Factors affecting macromolecule orientations in thin films formed in cryo-EM

Swati Yadav [a] and Kutti R. Vinothkumar [a] *

[a] National Centre for Biological Sciences, Tata Institute of Fundamental Research

GKVK Post, Bellary Road, Bangalore, 560065.

Correspondence – vkumar@ncbs.res.in

Abstract: The formation of a vitrified thin film embedded with randomly oriented macromolecules is an essential prerequisite for cryogenic sample electron microscopy. Most commonly, this is achieved using the plunge freeze method described nearly 40 years ago. Although this is a robust method, the behaviour of different macromolecules shows great variation upon freezing and often needs to be optimized to obtain an isotropic, high-resolution reconstruction. For a macromolecule in such a film, the probability of encountering the air-water interface in the time between blotting and freezing and adopting preferred orientations is very high. 3D reconstruction using preferentially oriented particles often leads to anisotropic and uninterpretable maps. Currently, there are no general solutions to this prevalent issue, but several approaches largely focusing on sample preparation with the use of additives and novel grid modifications have been attempted. In this study, the effect of physical and chemical factors on the orientations of macromolecules was investigated through an analysis of select well-studied macromolecules and important parameters that determine the behaviour of proteins on cryo-EM grids were revealed. These insights highlight the nature of interactions that cause preferred orientations and can be utilized to systematically address orientation bias for any given macromolecule and provide a framework to design small-molecule additives to enhance sample stability and behaviour.

Keywords: cryo-EM; thin films; preferred orientation; surfactants; temperature

1. Introduction

Single particle cryo-EM has become an indispensable technique in structural biology owing to its relative ease of image acquisition, reconstruction, and structure determination. Consequently, there has been a steady increase in the number of structures deposited in the Electron Microscopy Database (EMDB) in the past few years (Patwardhan, 2017). Near atomic resolution maps of biological macromolecules can now be routinely obtained due to improved hardware and specimen preparation methods along with developments in algorithms for image processing (Bai, 2021; Y. Cheng, 2015; Chua et al., 2022; Lyumkis, 2019; Nogales, 2015; Vinothkumar & Henderson, 2016; Zheng et al., 2023). Obtaining a thin film of purified macromolecules in a vitreous layer of ice transparent to electrons is an important first step in cryo-EM structure determination. Currently, the plunge freeze method is the most commonly used technique for specimen preparation. This method involves placing a drop of the sample on a plasma-cleaned metal grid (typically coated with carbon and has an array of holes) on which a thin film is formed as the excess fluid is blotted away using blotting paper (Dubochet et al., 1982; Dubochet & McDowall, 1981). The grid is then plunged into a cryogen, causing the vitrification of water and preserving the native structures of macromolecules (Dubochet et al., 1985; Dubochet & McDowall, 1981; Taylor & Glaeser, 1974).

The freezing and blotting conditions must be extensively optimised for temperature and humidity to reproducibly create a single layer of macromolecules on a thin film of ice suitable for imaging. Because of their large size, the first test specimens used to optimise these conditions were liposomes and viruses. They lose their structural integrity if the osmolarity of the buffer changes owing to slow cooling (Adrian et al., 1984; Dubochet et al., 1988). It has been shown through recent studies that during freezing, very often macromolecules can interact with the air-water interface (AWI). This interaction can cause the macromolecules to denature, dissociate, adapt preferential orientations, or avoid the holes completely and stick to the carbon support (D'Imprima et al., 