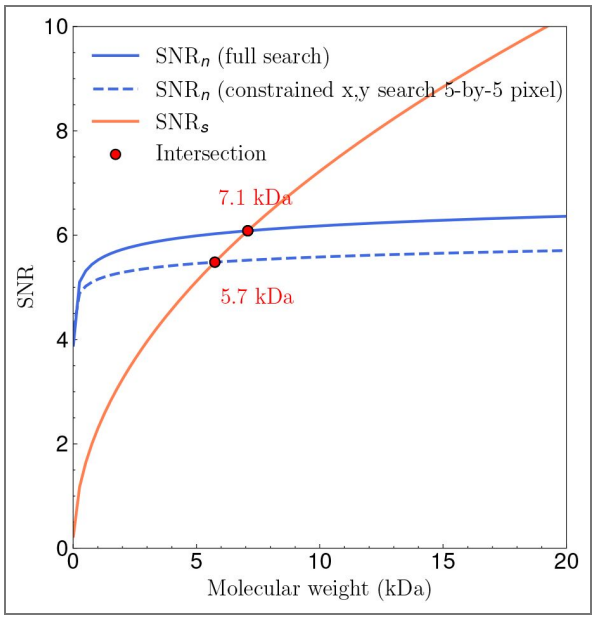
Table 1. Single-particle cryo-EM lower molecular weight limits under different assumptions.

A constrained search restricts the *x* and *y* dimensions to a 5-by-5 pixel window.

Parameters	Values	Full Search	Constrained Search
Acceleration Voltage	300 keV	14.8	11.8
Defocus	1500 nm		
Total Exposure	45 e^- ($M = 45$)		
Resolution Limits	2-120 Å		
Perfect Beam	-		
+Inelastic Scattering	ice, 30 nm	16.3	13.0
+Particle Envelope	ice thickness = particle diameter	15.0	12.0
+Amplitude Contrast	0.1	14.8	11.8
+Phase Plate	90°	9.2	7.4
+Liquid Helium	1.8×	7.1	5.7

Figure 7. Theoretical lower molecular weight limit.

A constrained search restricts the *x* and *y* dimensions to a 5-by-5 pixel window. At the minimal molecular weight, the SNR calculated from alignment noise and phase contrast are equal.



Discussions

The reference quality may be a limiting factor of single-particle reconstruction

Images collected in (30) are of high quality, where particles in many images can be visually observed. Our CTF analysis shows that 2,108 images in the untilted dataset exhibited Thon rings extending beyond 4 Å. However, the traditional single-particle processing done in (30) was limited to 4.3 Å resolution, and features at the ATP-binding site were not resolved. In the same work, the authors analyzed two other sub-100 kDa complexes, the 82 kDa homodimeric enzyme alcohol dehydrogenase (ADH) and the 64 kDa methemoglobin (metHb), and found that very few particles from many images, rather than many particles from very few images, contributed to the final stack used for reconstruction. This was consistent with our results of selecting a subset of particles for the protein kinase dataset, where most images yielded fewer than 10 particles (Figure 1—figure supplement 1d and e).

So, what is the limiting factor of the single-particle pipeline? In our assessment, we speculate that the images themselves contain sufficient signal to support a reconstruction beyond 4.3 Å resolution. A possible contributing factor is the suboptimal reference used during 3D refinement. In the workflow in (30), an *ab initio* volume generated by cryoSPARC was low-pass filtered to 20 Å and then used as the initial reference for 3D auto-refinement in RELION. Because this starting map lacked high-resolution features, particle alignment may have been less accurate. Once these misalignments were introduced, subsequent refinement iterations may not be able to fully correct them from the local refinement optimum that was created earlier. By contrast, the SAM-IV riboswitch, despite having a comparable molecular weight, has distinct asymmetric features even at low resolution. Combined with the stronger scattering of nucleic acids, this likely facilitated more accurate alignment, allowing the traditional single-particle workflow to achieve a high-resolution reconstruction (43).

Remaining gaps between experimental and theoretical limits

Despite the progress in cryo-EM, there is still a gap between our predicted lower molecular weight limit in Table. 1 and the smallest template used in our tests (36.3 kDa). Several factors may contribute to this difference:

Beam-induced motion particularly affects small particles (44). Two distinct contributions are often considered. The first arises from rapid initial specimen deformation (e.g., cryo-crinkling or buckling of the ice film), which can be substantially reduced through grid design and improved supports (30, 45). The second is a stochastic, pseudo-Brownian motion of particles that accumulates with dose and attenuates high-resolution signal. Based on the analysis of beam-induced motion in amorphous ice and scaling with particle size, this effect is expected to be on the order of a mean-squared displacement of 0.1 $Å^2$ per ($e^-/Å^2$) for a ~ 10 kDa particle (6). Over a typical total exposure of 40-60 $e^-/Å^2$, this corresponds to an accumulated RMS displacement of ~ 2-2.5 Å, which is sufficient to attenuate medium-to high-resolution signal. Although Bayesian polishing can partially compensate for beam-induced motion, it relies on sufficient particle signal to estimate individual trajectories and assumes spatially smooth motion between nearby particles (46). For very small particles, the per-particle signal may be too weak to reliably support this analysis, and stochastic particle-level motion may therefore remain only partially corrected. The protein kinase dataset was collected using UltrAuFoil grids (30), which reduce initial beam-induced specimen motion; nevertheless, residual motion, particularly pseudo-Brownian motion of small particles, remains a factor limiting the recoverable high-resolution signal in cryo-EM images.

Another factor is the **electron beam energy** used for imaging. The images were collected at 200 keV whereas the theoretical calculation presented here assumes 300 keV. Both elastic and inelastic scattering cross-sections increase at lower voltage. A recent study of sub-200 kDa complexes showed that there is no obvious choice of electron energy for imaging smaller complexes (47). Any gain in image contrast from stronger elastic scattering may be offset by the increasing inelastic

scattering (47). However, the lens aberrations are more noticeable at lower energies (48). Detectors are also generally optimized for 300 keV electrons, and practical detector DQE is generally more favorable at 300 keV than at 200 keV. Our calculation assumed a perfect beam whereas the experimental images suffer from spherical and chromatic aberrations. New developments in aberration correction will help obtain atomic-resolution structures of small complexes (49–52).

Non-ideal experimental conditions will also widen the gap between theoretical and experimental limits. For example, we assumed the ice is just thick enough to accomodate the particle whereas the images in the protein kinase dataset have varying ice thickness as shown in Figure 1—figure supplement 1c. Additionally, we did not perform a defocus search during 2DTM, instead, we used the average defocus values estimated by CTFFind. Incorporating the defocus search directly into 2DTM may give more accurate estimations in certain cases. Another contributing factor is the **template generation strategy**. In Figure 1—figure supplement 4, the difference map between the template and the reconstruction for the experiment in Figure 1a was generated using the `diffmap.exe` program (53) with a protein soft mask. Under this masking setup, coherent difference densities are most prominent at the ATP binding site and residues 222–227, which were deleted from the template. Residual noisy densities in other regions indicate a limitation of the current template generation approach, which may not fully capture the solvent background or accurately model the electrostatic potential of atoms within their molecular environment. As shown by recent work (54) on empirical scattering-factor inference from cryo-EM maps, environment-aware scattering factors that incorporate bonding and charge redistribution effects can improve agreement between model and map. Incorporating such learned scattering factors into template construction would provide a more accurate forward model of the electrostatic potential, potentially reducing spurious residual densities and improving template–map consistency. Additionally, **structural flexibility and conformational heterogeneity** further attenuate high-resolution signal, effectively acting as an additional envelope that reduces the detectable signal relative to the idealized rigid-particle model considered here.

Taken together, these factors explain why the current 2DTM implementation has not yet reached the theoretical lower molecular-weight limit estimated here, and why reliable alignment of proteins much below ~ 40 kDa remains challenging. Advancements in both microscope optics, detector performance, image processing algorithms, and sample preparation strategies will be essential to close this experimental-theoretical gap and make high-resolution cryo-EM of small complexes a reality.

Implications for structure-based drug design

For structure-based drug design (SBDD), obtaining a high-resolution structure of the target (and its ligand complex) is a critical step. X-ray crystallography has long been the dominant method for this purpose, but many protein targets are hard to crystallize, especially those that are flexible or membrane-embedded (55). NMR can handle very small proteins or nucleic acids in solution, but typically offers lower resolution information and is limited to proteins below 25 kDa, otherwise spectra become too complex to interpret without isotope labeling and extra processing (8, 9). Cryo-EM is emerging as a powerful alternative and complementary method and enables structural and functional studies under conditions more closely resembling the native cellular context. One of the major caveat for cryo-EM, however, is the reconstruction of sub-50 kDa complexes.

The 2DTM-based single-particle alignment and reconstruction workflow we proposed here **simplifies** the conventional single-particle pipeline by foregoing iterative rounds of 2D classification, *ab initio* modeling, 3D classification and refinement. 2DTM directly returns particles with their *x*, *y* position, Euler angles and defocus in *one pass*. Aside from the computational cost of the search itself, the workflow is trivial compared to a single-particle pipeline. For a typical single-particle dataset of ~ 2,000 micrographs (5k × 4k pixels), a 2DTM search without defocus refinement completes in approximately one day on 64 NVIDIA A6000 GPUs. Once particles are located with their orientations and positions, a single 3D reconstruction is sufficient without further refinement, eliminating the iterative 2D classification, *ab initio* modeling, 3D classification and refinement steps of a conventional pipeline.

Another motivation for 2DTM is overcoming the **size limit** of cryo-EM. Conventional single-particle workflows have struggled to achieve good reconstructions for particles below 50 kDa, where low contrast makes alignment unreliable with an imperfect reference. To date, only a handful of isolated sub-100 kDa structures have been solved at <4 Å resolution. Notable examples include the 52 kDa streptavidin tetramer resolved to 3.2 Å using a Volta phase plate and C_s corrector (56), and to 2.6 Å using graphene grids (57); the 64 kDa methemoglobin at ~ 2.8 Å resolution (30); the 39 kDa SAM-IV riboswitch at 3.7 Å resolution (43), and the previously mentioned 40 kDa Aca2–RNA complex at 2.5 Å (42). In a recent study reporting the 3D reconstruction of the protein kinase also used in the present study, the authors achieved near-atomic resolution with a carefully tuned *ab initio* procedure, using a narrow band of spatial frequencies that was centered at 5 Å and incremented in small steps between reconstruction iterations to 2.3 Å resolution (58). These complementary findings highlight the importance of high-resolution information for small particle alignment. In our case, the 2DTM-based reconstruction method improves alignment by explicitly utilizing the high-resolution information from a perfect reference, enabling robust recovery with far fewer particles. Our theoretical estimation of the lower molecular weight limit further highlights that for perfect images and a perfect reference, complexes as small as 5.7 kDa can be accurately aligned when using liquid helium cooling and a phase plate.

The workflow we describe here also has the potential to operate as a **screening platform**, enabling structure-guided optimization of small-molecule ligands even for targets below the traditional cryo-EM size limit. High-resolution cryo-EM is already being applied to drug targets that resist crystallization: for example, the high-resolution (up to 1.8 Å) maps of human CDK-activating kinase bound to 15 different inhibitors revealed detailed inhibitor interactions and water networks in the active sites (59). In another case, cryo-EM captured a novel allosteric mechanism for protein inhibition of the human ATP-citrate lyase that enhances the target's "druggability" (60). These studies show that resolving the ligand-bound structures can directly guide design of novel therapeutics. Our workflow can broaden this approach to smaller drug targets. In principle, one could incubate a sub-50 kDa target with various inhibitors then apply 2DTM using the apo structure as the template, which can be determined from *in vitro* experiment or AlphaFold predictions. The ligand-bound complex can be located, aligned and reconstructed using 2DTM. 2DTM thus offers a structure-based assay: binding of each inhibitor produces a distinct density feature at the binding site, streamlining hit validation.

In summary, by overcoming the ~50 kDa barrier, 2DTM broadens the possibilities for structural studies of many previously inaccessible drug targets that have remained difficult to reconstruct, with the ultimate goal of integrating cryo-EM into high-throughput SBDD.

Toward data-driven refinement of AlphaFold3 models via 2DTM

Our results demonstrate that AlphaFold3-predicted structures can be used directly as templates in 2DTM searches to pick particles and reconstruct high-resolution maps, even in the absence of any experimentally determined model. This validates the use of AlphaFold3 as a starting point for structure determination. While the predicted structure may deviate from the true conformation, as shown in Figure 6a, the resulting cryo-EM map provides information that can guide further refinement of the AlphaFold3 model. This refined model could then be used as an improved template in a second round of particle picking and reconstruction, allowing recovery of more accurate particles and a higher-quality map. In Figure 6d, we show a deleted residue and the reconstructed densities obtained using both the X-ray template and the AlphaFold3 template. Although the densities are weak and discontinuous, both maps are consistent with the side chain conformation observed in the X-ray model. While we have not yet tested this type of refinement, these preliminary results suggest that this data-driven approach could progressively improve both the structural model and the final reconstruction, starting entirely from prediction. We have demonstrated residue-level omission validation using the X-ray template via a composite omit map (Figure 5); however, extending this approach to AlphaFold3-based templates remains computationally expensive and we consider this an important direction for future work.

Methods

Cryo-EM data set and image processing

Unaligned movies of the protein kinase were downloaded from EMPIAR-10252 (30). The dataset contains 4,809 images, among which 2,488 are from untilted samples and 2,321 are collected at 30° tilt. Motion correction and exposure-weighting were performed using the MotionCor2 program (61). We kept the same procedure as the original publication by using 5×5 tiled frames with a B-factor of 250 $Å^2$ and a binning factor of 2. We used the exposure weighted summed frames for CTF fitting using CTFFind5 (37) with a box size of 512 and a resolution range of 4-30 Å. We selected images with CTF fitting scores between 0.05 and 0.2 for the untilted dataset and 0.04-0.13 for the tilted dataset. Images that were excluded are shown in Figure 1—figure supplement 2 for comparison.

2DTM search

After selecting good micrographs based on their CTF fit, 2,314 images of untilted and 2,252 of tilted samples with a pixel size of 1.117 Å/pixel were used for 2DTM. For the experiment in Figure 1, ATP, Mn^{2+}, and residues 222–227 were deleted from the X-ray model (PDBID: 1ATP) before template simulation. For the experiment in Figure 2, residues within a radius of 3 Å and 5.5 Å from ATP were deleted from the X-ray model. IP20 was also deleted for the test in Figure 2b. Modified atomic coordinates were generated using UCSF ChimeraX (62). High-resolution 3D templates were then generated from the modified models using program `simulate` in *cis*TEM (32). A uniform B-factor of 30 $Å^2$ was applied to all atoms. 2DTM searches were done using an angular search step of 2.5° for out-of-plane angles and 1.5° for in-plane angle for all tests with no defocus search.

Particle extraction and selection

To streamline post-processing of 2DTM, we implemented a dedicated Python toolkit `2DTM_postprocess_tool`. The module contains two command line tools:

1. `extract-particles`

Input arguments are (i) the *cis*TEM 2DTM SQLite project database, (ii) the IDs of the search and associated CTF-estimation jobs, and (iii) the image pixel size. The program extracts candidate particles in each image in the specified job using one of the three 2DTM metrics: SNR, z-score, or p-value.

For the p-value, the user may choose first-quadrant (1Q) or three-quadrant (3Q) definitions (33): after probit transformation of z-score and SNR, 1Q keeps candidates where both probit-zscore and probit-SNR are greater than 0, while 3Q keeps candidates where at least one of the two values is greater than 0. In the protein kinase example, we first located local maxima in 2DTM SNR maps (exclusion radius = 10 pixels; micrograph border mask = 92 pixels to avoid truncated particles), then calculated the three-quadrant p-values and stored particles with p-values larger or equal to 8.0.

2. `filter-particles`

This function provides secondary quality selection metrics based on the CTF fitting quality, sample thickness, and particle-level statistics from the 2DTM angular search (mean and standard deviation of per-pixel cross-correlations across sampled orientations). For the present study we required a SNR> 6.0, an ice thickness between 100 and 800 Å, and a CTF fitting score between 0.05-0.2 for untilted images (or 0.04-0.13 for tilted). Particles whose angular search mean cross-correlation was negative or standard deviation of cross-correlations exceeded 1.1 were also discarded.

Both steps output a star file that can be used for the following 3D reconstruction.

3D reconstruction

Particle stacks processed by `filter-particles` and alignment parameters were imported into *cis*TEM as a refinement package for single-particle processing. A 3D reconstruction was generated using the *cis*TEM program `reconstruct3d`. UCSF ChimeraX was used for visualizing the final reconstructions.

RELION processing

Particles selected by 2DTM were subjected to 3D classification in RELION using a tau fudge factor of 4 and an E-step resolution limit of 7 Å, resulting in five classes. The best-resolved classes were subsequently used for either (i) 3D reconstruction without angular refinement, using the 2DTM-derived orientations directly, or (ii) 3D auto-refinement with alignment, using C1 symmetry and a 3.7° angular sampling step with the corresponding 10 Å low-pass filtered template as reference (`--ini_high` 10 Å). A soft mask was then generated from the resulting map by applying a 15 Å low-pass filter and using a soft-edge of 6 pixels. The final map was produced by post-processing with B-factor sharpening and low-pass filtering. To test the effect of the initial reference resolution, auto-refinement was additionally repeated with `--ini_high` values of 3, 5, and 15 Å. The all-five-class reconstruction was performed using `--ini_high` 10 Å with the same masking and postprocessing protocol.

Template bias quantification

To quantify template bias in the omitted region, we computed the metric $\Omega = (\sum_{mask} V_{full} - \sum_{mask} V_{omit}) / \sum_{mask} V_{full}$, where V_{full} and V_{omit} are 3D reconstructions generated using orientations and particles derived from independent 2DTM searches with the full and omit templates, respectively (29). The omission mask was derived by subtracting the omit template from the full template and thresholding at one-fifth of the mean of the top-100 difference voxels, yielding 959 masked voxels concentrated at the ATP-binding site and residues 222–227 (Figure 4—figure supplement 1). Ω = 0 indicates no template bias (the omit reconstruction recovers identical density), while Ω = 1 indicates full bias (all density vanishes when not included in the template). Bias was computed using the `measure-template-bias` function in the `2DTM_postprocess_tool` Python package (*https://github.com/kekexinz/2DTM_postprocess_tool*), also available in the official *cis*TEM repository (*https://github.com/timothygrant80/cisTEM*) (29, 33).

Composite omit map construction

To avoid template bias and test spatial generalization of density recovery, we constructed a composite omit map from multiple independent reconstructions. Starting from the X-ray model (PDBID: 1ATP), we generated 36 omit templates, each deleting a set of ~ 10 non-overlapping residues distributed across the structure. For each omit template, a separate 2DTM search was performed, and particle orientations obtained from that search were used to compute a corresponding 3D reconstruction. For each omit reconstruction k, a binary mask $M_k(\mathbf{x})$ was defined to isolate density associated with the omitted residues. Voxels within 3 Å of any atom in the omitted residues were included, while voxels within 2 Å of backbone Cα atoms of adjacent residues (i ± 1 in sequence) were excluded to reduce contamination from neighboring template-included regions.

To ensure that each voxel in the final composite map is contributed by at most one reconstruction, overlapping mask regions were resolved deterministically. A support map was first computed to identify voxels covered by multiple masks. For voxels with multiple assignments, the contributing reconstruction was selected based on the minimal distance between the voxel and the centroid of the omitted residues associated with each mask.

Each binary mask was then converted to a soft mask by applying a cosine-edge taper using a distance transform, with a taper width of two voxels, ensuring smooth transitions without expanding the support beyond the assigned region. The final composite map was constructed as

$$\rho_{\text{composite}}(\mathbf{x}) = \sum_k M_k^{\text{soft}}(\mathbf{x})\rho_k(\mathbf{x}), \tag{4}$$

where $\rho_k(\mathbf{x})$ is the reconstruction corresponding to omit template k, and $M_k^{\text{soft}}(\mathrm{x})$ is the tapered mask. By construction, each voxel in the composite map is derived from a single omit reconstruction.

Data availability

The current manuscript is a computational study, so no data have been generated for this manuscript. The `2DTM_postprocess_tool` Python package, which includes the `extract-particles`, `filter-particles`, and `measure-template-bias` command line tools used in this work, is available at *https://github.com/kekexinz/2DTM_postprocess_tool* and in the official *cis*TEM repository at *https://github.com/timothygrant80/cisTEM*..

Acknowledgements

We thank the members of the Grigorieff lab for the fruitful discussion of this work. We are especially grateful to Dongjie Zhu for sharing and testing his new methods and for many insightful conversations.

Supplementary Note 1: Estimating the lower molecular weight limit

A. SNR from alignment noise

For two independent Gaussian noise images of N_p pixels, the normalized cross-correlation, r_n, is zero on average. The distribution of r_n when N_p is large can be approximated by a Gaussian distribution

$$r_n \sim \mathcal{N}\left(0, \frac{1}{N_p}\right). \tag{S1}$$

In cryo-EM particle alignment, normalized cross-correlations are computed between a noisy particle image and a set of clean 2D projections generated at sampled orientations, in order to determine the best-matching position (x, y) and orientation. For each particle image, N_s cross-correlations are evaluated, and the alignment is assigned based on the position and orientation corresponding to the maximum cross-correlation. This process is equivalent to drawing the maximal value from N_s Gaussian distributed random variables with zero mean and variance of $\frac{1}{N_p}$. The upper bound of the expectation of this maximum is (64)

$$\mathbb{E}(r_{n,\max}) \leq \sigma_{cc}\sqrt{2\log N_s} = \frac{\sqrt{2\log N_s}}{\sqrt{N_p}}. \tag{S2}$$

We define the signal-to-noise ratio (SNR) of alignment noise as the number of standard deviations (SDs) by which the maximum cross-correlation exceeds the SD of cross-correlations computed across a pure noise image:

$$\text{SNR}_n = \frac{\sqrt{2\log N_s}}{\sqrt{N_p}} / \sqrt{1/N_p} = \sqrt{2\log N_s}. \tag{S3}$$

Assuming that the high-resolution limit for alignment is $d = 1/k_{max}$ (Å), where k_{max} is the maximum spatial frequency (Å^{-1}). The ideal pixel size is then $p = d/2$ (Å/pixel). For a particle of diameter D (Å), a five-dimensional search includes the following components:

1. in-plane rotations: $\frac{\pi D}{p}$.
2. Out-of-plane viewing directions:

$$N_{\text{views}} \approx \frac{4\pi}{(\Delta\theta)^2} = \frac{4\pi}{(2p/D)^2} = \pi\left(\frac{D}{p}\right)^2$$

3. x, y shifts: $\frac{D^2}{p^2}$.

The total number of correlations calculated during a five-dimensional search is then

$$N_s = 32\pi^2\left(\frac{D}{d}\right)^5. \tag{S4}$$

B. SNR from phase contrast

Fraction of electrons being elastically scattered up to a resolution limit

When the image contains phase contrast, the SNR is defined as the number of SDs by which the normalized cross-correlation exceeds the background SD, consistent with the definition used in the 2DTM implementation (26).

We assume that the particle consists of only randomly positioned carbon atoms. The electron atomic scattering factor of carbon can be approximated as a sum of (normally) five Gaussians (65, 66):

$$f^{(e)}(k) = \sum_{i=1}^{5} a_i \exp\left(-b_i k^2\right), \tag{S5}$$

where a_i and b_i are the fitting parameters up to 6 Å^{-1} (67), and k is the spatial frequency (Å^{-1}). The differential scattering cross-section is:

$$\frac{d\sigma_e}{d\Omega} = \left|f^{(e)}(\theta)\right|^2, \tag{S6}$$

where θ is the scattering angle and

$$k = \frac{\sin\left(\frac{\theta}{2}\right)}{\lambda}, \tag{S7}$$

where λ is the electron wavelength. The differential cross-section for a single, isolated atom is related to θ by:

$$\frac{d\sigma_e}{d\Omega} = \frac{1}{2\pi\sin\theta}\frac{d\sigma_e}{d\theta}. \tag{S8}$$

We now integrate to calculate the total scattering cross-section:

$$\sigma_e = \int_0^{\theta_{\max}} 2\pi\sin\theta\left|f^{(e)}(k)\right|^2 d\theta \tag{S9}$$

Assuming protein density $\rho \approx 0.8$ Da/Å^3, the number of carbon atoms equivalent to a spherical protein of diameter D (Å) is (16):

$$N_c = \frac{M_w \cdot 9.1}{110} = \frac{4}{3}\pi\left(\frac{D}{2}\right)^3 \rho \cdot \frac{9.1}{110}. \tag{S10}$$

The fraction of electrons being elastically scattered by the particle is:

$$f = \frac{\sigma_e(\mathrm{C})N_c}{D^2}. \tag{S11}$$

The exit wave function at distance z below the specimen is (68)

$$\Psi_{\mathrm{exit}}(x,z) = \Psi_{\mathrm{inc}}(z) + \Psi_{\mathrm{inc}}(z)\cdot i\sigma o(x), \tag{S12}$$

where $o(x)$ is the projected potential of the molecule. We can now relate the molecular weight of the particle to the Fourier component of the image:

$$\begin{aligned}\frac{\text{total intensity of scattered wave}}{\text{total intensity of incident wave}} &= \frac{\int_{N_p}\left|\Psi_{\mathrm{inc}}(z)\cdot i\sigma o(x)\right|^2 dx}{\int_{N_p}\left|\Psi_{\mathrm{inc}}(z)\right|^2 dx}\\ &= \frac{\int_{N_p}\sigma^2|o(x)|^2 dx}{\int_{N_p}1dx}\\ &= \frac{\sigma^2}{N_p}\int_{N_p}|o(x)|^2 dx\\ &= \frac{\sigma^2}{N_p}\int_{k_{\min}}^{k_{\max}}|F(k)|^2 dk\\ &= f,\end{aligned} \tag{S13}$$

where $F(k)$ denotes the 2D Fourier transform of the projected Coulomb potential $o(x)$ of the particle.

Image formation model

The wave function in Fourier space after lens aberration is (6, 68):

$$\begin{aligned}\Psi_{\mathrm{pc}}(k) &= \Psi_{\mathrm{exit}}(k,z)\cdot\exp[-i\chi(k)]\\ &= \Psi_{\mathrm{inc}}(z)[\delta(k) + i\sigma F(k)]\cdot\exp(-i\chi(k)),\end{aligned}$$

in which the lens aberration function (69) is

$$\chi(k) = \pi\Delta f\lambda k^2 - \frac{\pi}{2}C_s\lambda^3 k^4 \tag{S14}$$

Here, Δf is the defocus (positive for underfocus) and C_s is the spherical aberration. The contrast transfer function (CTF) is defined as

$$\mathrm{CTF}\ (k) = -\sin\chi(k). \tag{S15}$$

The Fourier transform of the linearized intensity is (6, 68):

$$I_0(k) = \left|\Psi_{\mathrm{pc}}(k)\right|^2 \approx \delta(k) + 2\sigma F(k)\,\mathrm{CTF}(k). \tag{S16}$$

Eq. S16 is the *linear model of image formation* in cryo-EM.

Given the per-frame per-unit area exposure $\frac{d^2}{4}N_0$ where N_0 is the exposure per-frame, the observed noisy image of a single frame is

$$I_i^{\text{noisy}}(x) \sim \text{Poisson}\left[\frac{d^2}{4} N_0 \mathscr{F}^{-1}[\delta(k) + 2\sigma F(k)\,\text{CTF}(k)]\right] \approx I_i(x) + n_i(x) + \text{DC}, \tag{S17}$$

where

$$I_i(x) = \frac{d^2}{4} N_0 \mathscr{F}^{-1}[2\sigma F(k)\,\text{CTF}(k)]. \tag{S18}$$

Equivalently, in Fourier space

$$I_i^{\text{noisy}}(k) \approx S_i(k) + n_i(k), \tag{S19}$$

where

$$S_i(k) = \frac{d^2}{4} N_0 2\sigma F(k)\,\text{CTF}(k). \tag{S20}$$

And the noise term $n_i(x)$ is additive white Gaussian noise:

$$n_i(x) \sim \mathcal{N}\left(0, \sigma_0^2\right), \tag{S21}$$

where $\sigma_0^2 = \frac{d^2}{4} N_0$. We will ignore the contribution of the DC term in the cross-correlation.

Matching with a perfect reference

We now define an SNR value as the expected value of cross-correlations generated from phase contrast, divided by the SD of the correlations from noise

$$\text{SNR}_s = \frac{CC_s}{\sigma(CC_n)}. \tag{S22}$$

For an image summed over M frames, Eq. S19 is updated with an exposure filter $Q(k, N_i)$ calculated from (19). Here, $Q(k, N_i)$ denotes the normalized exposure-weighting transfer function at spatial frequency k for frame dose N_i ($N_i = i\,N_0$). The unnormalized exposure-dependent weight is

$$q(k, N) = e^{\frac{-N}{2N_E(k)}}, \tag{S23}$$

and the normalized filter is

$$Q(k, N_i) = \frac{q(k, N_i)}{\sqrt{\sum_i q(k, N_i)^2}}. \tag{S24}$$

The processed image is then

$$I_i^{\text{proc}}(k) = Q(k, N_i)\left[S_i(k) + n_i(k)\right]. \tag{S25}$$

It follows that

$$\sum_{i=1}^{M} |Q(k, N_i)|^2 = 1, \tag{S26}$$

ensuring the noise in the summed frame remains "white" with variance σ_0^2. Additional instrument-dependent attenuation factors, such as detector DQE and temporal or spatial coherence envelope functions, were not explicitly included. These effects act as additional frequency-dependent weighting terms in Fourier space and could be incorporated into the same filter without changing the structure of the derivation.

The summed image:

$$\begin{aligned} I_{\mathrm{sum}}(k) &= \sum_{i=1}^{M} I_i^{\mathrm{proc}}(k) \\ &= \sum_{i=1}^{M} Q\left(k, N_i\right) S_i(k) + \sum_{i=1}^{M} Q\left(k, N_i\right) n_i(k) \\ &= S_{\mathrm{sum}}(k) + n_{\mathrm{sum}}(k) \end{aligned} \tag{S27}$$

where

$$S_{\mathrm{sum}}\left(k\right) = \frac{d^2}{4} N_0 2\sigma F(k)\,\mathrm{CTF}(k) \sum_{i=1}^{M} Q\left(k, N_i\right). \tag{S28}$$

Using a perfect reference that is also exposure filtered

$$R(k) = \frac{S_{\mathrm{sum}}^*(k)}{\sigma_0^2}, \tag{S29}$$

the expected CC from the signal is

$$CC_s = \int_k \frac{S_{\mathrm{sum}}(k) \cdot S_{\mathrm{sum}}^*(k)}{\sigma_0^2} dk = \frac{1}{\sigma_0^2} \int_k \left|S_{\mathrm{sum}}(k)\right|^2 dk \tag{S30}$$

For noise, the mean CC is zero and the variance is:

$$\mathrm{Var}\left(CC_n\right) = \int_k \left|\frac{S_{\mathrm{sum}}(k)}{\sigma_0^2}\right|^2 \cdot \sigma_0^2 dk = \frac{1}{\sigma_0^2} \int_k \left|S_{\mathrm{sum}}(k)\right|^2 dk \tag{S31}$$

Thus, the SNR from phase contrast, assuming Wilson statistics (70) (flat spectrum of randomly positioned carbon atoms), is:

$$\begin{aligned} \frac{CC_s}{\sqrt{\mathrm{Var}\left(CC_n\right)}} &= \sqrt{\frac{1}{\sigma_0^2} \int_k \left|S_{\mathrm{sum}}\left(k\right)\right|^2 dk} \\ &= \sqrt{\frac{1}{\sigma_0^2} \int_k \left(\frac{d^2}{4} N_0 2\sigma F(k)\,\mathrm{CTF}(k) \sum_{i=1}^{M} Q\left(k, N_i\right)\right)^2 dk} \\ &= \sqrt{\frac{d^2}{4} N_0 \int_k 4\sigma^2 |F(k)|^2 \left(\mathrm{CTF}(k) \sum_{i=1}^{M} Q\left(k, N_i\right)\right)^2 dk} \\ &= \sqrt{4 f N_0 D^2 \int_k \mathrm{CTF}(k)^2 \left(\sum_{i=1}^{M} Q\left(k, N_i\right)\right)^2 dk} \\ &= D\sqrt{4 f N_0 T} \end{aligned} \tag{S32}$$

where the integral T can be calculated numerically.

C. Interpretation when $SNR_n = SNR_s$

Eq. 1 describes an *expected* alignment-noise level from the maximum over N_s tested hypotheses, where N_s includes in-plane rotation, out-of-plane viewing directions, and *x*, *y* shifts for a per-particle 5D search. Therefore, $SNR_n = SNR_s$ should be interpreted as a threshold point in expectation, not a deterministic boundary for each particle. When SNR_s is only slightly above SNR_n, correct alignments are favored on average and real density can accumulate with sufficiently large particle numbers, but residual pose errors still attenuate high-resolution amplitudes (effectively a larger positive B-factor). In this regime, sharpening (negative-B correction) can improve visibility after averaging, but cannot recover information lost by misalignment.

D. Advanced assumptions

Inelastic scattering

We further add a correction for inelastic scattering (assuming the use of an energy filter to remove inelastically scattered electrons) using:

$$N_{\text{eff}} = N_e \cdot \exp\left(\frac{-t}{\Lambda_{\text{in}}}\right), \tag{S33}$$

where t is the sample thickness and Λ_{in} is the inelastic scattering mean free path of the solution. The inelastic mean free path depends on the incident electron energy and is related to the scattering cross-section as:

$$\Lambda_{\text{in}} = \frac{M_w}{\rho N_A \sigma_{\text{in,sol}}}, \tag{S34}$$

where $\sigma_{\text{in,sol}}$ is the weighted inelastic scattering cross-section calculated by individual atom's inelastic scattering cross-sections (71, 72):

$$\sigma_{\text{in}} = \frac{1.5 \cdot 10^{-6} Z^{\frac{1}{2}}}{\beta^2} \ln \frac{\beta^2 \left(U_0 + mc^2\right)}{10} \quad \left[\text{nm}^2\right] \tag{S35}$$

where Z is atomic number, β the ratio between velocity of the electron and light, U_0 the incident electron energy and mc^2 the rest energy of the electron.

For vitreous ice,

$$\sigma_{\text{in, sol}} = 2\sigma_{\text{in, H}} + 1\sigma_{\text{in, O}}. \tag{S36}$$

For example, at 300 keV, $\sigma_{\text{in, sol}}$ = 343 nm.

Phase plate

Based on (36), Eq. S15 can be extended to

$$\text{CTF}\left(\lambda, k, \Delta f, C_s, \Delta\phi, w\right) = -\sin\left[\chi\left(\lambda, |k|, \Delta f, C_s, \Delta\phi, w\right)\right] \tag{S37}$$

where

$$\begin{aligned}\chi\left(\lambda, k, \Delta f, C_s, \Delta\phi, w\right) = \pi\lambda k^2 \left(\Delta f - \frac{1}{2}\lambda^2 k^2 C_s\right) \\ + \Delta\phi + \arctan\left(\frac{w}{\sqrt{1-w^2}}\right).\end{aligned} \tag{S38}$$

Here, $\Delta\Phi$ is an additional phase shift introduced by a phase plate and w is the fraction of amplitude contrast (e.g., 0.07 or 0.1).

Defocus spread from particle thickness

Defocus has variation due to the particle "depth" D. We further account for this in the CTF based on (6, 37). For each depth slice z, the updated phase shift (Eq. S38) is

$$\begin{aligned}\chi^{'}(z) &= \pi\lambda k^2\left(\Delta f + z - \frac{1}{2}\lambda^2 k^2 C_s\right) + \Delta\phi + \arctan\left(\frac{w}{\sqrt{1-w^2}}\right)\\ &= \chi\left(\lambda, k, \Delta f, C_s, \Delta\phi, w\right) + \pi\lambda k^2 z\end{aligned} \tag{S39}$$

Hence, the depth-averaged CTF is

$$\begin{aligned}\mathrm{CTF}_D\left(\lambda, k, \Delta f, C_s, \Delta\phi, w\right) &= -\frac{1}{D}\int_{-D/2}^{D/2}\sin\left[\chi^{'}(z)\right]dz\\ &= -\frac{1}{D}\int_{-D/2}^{D/2}\sin\chi\left(\lambda, k, \Delta f, C_s, \Delta\phi, w\right)\cos\left(\pi\lambda k^2 z\right)\\ &\quad + \cos\chi\left(\lambda, k, \Delta f, C_s, \Delta\phi, w\right)\sin\left(\pi\lambda k^2 z\right)dz\\ &= -\frac{\sin\chi\left(\lambda, k, \Delta f, C_s, \Delta\phi, w\right)}{D}\int_{-D/2}^{D/2}\cos\left(\pi\lambda k^2 z\right)dz\\ &\quad - \frac{\cos\chi\left(\lambda, k, \Delta f, C_s, \Delta\phi, w\right)}{D}\int_{-D/2}^{D/2}\sin\left(\pi\lambda k^2 z\right)dz\\ &= -\sin\chi\left(\lambda, k, \Delta f, C_s, \Delta\phi, w\right)\cdot\frac{\sin\left(\pi\lambda k^2 D/2\right)}{\pi\lambda k^2 D/2} + 0\\ &= -\sin\chi\left(\lambda, k, \Delta f, C_s, \Delta\phi, w\right)\cdot\mathrm{sinc}\left[\frac{\pi\lambda k^2 D}{2}\right]\end{aligned} \tag{S40}$$

Liquid helium cooling

Based on the fact that electron diffraction spots fade 1.2-1.8× slower at liquid helium compared to using liquid nitrogen (24), the exposure filter function can be updated

$$q(k, N) = \exp\left(-\frac{N}{2\times 1.8 N_E(k)}\right) = \exp\left(-\frac{N}{3.6 N_E(k)}\right). \tag{S41}$$

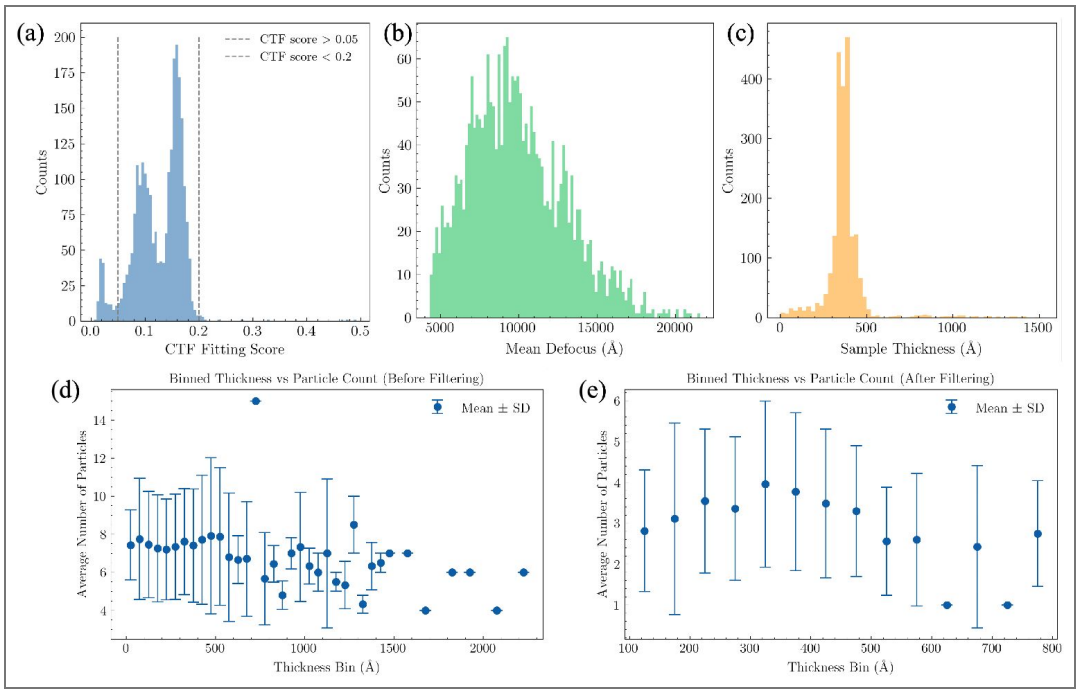


Figure 1—figure supplement 1. Image statistics of the untilted micrographs in EMPIAR-10252. (a) CTF fitting scores for 2,488 untilted images, calculated using `ctffind5`. Images with scores above 0.2 or below 0.05 were excluded from 2DTM searches. (b) Mean defocus values of the 2,314 images retained for 2DTM. (c) Sample thickness estimates from `ctffind5`, with a median thickness of 367 Å. Negative thickness estimates were excluded from the histogram. (d) Particle counts per thickness bin, based on 17,274 particles extracted from the 2,314 images using `extract-particles`. (e) Particle counts per thickness bin after particle selection using `filter-particles`, showing the final stack of 7,353 particles.

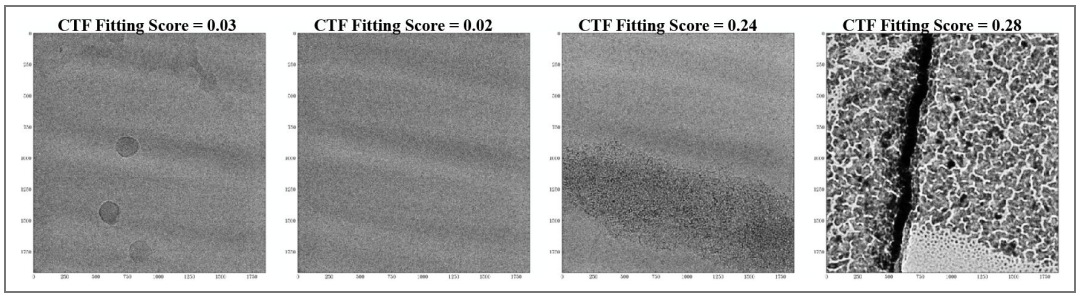


Figure 1—figure supplement 2. Examples of micrographs excluded from 2DTM search. (a) Very low contrast and ice contamination. **(b)** Extremely low contrast, likely drift or astigmatism. **(c)** Particle aggregation or contamination. **(d)** Cross-grating calibration grid.

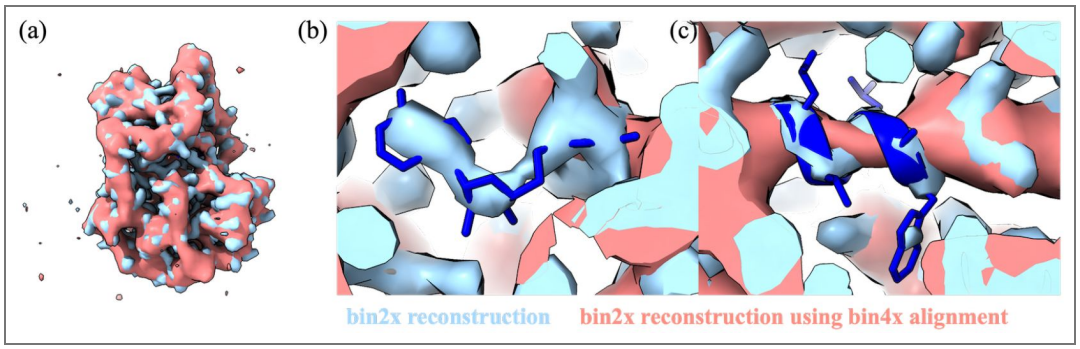


Figure 1—figure supplement 3. Effect of template resolution on reconstruction quality. Comparison of 3D reconstructions using particles aligned at bin2x (1.117 Å/pixel) versus bin4x (2.234 Å/pixel) resolution. For the bin4x experiment, 2DTM was performed on bin4x images and the detected particle coordinates were used to extract particles from the bin2x images for reconstruction. Densities at the ATP-binding site and deleted residues 222–227 are shown. The bin4x-aligned reconstruction shows loss of ATP density and degraded backbone density for residues 222–227, demonstrating that high-resolution signal in the template is critical for accurate particle alignment and recovery of omitted densities. Map contour σ = 5.

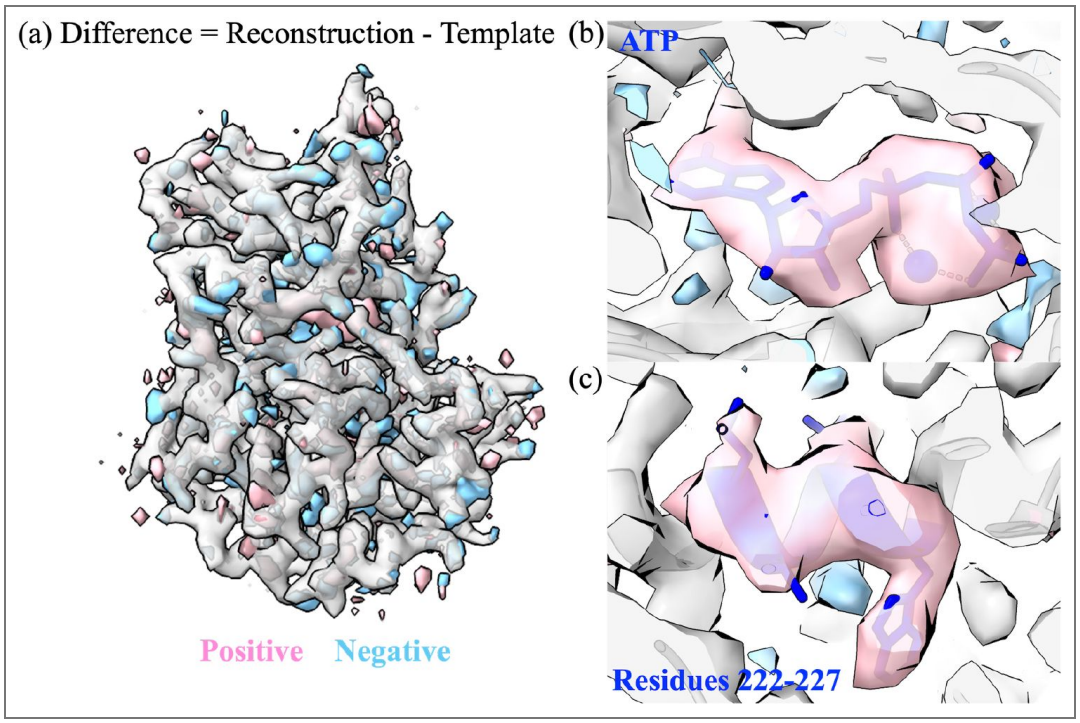


Figure 1—figure supplement 4. Difference map between the template (gray, map contour σ = 15) and the reconstruction shown in Figure 1. The difference map was generated using the `diffmap.exe` program (53) with a protein soft mask. Positive (pink) and negative (light blue) difference densities are shown at map contour σ = 20. Positive densities indicate regions present in the reconstruction but absent from the template, with coherent features at the ATP binding site and residues 222–227, which were deleted from the template. Residual noisy densities elsewhere reflect limitations of the forward model.

Group	Residue	Occupancy
Omitted residues	Trp222	0.70
	Ala223	0.75
	Leu224	0.80
	Gly225	0.55
	Val226	0.70
	Leu227	0.79
	Mean	*0.72*
Ligands	ATP	0.61
	Mn^{2+} (1)	0.28
	Mn^{2+} (2)	0.28
Control residues (in template)	Ile150	1.00
	Val151	1.00
	Leu152	1.00
	Thr153	0.89
	Phe154	1.00
	Glu155	0.88
	Mean	*0.96*

Figure 1—source data 1. Grouped occupancy refinement of omitted and control residues. Occupancies were refined using Phenix real-space refinement against the omit reconstruction (Figure 1), with the full 1ATP model docked into the map. Omitted residues (222–227), ATP, and Mn^{2+} were absent from the template used for 2DTM alignment. Control residues (150–155) were included in the template. Occupancies near 1.0 for control residues and intermediate values (0.55–0.80) for omitted residues confirm partial, unbiased recovery of the omitted-region density.

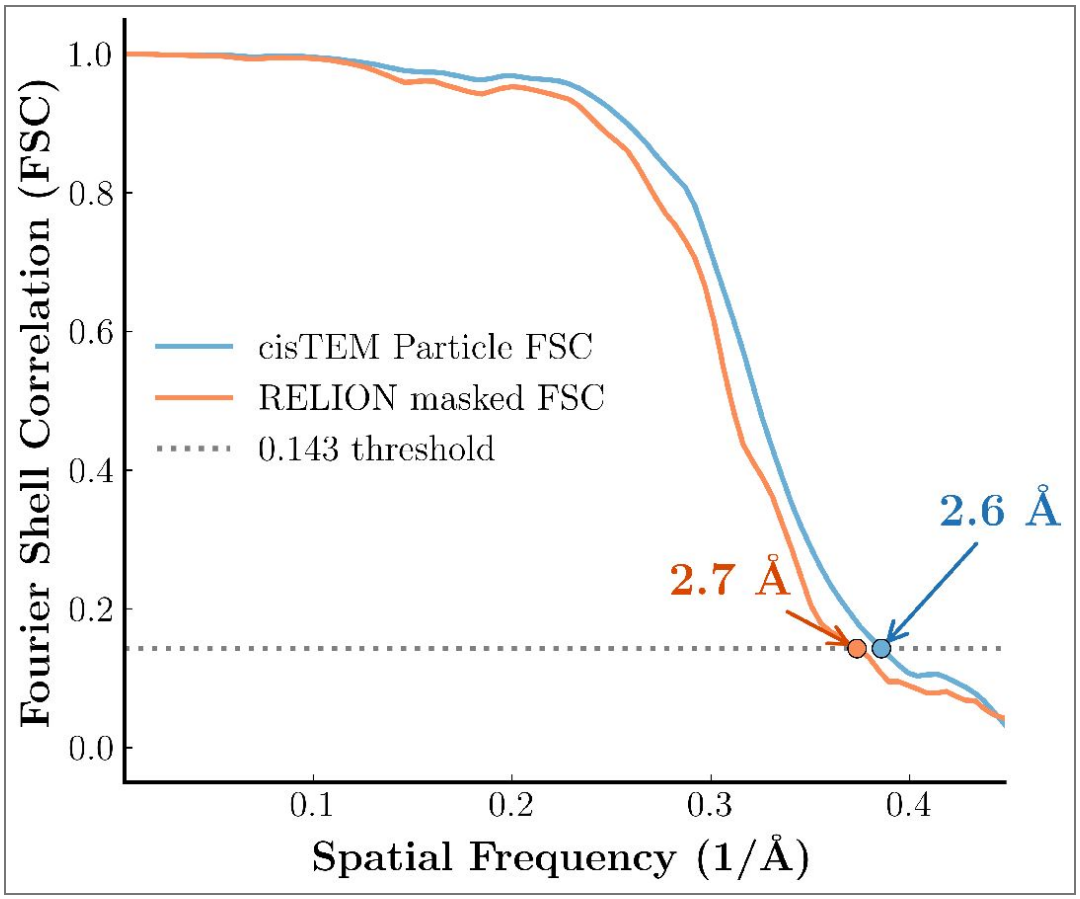


Figure 3—figure supplement 1. FSC comparison between *cis*TEM and RELION on the same half-maps. Both curves were computed from the same *cis*TEM half-maps (Figure 1 reconstruction). The *cis*TEM Particle FSC (blue) uses a spherical mask with an analytical solvent-fraction correction, while the RELION masked FSC (orange) uses a tight 3D protein mask applied directly to the half-maps. Both cross the FSC = 0.143 threshold at ~2.6–2.7 Å, confirming that the two packages give consistent resolution estimates when applied to the same data.

	p-value = 8.0	p-value = 7.0	SNR = 7.5	+ tilt images
ATP	0.60	0.58	0.54	0.57
Residues 222–227	0.57	0.56	0.45	0.54

Figure 4—source data 1. Q-scores for omitted regions across particle selection conditions (Figure 4). Values are average Q-scores over all atoms in each group, calculated using the MapQ command line tool (34, 35). The ATP Q-score is averaged over all 31 non-hydrogen atoms; the residues 222–227 Q-score is averaged over all non-hydrogen atoms in the six residues (Trp, Ala, Leu, Gly, Val, Leu).

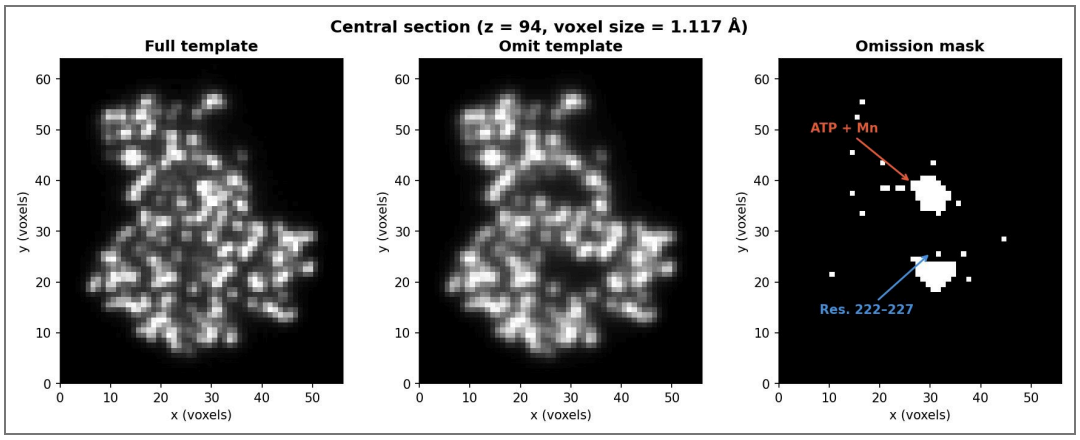


Figure 4—figure supplement 1. Central section of the omission mask used for template bias (Ω) calculation. The mask was derived from the difference between the full and omit templates (threshold = average of top-100 voxels / 5; 959 masked voxels). Left: full template. Center: omit template (residues 222–227, ATP, Mn, and H2O deleted). Right: binary omission mask, with the two omitted regions (ATP + Mn and residues 222–227) annotated. The mask isolates the omitted region for the Ω calculation shown in Figure 4. Voxel size 1.117 Å.

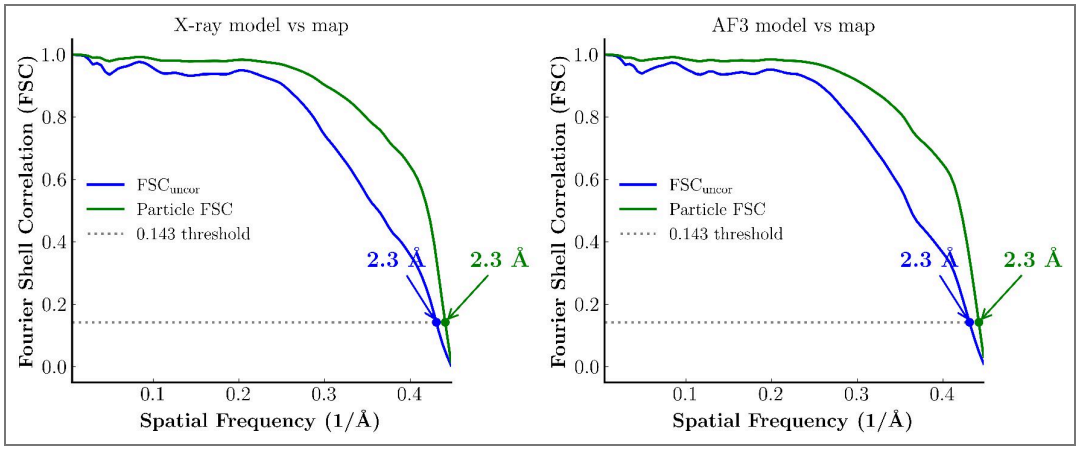


Figure 6—figure supplement 1. Map–model FSC for X-ray and AlphaFold3 templates in Figure 6. Map–model FSC was computed between each atomic model-derived template and the corresponding 2DTM reconstruction using *cis*TEM `calculate_fsc`, with a protein envelope mask. Left: X-ray-derived template versus X-ray-derived 2DTM reconstruction. Right: AlphaFold3-derived template versus AlphaFold3-derived 2DTM reconstruction. Both curves cross the FSC = 0.143 threshold at ~2.3 Å.

Additional information

Funding

Funder	Grant reference number	Author
Howard Hughes Medical Institute (HHMI)		Kexin Zhang
Howard Hughes Medical Institute (HHMI)		Nikolaus Grigorieff

Author ORCID iDs

Kexin Zhang: https://orcid.org/0000-0002-3273-2797
Nikolaus Grigorieff: https://orcid.org/0000-0002-1506-909X

References

1. **Kühlbrandt Werner** (2014) The resolution revolution. *Science* **343**:1443-1444 https://doi.org/10.1126/science.1251652 | PubMed
2. **Nakane Takanori,** Kotecha Abhay, Sente Andrija, McMullan Greg, Masiulis Simonas, Brown Patricia MGE, Grigoras Ioana T, Malinauskaite Lina, Malinauskas Tomas, Miehling Jonas, *et al.* (2020) Single-particle cryo-em at atomic resolution. *Nature* **587**:152-156 https://doi.org/10.1038/s41586-020-2829-0 | PubMed
3. **Yip Ka Man,** Fischer Niels, Paknia Elham, Chari Ashwin, Stark Holger (2020) Atomic-resolution protein structure determination by cryo-em. *Nature* **587**:157-161 https://doi.org/10.1038/s41586-020-2833-4 | PubMed
4. **Burley Stephen K,** Berman Helen M, Kleywegt Gerard J, Markley John L, Nakamura Haruki, Velankar Sameer (2017) Protein data bank (pdb): the single global macromolecular structure archive. *Protein crystallography: methods and protocols* 627-641 https://doi.org/10.1007/978-1-4939-7000-1_26 | PubMed
5. **Sigworth Fred J** (2004) Classical detection theory and the cryo-em particle selection problem. *Journal of structural biology* **145**:111-122 https://doi.org/10.1016/j.jsb.2003.10.025 | PubMed
6. **McMullan G,** Vinothkumar KR, Henderson R (2015) Thon rings from amorphous ice and implications of beam-induced brownian motion in single particle electron cryo-microscopy. *Ultramicroscopy* **158**:26-32 https://doi.org/10.1016/j.ultramic.2015.05.017 | PubMed
7. **Wentinck Koen,** Gogou Christos, Meijer Dimphna H (2022) Putting on molecular weight: Enabling cryo-em structure determination of sub-100-kda proteins. *Current Research in Structural Biology* **4**:332-337 https://doi.org/10.1016/j.crstbi.2022.09.005 | PubMed
8. **Foster Mark P,** McElroy Craig A, Amero Carlos D (2007) Solution nmr of large molecules and assemblies. *Biochemistry* **46**:331-340 https://doi.org/10.1021/bi0621314 | PubMed
9. **Puthenveetil Robbins,** Vinogradova Olga (2019) Solution nmr: A powerful tool for structural and functional studies of membrane proteins in reconstituted environments. *Journal of Biological Chemistry* **294**:15914-15931 https://doi.org/10.1074/jbc.rev119.009178 | PubMed
10. **Perilla Juan R,** Zhao Gongpu, Lu Manman, Ning Jiying, Hou Guangjin, Byeon In-Ja L, Gronenborn Angela M, Polenova Tatyana, Zhang Peijun (2017) Cryoem structure refinement by integrating nmr chemical shifts with molecular dynamics simulations. *The Journal of Physical Chemistry B* **121**:3853-3863 https://doi.org/10.1021/acs.jpcb.6b13105 | PubMed
11. **Wu Shenping,** Avila-Sakar Agustin, Kim JungMin, Booth David S, Greenberg Charles H, Rossi Andrea, Liao Maofu, Li Xueming, Alian Akram, Griner Sarah L, *et al.* (2012) Fabs enable single particle cryoem studies of small proteins. *Structure* **20**:582-592 https://doi.org/10.1016/j.str.2012.02.017 | PubMed
12. **Wu Xudong,** Rapoport Tom A (2021) Cryo-em structure determination of small proteins by nanobody-binding scaffolds (legobodies). *Proceedings of the National Academy of Sciences* **118**:e2115001118 https://doi.org/10.1073/pnas.2115001118 | PubMed
13. **Uchański Tomasz,** Masiulis Simonas, Fischer Baptiste, Kalichuk Valentina, López-Sánchez Uriel, Zarkadas Eleftherios, Weckener Miriam, Sente Andrija, Ward Philip, Wohlkönig Alexandre, *et al.* (2021) Megabodies expand the nanobody toolkit for protein structure determination by single-particle cryo-em. *Nature methods* **18**:60-68 https://doi.org/10.1038/s41592-020-01001-6 | PubMed
14. **Lu Xin,** Yan Ming, Cai Yang, Song Xi, Chen Huan, Du Mengtan, Wang Zhenyi, Li Jia'an, Niu Liwen, Zeng Fuxing, *et al.* (2025) A large, general and modular darpin–apoferritin scaffold enables the visualization of small proteins by cryo-em. *IUCrJ* **12** https://doi.org/10.1107/s2052252525003021 | PubMed
15. **Bonilla Steve L,** Jang Karen (2024) Challenges, advances, and opportunities in rna structural biology by cryo-em. *Current Opinion in Structural Biology* **88**:102894 https://doi.org/10.1016/j.sbi.2024.102894 | PubMed

16. **Henderson Richard** (1995) The potential and limitations of neutrons, electrons and x-rays for atomic resolution microscopy of unstained biological molecules. *Quarterly reviews of biophysics* **28**:171-193 https://doi.org/10.1017/s003358350000305x | PubMed

17. **Glaeser Robert M** (1999) Electron crystallography: present excitement, a nod to the past,1 anticipating the future. *Journal of structural biology* **128**:3-14 https://doi.org/10.1006/jsbi.1999.4172 | PubMed

18. **McMullan G,** Faruqi AR, Clare D, Henderson R (2014) Comparison of optimal performance at 300 kev of three direct electron detectors for use in low dose electron microscopy1. *Ultramicroscopy* **147**:156-163 https://doi.org/10.1016/j.ultramic.2014.08.002 | PubMed

19. **Grant Timothy,** Grigorieff Nikolaus (2015) Measuring the optimal exposure for single particle cryo-em using a 2.6 å reconstruction of rotavirus vp6. *eLife* **4**:e06980 https://doi.org/10.7554/eLife.06980 | PubMed

20. **Schwartz Osip,** Axelrod Jeremy J, Campbell Sara L, Turnbaugh Carter, Glaeser Robert M, Müller Holger (2019) Laser phase plate for transmission electron microscopy. *Nature methods* **1 16**:1016-1020 https://doi.org/10.1038/s41592-019-0552-2 | PubMed

21. **Remis Jonathan,** Petrov Petar N, Zhang Jessie T, Axelrod Jeremy J, Cheng Hang, Sandhaus Shahar, Mueller Holger, Glaeser Robert M (2024) Cryo-em phase-plate images reveal unexpected levels of apparent specimen damage. *Journal of Structural Biology* **216**:1 https://doi.org/10.1016/j.jsb.2024.108150 | PubMed

22. **Axelrod Jeremy J,** Petrov Petar N, Remis Jonathan, Glaeser Robert M, Mueller Holger (2023) Overcoming resolution-loss mechanisms in laser phase plate cryo-electron microscopy1. *Biophysical Journal* **122**:316a https://doi.org/10.1016/j.bpj.2022.11.1775

23. **Axelrod Jeremy J,** Zhang Jessie T, Petrov Petar N, Glaeser Robert M, Müller Holger (2024) Modern approaches to improving phase contrast electron microscopy. *Current Opinion in Structural Biology* **86**:102805 https://doi.org/10.1016/j.sbi.2024.102805 | PubMed

24. **Naydenova Katerina,** Kamegawa Akiko, Peet Mathew J, Henderson Richard, Fujiyoshi Yoshinori, Russo Christopher J (2022) On the reduction in the effects of radiation damage to two-dimensional crystals of organic and biological molecules at liquid-helium temperature1. *Ultramicroscopy* **237**:113512 https://doi.org/10.1016/j.ultramic.2022.113512 | PubMed

25. **Dickerson Joshua L,** Naydenova Katerina, Peet Mathew J, Wilson Hugh, Nandy Biplob, McMullan Greg, Morrison Robert, Russo Christopher J (2025) Reducing the effects of radiation damage in cryo-em using liquid helium temperatures. *Proceedings of the National Academy of Sciences* **122**:e2421538122 https://doi.org/10.1073/pnas.2421538122 | PubMed

26. **Rickgauer J Peter,** Grigorieff Nikolaus, Denk Winfried (2017) Single-protein detection in crowded molecular environments in cryo-em images. *eLife* **6**:e25648 https://doi.org/10.7554/eLife.25648 | PubMed

27. **Lucas Bronwyn A,** Himes Benjamin A, Xue Liang, Grant Timothy, Mahamid Julia, Grigorieff Nikolaus (2021) Locating macromolecular assemblies in cells by 2d template matching with cistem. *eLife* **10**:e68946 https://doi.org/10.7554/eLife.68946 | PubMed

28. **Lucas Bronwyn A,** Zhang Kexin, Loerch Sarah, Grigorieff Nikolaus (2022) In situ single particle classification reveals distinct 60s maturation intermediates in cells. *eLife* **11**:e79272 https://doi.org/10.7554/eLife.79272 | PubMed

29. **Lucas Bronwyn A,** Himes Benjamin A, Grigorieff Nikolaus (2023) Baited reconstruction with 2d template matching for high-resolution structure determination in vitro and in vivo without template bias. *eLife* **12**:RP90486 https://doi.org/10.7554/eLife.90486 | PubMed

30. **Jr Mark A Herzik,** Wu Mengyu, Lander Gabriel C (2019) High-resolution structure determination of sub-100 kda complexes using conventional cryo-em. *Nature communications* **10**:1032 https://doi.org/10.1038/s41467-019-08991-8 | PubMed

31. **Zheng J,** Trafny Elzbieta A, Knighton Daniel R, Xuong N-H, Taylor Susan S, Ten Eyck Lynn F, Sowadski Janusz M (1993) 2.2 å refined crystal structure of the catalytic subunit of camp-dependent protein kinase complexed with mnatp and a peptide inhibitor. *Biological Crystallography* **49**:362-365 https://doi.org/10.1107/s0907444993000423 | PubMed

32. **Himes Benjamin,** Grigorieff Nikolaus (2021) Cryo-tem simulations of amorphous radiation-1 sensitive samples using multislice wave propagation. *IUCrJ* **8**:943-953 https://doi.org/10.1107/s2052252521008538 | PubMed

33. **Zhang Kexin,** Cossio Pilar, Rangan Aaditya V, Lucas Bronwyn A, Grigorieff Nikolaus (2025) A new statistical metric for robust target detection in cryo-em using 2d template matching1. *IUCrJ* **12** https://doi.org/10.1107/s2052252524011771 | PubMed

34. **Pintilie Grigore,** Zhang Kaiming, Su Zhaoming, Li Shanshan, Schmid Michael F, Chiu Wah (2020) Measurement of atom resolvability in cryo-em maps with q-scores. *Nature methods,1* **17**:328-334 https://doi.org/10.1038/s41592-020-0731-1 | PubMed

35. **Pintilie Grigore,** Shao Chenghua, Wang Zhe, Hudson Brian P, Flatt Justin W, Schmid Michael F, Morris Kyle, Burley Stephen K, Chiu Wah (2025) Q-score as a reliability measure for protein, nucleic acid, and small molecule atomic coordinate models derived from 3dem density maps. *bioRxiv* https://doi.org/10.1101/2025.01.14.633006 | PubMed

36. **Rohou Alexis,** Grigorieff Nikolaus (2015) Ctffind4: Fast and accurate defocus estimation from electron micrographs. *Journal of structural biology* **192**:216-221 https://doi.org/10.1016/j.jsb.2015.08.008 | PubMed

37. **Elferich Johannes,** Kong Lingli, Zottig Ximena, Grigorieff Nikolaus (2024) Ctffind5 provides improved insight into quality, tilt, and thickness of tem samples. *eLife* **13**:RP97227 https://doi.org/10.7554/eLife.97227 | PubMed

38. **Zhu Jianying,** Zhang Qi, Zhang Hui, Shi Zuoqiang, Hu Mingxu, Bao Chenglong (2023) A minority of final stacks yields superior amplitude in single-particle cryo-em. *Nature Communications* **114**:7822 https://doi.org/10.1038/s41467-023-43555-x | PubMed

39. **Abramson Josh,** Adler Jonas, Dunger Jack, Evans Richard, Green Tim, Pritzel Alexander, Ronneberger Olaf, Willmore Lindsay, Ballard Andrew J, Bambrick Joshua, *et al.* (2024) Accurate structure prediction of biomolecular interactions with alphafold 3. *Nature* **630**:493-500 https://doi.org/10.1038/s41586-024-07487-w | PubMed

40. **Lander Gabriel C,** Glaeser Robert M (2021) Conquer by cryo-em without physically dividing1. *Biochemical Society Transactions* **49**:2287-2298 https://doi.org/10.1042/bst20210360 | PubMed

41. **Zhang Yue,** Tammaro R, Peters Peter J, Ravelli RBG (2020) Could egg white lysozyme be solved by single particle cryo-em?. *Journal of chemical information and modeling* **60**:2605-2613 https://doi.org/10.1021/acs.jcim.9b01176 | PubMed

42. **Kimanius Dari,** Jamali Kiarash, Wilkinson Max E, Lövestam Sofia, Velazhahan Vaithish, Nakane Takanori, Scheres Sjors HW (2024) Data-driven regularization lowers the size barrier of cryo-em structure determination. *Nature Methods* **21**:1216-1221 https://doi.org/10.1038/s41592-024-02304-8 | PubMed

43. **Zhang Kaiming,** Li Shanshan, Kappel Kalli, Pintilie Grigore, Su Zhaoming, Mou Tung-Chung, Schmid Michael F, Das Rhiju, Chiu Wah (2019) Cryo-em structure of a 40 kda sam-iv riboswitch rna at 3.7 å resolution. *Nature communications* **10**:5511 https://doi.org/10.1038/s41467-019-13494-7 | PubMed

44. **Scheres Sjors HW** (2014) Beam-induced motion correction for sub-megadalton cryo-em particles. *eLife* **3**:e03665 https://doi.org/10.7554/eLife.03665 | PubMed

45. **Russo Christopher J,** Passmore Lori A (2016) Progress towards an optimal specimen support for electron cryomicroscopy. *Current opinion in structural biology* **37**:81-89 https://doi.org/10.1016/j.sbi.2015.12.007 | PubMed

46. **Zivanov Jasenko,** Nakane Takanori, Scheres Sjors HW (2019) A bayesian approach to beam-induced motion correction in cryo-em single-particle analysis. *IUCrJ* **6**:5-17 https://doi.org/10.1107/s205225251801463x | PubMed

47. **Wu Mengyu,** Lander Gabriel C (2020) How low can we go? structure determination of small biological complexes using single-particle cryo-em. *Current opinion in structural biology* **64**:9-16 https://doi.org/10.1016/j.sbi.2020.05.007 | PubMed

48. **Peet Mathew J,** Henderson Richard, Russo Christopher J (2019) The energy dependence of contrast and damage in electron cryomicroscopy of biological molecules. *Ultramicroscopy* **203**:125-131 https://doi.org/10.1016/j.ultramic.2019.02.007 | PubMed

49. **Haider M,** Hartel P, Müller H, Uhlemann S (2010) Information transfer in a tem corrected for spherical and chromatic aberration. *Microscopy and Microanalysis* **16**:393-408 https://doi.org/10.1017/s1431927610013498 | PubMed

50. **Hosokawa Fumio,** Sawada Hidetaka, Kondo Yukihito, Takayanagi Kunio, Suenaga Kazutomo (2013) Development of cs and cc correctors for transmission electron microscopy. *Microscopy* **62**:23-41 https://doi.org/10.1093/jmicro/dfs134 | PubMed

51. **Linck Martin,** Müller Heiko, Hartel Peter, Perl Svenja, Uhlemann Stephan, Haider Max (2020) On the benefit of aberration correction in cryo electron microscopy. *Microscopy and Microanalysis* **26**:2156-2157 https://doi.org/10.1017/s1431927620020644

52. **Chua Eugene YD,** Mendez Joshua H, Rapp Micah, Ilca Serban L, Tan Yong Zi, Maruthi Kashyap, Kuang Huihui, Zimanyi Christina M, Cheng Anchi, Eng Edward T, *et al.* (2022) Better, faster, cheaper: recent advances in cryo–electron microscopy. *Annual review of biochemistry* **91**:1-32 https://doi.org/10.1146/annurev-biochem-032620-110705 | PubMed

53. **Grigorieff Nikolaus** (2021) diffmap.exe. Grigorieff Lab. https://grigoriefflab.umassmed.edu/diffmap

54. **Shtyrov Alexander,** Wilson Hugh, Slowik Daria, Yamashita Keitaro, Li Jade, Wojdyr Marcin, Chen Shaoxia, McMullan Greg, Short Jude, Russo Christopher J, *et al.* (2025) Measurement of atomic scattering factors by cryo-electron microscopy. *bioRxiv* 2025.10.24.683059 https://doi.org/10.1101/2025.10.24.683059

55. **Piper Sarah J,** Johnson Rachel M, Wootten Denise, Sexton Patrick M (2022) Membranes under the magnetic lens: a dive into the diverse world of membrane protein structures using cryo-em. *Chemical reviews* **122**:13989-14017 https://doi.org/10.1021/acs.chemrev.1c00837 | PubMed

56. **Fan Xiao,** Wang Jia, Zhang Xing, Yang Zi, Zhang Jin-Can, Zhao Lingyun, Peng Hai-Lin, Lei Jianlin, Wang Hong-Wei (2019) Single particle cryo-em reconstruction of 52 kda streptavidin at 3.2 angstrom resolution. *Nature communications* **10**:2386 https://doi.org/10.1038/s41467-019-10368-w | PubMed

57. **Han Yimo,** Fan Xiao, Wang Haozhe, Zhao Fang, Tully Christopher G, Kong Jing, Yao Nan, Yan Nieng (2020) High-yield monolayer graphene grids for near-atomic resolution cryoelectron microscopy. *Proceedings of the National Academy of Sciences* **117**:1009-1014 https://doi.org/10.1073/pnas.1919114117 | PubMed

58. **Kim Kookjoo,** Li Huan, Clarke Oliver B (2025) High-resolution ab initio reconstruction enables cryo-em structure determination of small particles. *bioRxiv* 2025.09.08.674935 https://doi.org/10.1101/2025.09.08.674935

59. **Cushing Victoria I,** Koh Adrian F, Feng Junjie, Jurgaityte Kaste, Bondke Alexander, Kroll Sebastian HB, Barbazanges Marion, Scheiper Bodo, Bahl Ash K, Barrett Anthony GM, *et al.* (2024) High-resolution cryo-em of the human cdk-activating kinase for structure-based drug design. *Nature Communications* **15**:2265 https://doi.org/10.1038/s41467-024-46375-9 | PubMed

60. **Wei Jia,** Leit Silvana, Kuai Jun, Therrien Eric, Rafi Salma, Harwood H James, DeLaBarre Byron, Tong Liang (2019) An allosteric mechanism for potent inhibition of human atp-citrate lyase. *Nature* **568**:566-570 https://doi.org/10.1038/s41586-019-1094-6 | PubMed

61. **Zheng Shawn Q,** Palovcak Eugene, Armache Jean-Paul, Verba Kliment A, Cheng Yifan, Agard David A (2017) Motioncor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nature methods* **14**:331-332 https://doi.org/10.1038/nmeth.4193 | PubMed

62. **Pettersen Eric F,** Goddard Thomas D, Huang Conrad C, Meng Elaine C, Couch Gregory S, Croll Tristan I, Morris John H, Ferrin Thomas E (2021) Ucsf chimerax: Structure visualization for researchers, educators, and developers. *Protein science* **30**:70-82 https://doi.org/10.1002/pro.3943 | PubMed

63. **Grant Timothy,** Rohou Alexis, Grigorieff Nikolaus (2018) cis tem, user-friendly software for single-particle image processing. *eLife* **7**:e35383 https://doi.org/10.7554/eLife.35383 | PubMed

64. **Grigorieff Nikolaus** (2000) Resolution measurement in structures derived from single particles. *Acta Crystallographica Section D: Biological Crystallography* **56**:1270-1277 https://doi.org/10.1107/s0907444900009549 | PubMed

65. **Peng L-M,** Ren Gongxizi, Dudarev SL, Whelan MJ (1996) Robust parameterization of elastic and absorptive electron atomic scattering factors. *Acta Crystallographica Section A: Foundations of Crystallography* **52**:257-276 https://doi.org/10.1107/s0108767395014371

66. **Peng L-M** (1999) Electron atomic scattering factors and scattering potentials of crystals. *Micron* **30**:625-648 https://doi.org/10.1016/s0968-4328(99)00033-5

67. **Prince Edward** (2004) *International Tables for Crystallography, Volume C: Mathematical, physical and chemical tables* Springer Science & Business Media.

68. **Koeck PJB,** Karshikoff A (2015) Limitations of the linear and the projection approximations in three-dimensional transmission electron microscopy of fully hydrated proteins. *Journal of Microscopy* **259**:197-209 https://doi.org/10.1111/jmi.12253 | PubMed

69. **Mindell Joseph A,** Grigorieff Nikolaus (2003) Accurate determination of local defocus and specimen tilt in electron microscopy. *Journal of structural biology* **142**:334-347 https://doi.org/10.1016/s1047-8477(03)00069-8 | PubMed

70. **Rosenthal Peter B,** Henderson Richard (2003) Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. *Journal of molecular biology* **333**:721-745 https://doi.org/10.1016/j.jmb.2003.07.013 | PubMed

71. **Langmore John P,** Smith Michael F (1992) Quantitative energy-filtered electron microscopy of biological molecules in ice. *Ultramicroscopy* **46**:349-373 https://doi.org/10.1016/0304-3991(92)90024-e | PubMed

72. **Langmore JP** (1973) The collection of scattered electrons in dark-field electron microscopy. *Optik* **38**:335-350

Herzik Jr MA, Wu M, Lander GC (2019) Catalytic subunit of protein kinase A bound to ATP, manganese, and IP20 movies obtained using Talos Arctica operating at 200 kV equipped with a K2. EMPIAR. ID EMPIAR-10252 https://www.ebi.ac.uk/empiar/EMPIAR-10252/

Peer reviews

Reviewer #1 (Public review):

Summary:

This paper describes an application of the high-resolution cryo-EM 2D template matching technique to sub-50kDa complexes. The paper describes how density for ligands can be reconstructed without having to process cryo-EM data through the conventional single particle analysis pipelines.

Strengths:

Improved insights in which particles contribute to the density of ligands that is absent from the templates are valuable.

Weaknesses:

Although the convenient visualisation of small molecules bound to protein targets of a known structure would be relevant for the pharmaceutical industry, the evidence described for the claim that this technique "significantly" improves alignment of reconstruction of small complexes is incomplete. In a revised paper, the authors are encouraged to better evaluate the effects of model bias on the reconstructed densities.

In the revised version, the refinement of atomic occupancies in the 2DTM-generated maps has been insightful: densities only come back at values ranging from 0.55-0.80, whereas residues included in the template remain at 1, suggesting that the 2DTM-reconstruction does suffer from model bias. Their newly added Omega calculations, which are helpful, also suggest that model bias is present in the 2DTM-based reconstructions. These observations therefore contradict the first subsection heading of the Results, which claims "unbiased reconstruction of omitted residues".

Both the Omega analysis and the refined atomic occupancies provide insights into the "real-space aspect" of the model bias. The question to what extent the model bias affects the map in Fourier space remains unanswered. The authors base some of their claim in the paper on FSC curves in Figures 1b and 3b, but these will suffer from the same model bias. To assess this, I had requested the authors to reconstruct an OMIT map and to assess its resolution using FSCs. The authors have indeed performed a careful reconstruction of an OMIT map, which is currently shown in Figure 5. I liked how they implemented this, as described in detail in the Methods section. However, the measurement of how much model bias is present in this OMIT map by FSC calculations is still pending. This could be done in two ways, and I would encourage the authors to present the results of both in (hopefully a last) revised version of their manuscript. My original suggestion was to calculate a map-to-model FSC for the OMIT map and the full reference. This should be compared with a similar map-to-model FSC on the map where only the ligand was omitted. Alternatively, they can use the cisTEM FSC_uncorr procedure on the OMIT half-reconstructions and compare the resulting curve with the one presented in Figure 1b.

The reason that I am keen to see these FSCs is because high-resolution model bias is a fundamental danger of the 2DTM approach. It will therefore also be in the interest of the authors to quantify the extent to which it happens. For now, I have kept the above public review and short assessment the same as they were, but I will consider raising the assessment after the suggested experiments (which I hope will be relatively easy to do!) are incorporated.

https://doi.org/10.7554/eLife.109790.2.sa2

Reviewer #3 (Public review):

Summary:

Due to the low SNR of cryo-EM micrographs necessitated by radiation damage, determining the structure of proteins smaller than 50 kDa is exceedingly challenging, such that only a handful have been solved to date. This work aims to improve the reconstruction of small proteins in single-particle cryo-EM by using high-resolution 2D template matching, an algorithm previously used to locate and align macromolecules in situ, to align and reconstruct small proteins. This approach uses an existing macromolecular structure, either experimentally determined or predicted by AlphaFold, to simulate a noise-free 3D reference and generates whitened projections, crucially including high-spatial-frequency information, to align particles by the orientation with maximal cross-correlation. They demonstrate the success of this approach by generating a 3D reconstruction from an existing dataset of a 41.3 kDa protein kinase that had previously evaded attempts at high-resolution structure determination. To alleviate concerns that this is purely from template bias, they demonstrate

clear density at two regions that were not present in the template: 6 residues in an alpha helix and an ATP in the ligand binding pocket. The latter is particularly important for its implications in determining structures of ligand-bound proteins for drug discovery. They also produce a composite omit map from 36 partial-deletion reconstructions spanning the entire protein, demonstrating a reconstruction can be obtained without template bias. Additionally, the authors provide an update to the classic calculation in Henderson 1995 to predict the minimum molecular mass of a protein that can be solved by single-particle cryo-EM.

Strengths:

I am in no doubt that this technique can be used to gain valuable insights into the structures of small proteins, and this is an important advancement for the field. It is complementary to single-particle cryo-EM and provides an extra tool for the experimentalist that may work better in certain cases. For cases where only a small region of the structure is of interest, such as in drug screening, this method provides a simple workflow to screen many structures.

The claim that using high-spatial frequency information is essential for aligning small proteins is a valuable insight. A recent pre-print published at a similar time to this manuscript used high-resolution information in standard ab-initio reconstruction to generate a high-resolution reconstruction from the same dataset, supporting the claims made in the manuscript.

The theoretical section outlined in the appendix is also theoretically sound. It uses the same logic as Henderson, but applies more up-to-date knowledge, such as incorporating dose-weighting and altering the cross-correlation based noise estimation. This update is valuable for understanding factors preventing us from reaching the theoretical limit.

Weaknesses:

The applicability of this technique to more than a single target was not demonstrated. Nor was it compared to more recent strategies for processing SPA data from small molecules, such as Blush regularization or HR-HAIR. Additionally, although the authors have demonstrated convincingly that their method selects a stack of high-quality particles, it is less clear whether it performs better than RELION when using the same stack of particles, particularly in the ATP binding pocket. This places this method as a complementary technique, and whether it outperforms those methods for a wide variety of molecules is yet to be determined. The method presented here also introduces template bias, so only parts of the reconstruction not in the initial template are free of template bias. Producing a full reconstruction through a composite omit map is computationally expensive, meaning that unless this method outperforms modern SPA methods, its major use case will be ligand binding studies instead of 3D reconstructions.

https://doi.org/10.7554/eLife.109790.2.sa1

Author response:

The following is the authors' response to the original reviews.

> ***eLife Assessment***
>
> *This important study builds on previous work from the same authors to present a conceptually distinct workflow for cryo-EM reconstruction that uses 2D template matching to enable highresolution structure determination of small (sub-50 kDa) protein targets. The paper describes how density for small-molecule ligands bound to such targets can be reconstructed without these ligands being present in the template. However, the evidence described for the claim that this technique "significantly" improves*

> *the alignment of the reconstruction of small complexes is incomplete. The authors could better evaluate the effects of model bias on the reconstructed densities.*

We have addressed both concerns. Regarding the claim that 2DTM "significantly" improves alignment, the most direct evidence is the controlled comparison in Fig. 3: using the same particle stack and the same reconstruction software (RELION), 2DTM-derived orientations yield a 3.1 Å reconstruction whereas RELION auto-refinement of the same particles yields 3.7 Å. Because the orientations are the only variable, this comparison directly demonstrates that 2DTM produces more accurate alignments.

We further evaluated RELION auto-refinement with initial low-pass filters of 3, 5, 10, and 15 Å (Fig. 3c); the final resolution remained between 3.7 and 4.0 Å across all conditions, indicating that the achievable resolution difference reflects a fundamental distinction between the two approaches. 2DTM directly leverages high-resolution signal in the template during alignment, which is particularly advantageous for small particles.

To assess whether this improvement extends beyond the ligand pocket, we constructed a composite omit map (Fig. 5) assembled from 36 reconstructions, each generated using a template with a different subset of residues deleted. The composite shows that density can be recovered at distributed locations across the kinase, including peripheral and surface-exposed regions further away from the alignment center. Recovery varies across sites, with some regions exhibiting weaker or fragmented density, consistent with local differences in structural heterogeneity and residual alignment error. Together, these results indicate that the orientation estimates support global density recovery rather than being confined to the ligand-binding region.

Regarding model bias, we have strengthened both the quantitative and visual analyses. Specifically, we have (i) updated the template-bias metric Ω in Fig. 4, (ii) added grouped occupancy refinement showing that omitted residues 222–227 refine to 0.55–0.80 (mean 0.72), ATP to 0.61, and Mn to 0.28, while template-included control residues 150–155 remain near 1.0 (0.88–1.00; mean 0.96), and (iii) completed the composite omit map described above. Together, these results provide consistent evidence that densities corresponding to omitted regions are not driven by the template and can be recovered from the data, while template-included regions show some, albeit limited evidence of overfitting, as expected.

> ***Reviewer #1 (Public review):***
>
> *Summary:*
>
> *This paper describes an application of the high-resolution cryo-EM 2D template matching technique to sub-50kDa complexes. The paper describes how density for ligands can be reconstructed without having to process cryo-EM data through the conventional single particle analysis pipelines.*
>
> *Strengths:*
>
> *This paper contributes additional data (alongside other papers by the same authors) to convey the message that high-resolution 2D template matching is a powerful alternative for cryo-EM structure determination. The described application to ligand density reconstruction, without the need for extensive refinements, will be of interest to the pharmaceutical industry, where often multiple structures of the same protein in complex with different ligands are solved as part of their drug development pipelines. Improved insights into which particles contribute to the best ligand density are also highly valuable and transferable to other applications of the same technique.*
>
> *Weaknesses:*

> *Although the convenient visualisation of small molecules bound to protein targets of a known structure would be relevant for the pharmaceutical industry, the evidence described for the claim that this technique "significantly" improves alignment of reconstruction of small complexes is incomplete. The authors are encouraged to better evaluate the effects of model bias on the reconstructed densities in a revised paper.*

We thank the reviewer for these constructive comments. We have updated the template-bias metric Ω in Fig. 4 and added two further quantitative controls: grouped occupancy refinement of omitted residues and a composite omit map spanning the entire protein. Full details are provided in our responses to Comments 1 and 2 below.

> ***Reviewer #1 (Recommendations for the authors):***
>
> *Main Comments*
>
> *(1) For the 1ATP structure: Q-scores for deleted residues/ligands are worse than the Q-scores for residues in the template. This means that the reconstructed map must suffer from template bias. Another indication of this bias is that the density for the ATP (and the omitted residues) appears to be weaker than the density for the residues in the template (although this is not easy to assess from the figures). The authors should perform additional experiments to quantify this bias.*
>
> *(a) One option could be to do what the X-ray crystallographers call an OMIT map, and omit allresidues, a few at a time, from the template in multiple 2DTM runs. They could then assemble a density map from all the omitted residues together and measure the resolution of the omit map against the known template by FSC.*
>
> *(b) Another insightful experiment would be to take the various 2DTM reconstructed maps describedin the paper and perform a refinement of the atom occupancies of all residues in the structure. Residues included in the template should refine to values close to 1. In the absence of bias, the occupancies of the omitted residues should be 1 too; if the reconstructed map were completely biased, those occupancies would refine to 0. Therefore, the refined occupancies of omitted residues could perhaps serve as a measure for the amount of bias in the reconstructed map.*

We thank the reviewer for these detailed and constructive suggestions. We agree that the lower Q-scores for omitted regions indicate weaker density and that template bias exists at residues that are included in the template. To quantify this more directly, we corrected the template-bias metrics at the omitted region (mask from the full–omit template difference) in Fig. 4.

Following the reviewer's suggestion, we performed Phenix real-space grouped occupancy refinement against the omit reconstruction using the docked full model. The results are shown in Table. S2. We refined occupancies for the omitted residues (chain E 222–227), ATP, Mn, and template-included control residues (chain E 150–155), while excluding waters. The omitted residues refined to occupancies of 0.55–0.80 (mean 0.72), ATP to 0.61, and Mn to 0.28, whereas the control residues remained near 1.0 (0.88–1.00; mean 0.96). These results indicate substantial recovery of density in the omitted regions, but also some degree of bias.

The substantially lower refined occupancy of Mn^{2+} may reflect genuine partial occupancy in the dataset. While compact features can be especially sensitive to residual alignment error, we cannot conclude from the present analysis that alignment effects alone account for the weak Mn^{2+} density.

Finally, we have constructed a composite omit map to assess density recovery across the protein. We generated 36 omit templates, each deleting ~10 non-overlapping residues

scattered across the structure (including peripheral and surface-exposed regions). For each template, an independent 2DTM search and reconstruction was performed. Local density patches were extracted within 3 Å of the omitted atoms (with neighboring residues excluded as described in Methods) and assembled into a composite map (Fig. 5). The composite map shows that density can be recovered at distributed locations across the protein and is not restricted to the central binding pocket. Recovery is variable across sites, with some regions exhibiting weaker or fragmented density, consistent with local differences in signal-to-noise, structural heterogeneity, and residual alignment error.

> *(2) The claim that 2DTM leads to "Improved" reconstruction (title) and "alignment and reconstruction [...] can be significantly improved" (abstract) is not supported by the data presented in the paper. The smallest single particle structure to resolutions sufficient for de novo atomic modelling is currently the ACA2 complex, with an ordered mass of less than 40 kDa, which was reconstructed using Blush regularisation in RELION. This paper should be referenced, and statements about single particle analysis (SPA) not working for sub-50 kDa complexes should be toned down. In general, I would say that 2DTM and SPA are not competing techniques, and the paper would be better if it focused on the intrinsic advantages of 2DTM (like ease-of-use for screening of pharmaceutical compounds) and useful findings described that make 2DTM better, e.g., excluding thick ice.*

We thank the reviewer for this important perspective and have added the Blush regularization reference Kimanius et al. (2024) to the revised manuscript, noting that the 40 kDa Aca2–RNA complex was reconstructed to 2.5 Å resolution using this approach (at L451). Furthermore, Blush regularization could be applied to reconstructions derived from 2DTM-based particle stacks, and a combination of both approaches may yield further improvements.

We agree that 2DTM and SPA are complementary rather than competing techniques and have revised the manuscript to reflect this. We have also toned down claims in the abstract, which now states that 2DTM "reconstructed a previously intractable ~43 kDa kinase complex and improved the density of its ligand-binding site" rather than making broad claims about SPA limitations. In the discussion, we now describe 2DTM as broadening possibilities for structural studies of targets "that have remained difficult to reconstruct" rather than implying they are impossible by SPA.

Regarding the intrinsic advantages of 2DTM: beyond ligand screening, the composite omit map (Fig. 5, described in Comment 1) demonstrates that 2DTM-derived orientations support density recovery throughout the entire protein, including peripheral and surface-exposed residues, using roughly an order of magnitude fewer particles than conventional SPA workflows.

> *(3) Given the uncertainties about the amount of template bias in the reconstructed 2DTM densities, I have trouble interpreting the predictions in Table 1. Where would the 1ATP structure lie in Figure 8? How much bias would there be in a 2DTM reconstruction at SNR n = SNR s? Could the authors perform tests on simulated data to confirm these predictions? At the point of SNR n = SNR s, how would a 2DTM reconstruction look, and what would refined occupancies for deleted residues be?*
>
> *(This may reflect a misunderstanding on my part, but I don't really see how the SNR n = SNR s is completely dependent on the number of orientations searched (through Equation 1). In Figure 8, is the full search in a 4k x 4k micrograph, or inside a particle box? And what are the relevant search ranges? Perhaps as a consequence of this misunderstanding, I do not understand how one would decide on the amount of noise in the simulated data for these tests.)*

We thank the reviewer for this important question and agree that this point needed clearer explanation. In our framework, $\mathrm{SNR}_n = \sqrt{2 \log N_s}$ is the expected alignment-noise level from maximizing many cross correlations, where N_s is the total number of sampled hypotheses in the 5D search (in-plane angle, out of-plane angles, and *x*, *y* shifts), not only the number of orientations. Thus, the relevant search is the per-particle alignment search window (full or constrained), not a full 4k×4k micrograph area.

At $\mathrm{SNR}_n = \mathrm{SNR}_s$, the true-match and noise-maxima levels are at a threshold; one could imagine if SNR_s is only slightly larger than SNR_n, the correct pose is favored on average, so with sufficiently large particle numbers real omitted-region density should accumulate, but with residual pose errors that attenuate high-frequency amplitudes (effectively a large positive B-factor). In that regime, sharpening (negative-B correction) can improve visibility once signal is accumulated. Therefore, we expect partial recovery rather than fully unbiased recovery at this threshold, with omitted-region occupancies remaining between 0 and 1 and below template-included controls (consistent with our measured values), and improving as $\mathrm{SNR}_s - \mathrm{SNR}_n$ and particle number increase. Simulations at this exact threshold would require a very large particle number to achieve sufficient statistics, and we leave this to future work. We have added this clarification to the Supporting Information.

> *(4) The strong (> 5 sigma!!) and ubiquitous difference densities in Figure 9A imply that the authors have a serious problem with their forward model, which could explain some of the effects of model bias discussed above. I recommend they investigate these differences in detail. It would be good to see negative and positive densities in different colours to understand these differences better. The text speaks about incomplete capture of the solvent background, but the difference densities appear to be of much higher spatial frequencies than those typical for background/solvent effects (e.g., 15-20A). It may thus also be helpful to analyse these differences in Fourier space.*

We thank the reviewer for this important point. In our previous analysis, we did not incorporate an appropriate protein mask when generating the difference map, which contributed to widespread residual densities. We have now regenerated the map using the program diffmap.exe (https: //*grigoriefflab.umassmed.edu/diffmap*) with a protein soft mask and moved it to the Supplementary Information (Fig. Figure 1—figure supplement 4, contour SD = 20). With this controlled setup, the strongest coherent residual densities localize to the omitted ATP pocket and residues 222–227, consistent with recovery of omitted features. We have revised the figure/text accordingly and clarified that remaining diffuse residuals are likely due to forward-model mismatch (including solvent/background representation). We also added to the manuscript that improved template generation may be achieved by incorporating recent methods that learn environment-aware scattering factors directly from experimental cryo-EM maps.

> *Other Comments*
>
> *(1) P.1: Alongside reference 2, a reference to the 1.2 Å apoferritin structure from the Stark group should be included.*

We have added the reference at L30.

> *(2) P.2: "commond line tool"*

We have corrected the typo.

> *(3) P.2-3: Robust reconstruction of the ATP binding pocket: Auto-refinements in RELION without alignments do not exist, and corresponding statements need to be removed from the manuscript. If one wants to skip alignments, then there is no refinement left to*

> *be done. In that case, one should just perform a reconstruction of the 2 halves (e.g., using relion reconstruct) and then run a standard RELION postprocessing.*

We agree with the reviewer and have revised the manuscript accordingly. Technically, RELION's relion refine with the --skip align flag runs an iterative loop that re-estimates the per-particle noise model (spectral noise σ^2) and computes the gold-standard FSC between half-maps, but it does not modify the particle orientations or translations. As the reviewer correctly points out, this is effectively a 3D reconstruction followed by postprocessing, not a refinement. We have updated the text to replace "skip-alignment auto-refinement" with "3D reconstruction without angular refinement" to accurately reflect what was performed.

> *(4) P.3: What are "first-quadrant p-values" and "three-quadrant p-values"?*

We apologize for the ambiguity and now define these terms explicitly in the revised text (with citation to the p-value paper). After transforming z-score and SNR to probit coordinates, "first-quadrant" (1Q) p-values use only candidate points with both coordinates > 0 (i.e., both probit-zscore and probitSNR are positive). "Three-quadrant" (3Q) p-values include candidates where at least one coordinate is > 0 (equivalently, all points except the quadrant where both are < 0).

> *(5) P.5: In Equation (2), it is unclear what Q means from the main text. Would it be better to leave Equation (2) for the Appendix, and only show Equation (3) in the main text?*

Thank you for this suggestion. We kept Equation (2) in the main text to preserve the continuity of the derivation, but we now define $Q(k,N_i)$ explicitly at first use as the normalized exposure-weighting transfer function (following Grant 2015). The detailed derivation and assumptions remain in the Supporting Information.

> *(6) P.6: "Remaining gaps": this section considers differences between 200 keV and 300 keV electron beam energies. The main practical effect for cryo-EM data sets is that the current detectors are designed for detecting 300 keV electrons, and their DQE is thus a lot worse at 200 keV. The entire paper doesn't mention detectors. Perhaps because they are assumed to be perfect, but it is still far from the case.*
>
> *Also, why were defocus searches not performed if the thickness of micrographs was up to 1500 A?*
>
> *The conclusion of this section states "Considering all these factors…", but it then claims standard single particle analysis still remains an outstanding challenge. This concluding statement makes no sense, as this whole section was about 2DTM.*

Thank you for this comment. We agree and have revised the text to make these points explicit. First, we now state clearly that detector response (DQE) is generally more favorable at 300 keV than at 200 keV, which contributes to the experimental–theoretical gap. Second, we clarify why we did not perform a defocus search in 2DTM: after CTF/thickness filtering, the retained micrographs are predominantly in the thin-ice regime, so expected defocus spread is smaller, while adding a defocus dimension substantially increases computational cost. We also tested downstream refinement (including CTF/beam-tilt related refinement in *cis*TEM) and did not observe measurable improvement for this dataset (data not included in the manuscript). Finally, we revised the concluding sentence in this subsection to refer specifically to 2DTM-based alignment limits rather than standard SPA, so the section scope is now consistent.

> *(7) P.7: Data-driven refinement of AlphaFold3 models: it might be worth pointing out that removing residues a few at a time from AF3 models and checking their reconstructed density by 2DTM would come at a considerable computational cost.*

We agree. We have demonstrated residue-level omission validation using the X-ray template via a composite omit map (Fig. 5), confirming that the approach is feasible. We have updated the Discussion to reflect this: extending the composite omit approach to AlphaFold3-based templates remains computationally expensive — each omission design requires an independent 2DTM search and downstream reconstruction — and we present this as an important direction for future work.

> *(8) Figure 1: What is "full FSC" and what is "particle FSC"?*

Thank you for pointing this out. We have clarified the terminology in the figure legend and text using *cis*TEM and Frealign definitions (Grant et al., 2018). What was previously labeled "Full FSC" is now referred to as the uncorrected FSC (FSC_{uncor}), computed within a generous mask. "Particle FSC" denotes the solvent-corrected FSC, obtained from FSC_{uncor} using the mask-volume correction factor f as described in the *cis*TEM/Frealign framework (Grant et al., 2018).

> *(9) Figure 3: Why were particles in class 5 discarded? The 2DTM approaches described in this paper are all about carefully selecting good particles, yet now the authors use standard 3D classification to throw away another 156 particles. This seems to be an arbitrary choice. How different would the results have been if these had been included in the reconstruction? Alternatively, did these few particles have any 2DTM metrics that would justify their exclusion?*

We thank the reviewer for raising this point. Class 5 contained only 156 particles (~2% of the dataset). While the 2DTM p-value and SNR metrics provide principled criteria for particle selection, they are not perfect, and a small number of suboptimal particles may still pass these filters. To address the reviewer's concern, we repeated the reconstruction including all five classes. The resulting map achieved a resolution of 3.7 Å, identical to the reconstruction without class 5, confirming that including these particles does not affect the results. We have clarified this point in the manuscript.

> *(10) Figure 4C: What are the negative sample thicknesses here? Why use an inset?*

The negative sample thickness values are artifacts of the CTF-based thickness estimation algorithm in ctffind5. This algorithm fits oscillations in the 1-D power spectrum arising from the interaction between the CTF and the specimen's finite thickness (a sinc-modulated envelope). When the ice is very thin or the power spectrum is noisy, the optimizer can converge to a physically meaningless negative value. Of the 2,488 total micrographs across both sessions (after CTF score filtering, 2,314 retained), 136 (~5.9%) returned negative thickness estimates. We have revised Figure 1—figure supplement 1c (previously Figure 4c) to show only the physically meaningful positive thickness values without the inset, which gives a clearer view of the unimodal distribution peaked near 350–400 Å.

> ***Reviewer #2 (Public review):***
>
> *Summary:*
>
> *In this manuscript, Zhang et al describe a method for cryo-EM reconstruction of small (sub50kDa) complexes using 2D template matching. This presents an alternative, complementary path for high-resolution structure determination when there is a prior atomic model for alignment. Importantly, regions of the atomic model can be deleted to avoid bias in reconstructing the structure of these regions, serving as an important mechanism of validation.*
>
> *The manuscript focuses its analysis on a recently published dataset of the 40kDa kinase complex deposited to EMPIAR. The original processing workflow produced a medium*

> *resolution structure of the kinase (GSFSC ~4.3 Å, though features of the map indicate ~6-7 Å resolution); at this resolution, the binding pocket and ligand were not resolved in the original published map. With 2DTM, the authors produce a much higher resolution structure, showing clear density for the ATP binding pocket and the bound ATP molecule. With careful curation of the particle images using statistically derived 2DTM p-values, a high-resolution 2DTM structure was reconstructed from just 8k particles (2.6 Å non-gold standard FSC; ligand Q-score of 0.6), in contrast to the 74k particles from the original publication. This aligns with recent trends that fewer, higher-quality particles can produce a higher-quality structure. The authors perform a detailed analysis of some of the design choices of the method (e.g., p-value cutoff for particle filtering; how large a region of the template to delete).*
>
> *Overall, the workflow is a conceptually elegant alternative to the traditional bottom-up reconstruction pipeline. The authors demonstrate that the p-values from 2DTM correlations provide a principled way to filter/curate which particle images to extract, and the results are impressive. There are only a few minor recommendations that I could make for improvement.*

We appreciate the positive assessment. In response to the bias-related concerns raised elsewhere, we have: (i) updated the template-bias metric Ω reported in Fig. 4, (ii) added grouped occupancy refinement showing that omitted residues 222–227 refine to a mean occupancy of 0.72 while template-included control residues remain near 1.0, and (iii) assembled a composite omit map (Fig. 5) from 36 partial-deletion reconstructions spanning the entire protein. These additions are described in the revised Results and in the rebuttal below.

> ***Reviewer #2 (Recommendations for the authors):***
>
> *(1) On page 3, "Finally, by comparing Figure 2a and b, we observed that deleting IP20 strongly reduced signal at several residues." Looking at Figure 2a and 2b, it was unclear which residues they were referring to.*

We have revised the text to explicitly list the affected residues. In the updated Figure 2, we now label the omitted residues with the lowest backbone Q-scores in the structural views (column 2) and include per-residue backbone Q-score plots (column 4), making the comparison between panels (a) and (b) quantitative. For example, when IP20 is additionally deleted (Fig. 2b), residues Phe54, Gly55, Lys72, Glu127, Glu170, and Asp184 all fall below a backbone Q-score of 0.5, compared with only Ser53 and Glu127 in the within-3 Å deletion alone (Fig. 2a).

> *(2) Figure 1a. Both the published density map and the text "Template" are gray, but the 2DTM template density map is yellow.*

Thank you for catching this inconsistency. We have updated Figure 1a so that the 2DTM template density is now rendered in gray, consistent with the X-ray crystal structure (PDB) coloring. The published single-particle map is shown in wheat and the 2DTM reconstruction in blue, providing a clear three-way color distinction.

> *(3) Figure 1b. I would recommend the x-axis label of "spatial frequency" instead of "resolution" (which is overloaded). Furthermore, the fact that this is not a GSFSC should be clearly labeled in the figure to prevent confusion with a standard GSFSC.*

We agree with both suggestions. The x-axis has been relabeled "Spatial Frequency (1/Å)" in the revised figure. We have also added a note in the figure caption stating that these FSC curves are not gold-standard FSCs, as the reconstruction uses orientations determined by template matching rather than independent half-set refinement.

> *(4) Figure 2: The usage of the negative sign in the labels "-3 Å", "-5 Å" to indicate within a given radius is a bit confusing. "Within 3 Å", perhaps?*

Thank you for this suggestion. We have changed the labels in Figure 2 from "−3 Å" and "−5.5 Å" to "Within 3 Å" and "Within 5.5 Å." We have also added a fourth column to Figure 2 showing per-residue backbone Q-scores for each deletion experiment, with omitted residues distinguished by color and marker shape. The residues with the lowest backbone Q-scores among the omitted set are circled in red and correspond to the labeled residues in the structural views.

> *(5) Figure 4c: Why does the sample thickness histogram go to negative values (-20,000 A)?*

As noted in our response to Reviewer 1, the negative thickness values are artifacts of the ctffind5 thickness estimation, which fits a sinc-modulated envelope to the 1-D power spectrum. For micrographs with very thin ice or noisy power spectra, the fit can converge to unphysical negative values. These account for ∼5.9% of micrographs. We have revised Figure 1—figure supplement 1 (originally Fig. 4c) to display only positive thickness values, removing the inset and providing a clearer histogram.

> *(6) Figured 4d: Should the label be "(Before Filtering)" instead of After?*

Yes, thank you for catching this. The original Figure 4d was mislabeled—it showed particle counts before filtering but was titled "After Filtering." We have corrected the labels: Figure 1—figure supplement 1d (originally Fig. 4d) now reads "Before Filtering" and Figure 1—figure supplement 1e (originally Fig. 4e) reads "After Filtering."

> *(7) Supplementary Note 1: Please provide units for d, p, D, and k max in equation S4 and the preceding text.*

We have added units to the text preceding Eq. S4: $d = 1/k_{\max}$ is the high-resolution alignment limit (Å), $k_{\max}$ is the maximum spatial frequency (Å $^{-1}$), $p = d/2$ is the ideal pixel size (Å/pixel), and D is the particle diameter (Å).

> *(8) What does the map-model FSC look like with the template as the model vs. the AF3 structure as the model?*

We have computed the map–model FSC for both the X-ray crystallographic template (PDB 1ATP) and the AlphaFold3-predicted template against their respective 2DTM reconstructions (Fig. Figure 6—figure supplement 1). Both curves cross the FSC = 0.143 threshold at ∼2.3 Å. We note that the map–model FSC in this context should be interpreted with caution, because the vast majority of the structure lies outside the omitted region and is present in the template, so template bias in those regions will dominate the map–model FSC and obscure differences in the small omitted region.

> ***Reviewer #3 (Public review):***
>
> *Summary:*
>
> *Due to the low SNR of cryo-EM micrographs necessitated by radiation damage, determining the structure of proteins smaller than 50 kDa is exceedingly challenging, such that only a handful have been solved to date. This work aims to improve the reconstruction of small proteins in single-particle cryo-EM by using high-resolution 2D template matching, an algorithm previously used to locate and align macromolecules in situ, to align and reconstruct small proteins. This approach uses an existing macromolecular structure, either experimentally determined or predicted by AlphaFold, to simulate a noise-free 3D reference and generates whitened projections, crucially including high-spatial-frequency information, to align particles by the orientation with*

maximal cross-correlation. They demonstrate the success of this approach by generating a 3D reconstruction from an existing dataset of a 41.3 kDa protein kinase that had previously evaded attempts at high-resolution structure determination. To alleviate concerns that this is purely from template bias, they demonstrate clear density at two regions that were not present in the template: 6 residues in an alpha helix and an ATP in the ligand binding pocket. The latter is particularly important for its implications in determining structures of ligand-bound proteins for drug discovery. Additionally, the authors provide an update to the classic calculation in Henderson 1995 to predict the minimum molecular mass of a protein that can be solved by single-particle cryo-EM.

Strengths:

I am in no doubt that this technique can be used to gain valuable insights into the structures of small proteins, and this is an important advancement for the field. The ability to determine the structure of ligands in a binding site is particularly important, and this paper provides a method of doing that which outperforms traditional single-particle cryo-EM processing workflows.

The claim that using high-spatial frequency information is essential for aligning small proteins is a valuable insight. A recent pre-print published at a similar time to this manuscript used high-resolution information in standard ab-initio reconstruction to generate a high-resolution reconstruction from the same dataset, supporting the claims made in the manuscript.

The theoretical section outlined in the appendix is also theoretically sound. It uses the same logic as Henderson, but applies more up-to-date knowledge, such as incorporating dose-weighting and altering the cross-correlation-based noise estimation. This update is valuable for understanding factors preventing us from reaching the theoretical limit.

Weaknesses:

Given that this technique creates template bias, only parts of the reconstruction not in the template can be trusted, unlike standard single-particle processing, where the independent half-maps from separate, ab initio templates are used to generate a 3D reconstruction. Although, in principle, one could perform the search many times such that every residue has been omitted in at least one search, this will be extremely computationally intensive and was not demonstrated in this manuscript. It is therefore currently only realistically applicable when only a small portion of the sub-50 kDa protein is of interest.

The applicability of this technique to more than a single target was also not demonstrated, and there are concerns that it may not work effectively in many cases. The authors note in the results that "the ATP density was consistently recovered more robustly than nearby residues" and speculate that this may be because misalignments disproportionately blur peripheral residues. Since the region of interest in a structure is not necessarily in the center, this may need further investigation. The implications of this statement may also be unclear to the reader. For example, can this issue be minimized by having the region of interest centered in the simulated volume?

In Figure 3, the authors demonstrate that it is not solely improved particle filtering and a noise-free reference that improves alignment, but that the high spatial frequency information is important. This information is very valuable since it can be applied to other, more standard methods. However, this key figure is not as clear or convincing as it could be. The FSC curves are possibly misleading, since the reduced resolution could be explained by reduced template bias when auto-refining with a map initially low-pass filtered to 10 A. Moreover, although the helix reconstruction does look slightly better

> *using the 2DTM angles, the improvement in density for ATP in the binding pocket is not clear. A qualitative argument only clear in one out of two cases is not as convincing as a quantitative metric across more examples.*

We address these concerns in three ways: (i) we quantify template bias using Phenix real-space grouped occupancy refinement: omitted residues 222–227 refine to occupancies of 0.55–0.80 (mean 0.72) and ATP to 0.61, while template-included control residues 150–155 remain near 1.0 (mean 0.96), confirming that recovered density is genuine rather than a template artifact; (ii) we have now completed a composite omit-map experiment (Fig. 5), in which 36 partial-deletion templates, each omitting ~10 non-overlapping residues, were used to perform independent 2DTM searches and reconstructions; local density patches from all 36 reconstructions were assembled into a composite map showing density recovery at distributed locations across the protein, including peripheral and surface-exposed regions, although recovery is variable across sites; and (iii) we have expanded the discussion to clarify that, while the primary scope of this work is omitted-region validation for the ligand-binding site, the composite omit-map result demonstrates that the approach generalizes beyond the central pocket.

> ***Reviewer #3 (Recommendations for the authors):***
>
> *In addition to the comments on the public review, I have some more specific suggestions that could improve the manuscript.*
>
> *(1) Another recent pre-print posted on BioRxiv shortly before this manuscript (Kim et al. Highresolution ab initio reconstruction enables cryo-EM structure determination of small particles) determined a high-resolution structure of the same protein from the same dataset, as well as determining the structures of other small proteins. Since both manuscripts rely on high-spatial frequency information, I think that the paper strengthens the claims in this manuscript and should be cited.*

We thank the reviewer for this suggestion. We agree that the recent preprint by Kim et al. strengthens the relevance of high-spatial-frequency information for small-particle cryo-EM reconstruction. We have now added this work to the revised manuscript and included a brief discussion comparing its ab initio strategy with our 2DTM-based approach.

> *(2) The claim in the abstract that "we were able to reconstruct previously intractable targets under 50 kDa and improve the density of the ligand-binding sites in the reconstructions" should be altered to make it clear that this is only a single previously intractable target.*

We agree. The revised abstract now reads ". . . we reconstructed a previously intractable ~43 kDa kinase complex and improved the density of its ligand-binding site" making clear that a single target is demonstrated in this work.

> *(3) Q-scores in the manuscript were sometimes used to quantify the improvement in map to model fit for the ATP binding pocket, but never for the 6 residues of the alpha helix. They were also not reported in every case for the ATP-binding pocket. This could lead a reader to think it is only being reported when the Q-score matches the expectation. For transparency, I would suggest either using Q-scores in every comparison or in no cases and simply relying on the qualitative result.*

We agree with the reviewer. In the revised manuscript, we now report Q-scores consistently for both ATP and residues 222–227 across all conditions: individual residue Q-scores for the omitted residues 222–227 in Fig. 1 are reported in the main text and figure caption; per-residue backbone Q-score plots for all deletion experiments in Fig. 2 are shown as the fourth column of each panel; Fig. 3 (RELION reconstruction) does not include Q-scores as the focus is

on orientation accuracy rather than map-model fit; and average Q-scores for all four particle selection conditions in Fig. 4 are listed in Figure 4—source data 1.

> *(4) The sigma values used for viewing the maps should also be stated in several figures, particularly Figure 3 and Figure 6.*

We have added contour levels (σ) to the captions of Fig. 3 and Fig. 4 (originally Fig. 6) in the revised manuscript.

> *(5) I have a slight concern about how well this method applies away from the region centered in the alignment. If parts on the periphery of the structure are removed, do these also reconstruct? Is it required that the omitted region be centered in the simulation of the 3D volume for each alignment? If so, this should be clearly stated.*

2DTM determines particle orientations by matching the full projected template to the image, so alignment is driven by the global structure rather than a localized region. As a result, the recovered orientations define the reconstruction throughout the entire particle, not only near the center. The omitted region does not need to be centered in the template volume. Any region of the protein can be omitted and its density evaluated after reconstruction.

To directly test whether peripheral regions are recovered in the same manner as central ones, we performed a composite omit-map experiment. We generated 36 omit templates, each deleting ~10 non-overlapping residues distributed across the entire protein, including peripheral and surface-exposed regions. For each template, an independent 2DTM search and reconstruction was performed. Local density patches corresponding to the omitted regions were then extracted and assembled into a composite map (Fig. 5). The resulting map shows density at distributed locations across the protein, indicating that density recovery is not restricted to regions near the alignment center and that peripheral regions can be reconstructed under the same alignment framework, although the quality of recovery varies across sites.

> *(6) I was confused by the difference between the FSCs in Figure 1 and Figure 3. I understand Figure 1 is from cisTEM and Figure 3 from RELION, but I expected the unmasked FSC and full FSC to be similar. Do the authors have any insights into why there is such a large difference? I would also consider removing the FSCs in Figure 3, since the reduced resolution may only be due to reduced template bias, meaning including this may be misleading.*

Thank you for raising this point. The apparent discrepancy arises from multiple differences between the two figures: different FSC definitions, different half-maps (reconstructed with different software and slightly different particle sets), and different masks.

In *cis*TEM (Fig. 1), two FSC curves are reported: the uncorrected FSC (FSC_{uncor}), measured within a spherical mask, and the "Particle FSC", which applies an analytical solvent-fraction correction (Grant et al., 2018) to account for solvent dilution within the mask. The Particle FSC crossed the 0.143 threshold at ~2.6 Å, whereas FSC_{uncor} crossed at ~3.0 Å. In Fig. 3, RELION postprocess applied phase-randomization correction with a soft mask, yielding ~3.1 Å. However, the Fig. 3 FSC was computed on different half-maps (RELION skip-alignment reconstruction of 7,197 particles after 3D classification) with a different mask.

To directly compare the two packages, we computed the FSC on the same *cis*TEM half-maps using both methods (Figure 3—figure supplement 1). The *cis*TEM Particle FSC (spherical mask + solvent correction) gave ~2.6 Å, while RELION image handler with a tight 3D protein mask gave ~2.7 Å. These two approaches converge to a similar resolution through different mechanisms: *cis*TEM compensates for a generous spherical mask using the solvent-fraction correction, while RELION uses a tight mask that excludes most solvent directly. This confirms that when the same half-maps are used, the two packages give consistent results and the

apparent discrepancy between Figs. 1 and 3 is primarily due to differences in the reconstruction and particle set, not the FSC calculation.

We agree with the reviewer that the FSC values in Figure 3 should be interpreted with caution. In this case, the particle orientations are not independently refined but are instead inherited from the 2DTM alignment, so the two half-maps are not strictly independent. We have added clarifying language in the revised manuscript to make this point explicit (Fig. 1 caption).

> *(7) I would also like to see how RELION auto-refinement performs with different low-pass filtering. This could strengthen the argument that high-resolution information is necessary from the start to successfully align small particles.*

We thank the constructive suggestion from the reviewer. We performed RELION auto-refinement on the same 7,197-particle stack using different initial low-pass filter resolutions (--ini high) of 3, 5, 10, and 15 Å. The resulting post-processed resolutions were:

ini_high (Å)	Post-processed resolution (Å)
3	4.0
5	3.7
10	3.7
15	4.0

Author response table 1.

The results show that varying the initial low-pass filter has minimal effect on the final resolution. This is expected because RELION uses a gold-standard, maximum-likelihood framework in which the resolution used for alignment is determined iteratively from the data via a probability distribution, rather than being fixed by the initial reference. After the first iteration, the reference is updated from the data, and higher-resolution information is incorporated only to the extent supported by the definition of the current reconstruction. Consequently, differences in the initial low-pass filter have limited impact on the final refinement outcome.

This behavior contrasts with 2DTM, where alignment is performed by direct cross-correlation against a fixed template. In this case, high-resolution features in the template contribute directly to the scoring function and can improve alignment accuracy.

To directly test the importance of high-resolution information for 2DTM alignment, we performed an additional experiment in which 2DTM was run on bin4x images (2.234 Å/pixel), and the detected particle coordinates were used to extract particles from the corresponding bin2x images (1.117 Å/pixel) for reconstruction. Despite using the same bin2x images for reconstruction, the bin4x-aligned particles yielded a map in which ATP density was lost and backbone density for residues 222–227 was visibly degraded compared to the bin2x-aligned reconstruction (Fig. Figure 1—figure supplement 3). This demonstrates that access to high-spatial-frequency information during template matching is critical for accurate alignment of small particles.

> *(8) The caption in Figure 3 should be more descriptive about what is being shown in each panel.*

We have substantially expanded the Figure 3 caption. It now describes each panel explicitly: (a) 3D classification results with particle counts, percentages, and per-class resolutions; (b) side-by-side comparison of reconstructions using 2DTM orientations versus RELION auto-refine, including full maps, zoomed binding-pocket views with the atomic model overlaid,

orientation distributions, and FSC curves with reported resolutions; and (c) a table of RELION auto-refinement resolution as a function of the initial low-pass filter setting. We also added a new panel (c) showing that including all five classes yields the same 3.7 Å resolution, addressing the concern about Class 5 exclusion.

> *(9) Figures 4 and 5 may be better suited as supplementary figures.*

We agree. Figures 4 and 5 have been moved to the Supplementary Information in the revised manuscript.

> *(10) In Figure 4c, it is difficult to understand why the thickness distribution plot goes negative, especially to such a high magnitude as 1.5 microns.*

We agree this was confusing. The negative values are fitting artifacts from ctffind5's thickness estimation, which fits a sinc-modulated envelope to the power spectrum. When the ice is very thin or the spectrum is noisy, the optimizer can converge to unphysical negative values (affecting ~5.9% of micrographs). We have revised Figure 1—figure supplement 1c (previously Figure 4c) to show only positive thickness values, which now clearly displays the unimodal distribution peaked at 350–400 Å.

> *(11) In Figure 5d, the micrograph looks a lot like a cross-grating grid used for calibration instead of crystalline ice or a fractured film.*

We agree. We have updated the caption for Figure 1—figure supplement 2d (originally Figure 5d) to read "Cross-grating calibration grid"

> *(12) Figure 6 was very surprising to me if I am interpreting it correctly. It is not stated in the caption what omega is, but I am assuming it is a measurement of template bias. It is very surprising that the template bias drops when using more particles by reducing the p-value from 8.0 to 7.0. This goes against what I understood from Lucas et al. 2023, so I am curious as to why this is the case.*

We thank the reviewer for this question and apologize for the unclear presentation. We have revised Fig. 4 (previously Figure 6) and its caption to define Ω explicitly and updated the Ω values. We also identified that the mask used in the original computation was too loose; the revised mask is now constrained to the omitted region only (ATP, Mn^{2+}, and residues 222–227), derived from the difference between the full and omit templates and shown in Figure 4—figure supplement 1. Ω is adapted from the template-bias metric introduced in (Lucas et al., 2023) and measures how much of the density in the omitted region is attributable to using the full template rather than the omit template. Specifically, for each particle selection condition we reconstruct two maps using orientations and particles derived from independent 2DTM searches with the full and omit templates (V_{full} and V_{omit}, respectively). Ω is the fractional reduction in density within the omission mask: $\Omega = (\sum_{\mathrm{mask}} V_{\mathrm{full}} - \sum_{\mathrm{mask}} V_{\mathrm{omit}}) / \sum_{\mathrm{mask}} V_{\mathrm{full}}$. In the revised Fig. 4, Ω increases from 46% (p-value = 8.0) to 48% (p-value = 7.0), consistent with the expectation that including more, lower-quality particles increases the relative contribution of the template to the reconstruction. The Ω values are 48% for the SNR = 7.5 and 53% for the tilt conditions.

> *(13) It would be useful if the in-house Python script used to calculate template bias could be made publicly available.*

We agree. The template-bias calculation (measure-template-bias) is now included in the publicly available Python package at *https://github.com/kekexinz/2DTM_postprocess_tool*, and can also be accessed in the official *cis*TEM repository at *https://github.com/timothygrant80/cisTEM*. The package also contains the extract-particles and filter-particles tools described in the Methods section.

> *(14) The p-value used is said to be a three-quadrant p-value instead of a one-quadrant p-value. Although I assume this is simply replacing an 'and' statement with an 'or' statement, the exact difference could be made clearer to the reader.*

We have now defined these terms explicitly in the revised Methods. After probit transformation of z-score and SNR, the first-quadrant (1Q) p-value requires both values to be > 0 (logical AND), whereas the three-quadrant (3Q) p-value requires at least one to be > 0 (logical OR). The 3Q criterion is therefore looser, retaining more candidates—which is beneficial for small targets that may score well on one metric but not both.

> *(15) I was, perhaps naively, surprised that z-scores could not be used. It was my understanding that by removing the rotationally invariant component from the cross-correlation, the z-score would down-weight low-resolution information compared to the cross-correlation. Given that the manuscript suggests low-resolution alignment can cause getting stuck in local minima, this is surprising to me. The authors note it led to the rejection of most particles; were there simply too many false positives when a lower threshold was used?*

The reviewer is correct that subtracting the angular mean removes the rotationally invariant component of the cross-correlation. However, the resulting z-score primarily measures how strongly a specific orientation stands out relative to other orientations. In other words, it reflects the orientation discriminability (closely related to Fisher information) rather than the absolute correlation strength. For small particles the cross correlation often varies only weakly across orientations, so $CC_{\max} - CC_{\text{avg}}$ remains small even when the absolute correlation is significant. As a result, using the z-score alone as a selection criterion led to the rejection of many true particles.

> *Theoretical Section Improvements*
>
> *(a) The discussion on beam-induced motion could be improved by separating it into initial motion (e.g., cryo-crinkling, buckling) that can be eliminated through grid design, and pseudo-Brownian motion, which cannot. Pseudo-Brownian motion will become much more significant for small proteins (based on reference 5, for a 10 kDa protein, this would be a MSD of ~0.1 A 2/e−/A 2, or a B-factor of over 2 A 2/e−/A 2), and Bayesian Polishing is unlikely to correct this perfectly, given that it imposes a smoothness of motion between nearby particles. The impact of not correcting for this could be quantified more explicitly.*

We thank the reviewer for this helpful suggestion. As noted, pseudo-Brownian motion of particles within irradiated ice introduces stochastic displacements that accumulate with dose and are expected to be more significant for small particles. Based on the analysis in (Mcmullan et al., 2015), and scaling with particle size, this effect can be aproximated as a dose-dependent mean-squared displacement (MSD) of ~0.1 $Å^2$ per ($e^-/Å^2$) for a ~10 kDa particle. Over a typical total exposure of 40–60 $e^-/Å^2$, this corresponds to an accumulated RMS displacement of ~2–2.5 Å, sufficient to attenuate high-resolution signal.

In practice, such motion acts as an additional high-frequency attenuation in Fourier space, analogous to an envelope function, reducing the coherent signal available for template matching. While Bayesian polishing can partially correct beam-induced motion, it assumes spatially smooth trajectories between nearby particles and therefore may not fully compensate for stochastic, particle-specific motion.

Within the theoretical framework presented here, this effect can be interpreted as an additional frequency-dependent damping of the signal (B-factor). Its primary consequence would be to reduce the effective signal-to-noise ratio at high spatial frequencies and therefore shift the detectable molecular-weight limit somewhat upward, without altering the structure

of the derivation. We have added text in the manuscript to clarify this point and to indicate the expected magnitude of this effect.

> *(b) The inclusion of inelastic scattering assumes an energy filter is being used, and this should be clearly stated.*

We have added this clarification in the inelastic scattering paragraph of the Supplementary Information.

> *(c) The reasons for not including other factors, such as DQE and the temporal and spatial coherence envelope functions, could be stated.*

We have added a note in the dose-weighting section clarifying that these instrument-dependent attenuation factors were not explicitly included, and that they could be incorporated as additional frequency-dependent weighting terms without changing the structure of the derivation.

> *(d) The flexibility and heterogeneity in protein structures, especially at high spatial frequencies, must also be a reason for a gap from experiment to theory, but this is not clearly stated.*

We agree. We have added a statement in the "Remaining gaps" section noting that structural flexibility and conformational heterogeneity act as an additional envelope that attenuates high-resolution signal relative to the rigid-particle model assumed in our derivation.

> *Additional Minor Comments*
>
> *(15) It is noted in the discussion that 2DTM-based single-particle alignment simplifies the processing pipeline. Although true, I think stating the computation time would be useful for the reader.*

We have added computation times to the Discussion. For a typical single-particle dataset of ~2,000 micrographs (5k × 4k pixels), a 2DTM search without defocus refinement completes in approximately one day on 64 NVIDIA A6000 GPUs. Once particles are located with their orientations and positions, a single 3D reconstruction is sufficient without further refinement, eliminating the iterative 2D classification, *ab initio* modeling, 3D classification and refinement steps of a conventional pipeline.

> *(16) There are some formatting issues with e−/A 2, sometimes losing the minus sign.*

Thank you for catching this. We have corrected all instances to consistently use $e^{-}/Å^{2}$ throughout the manuscript.

https://doi.org/10.7554/eLife.109790.2.sa0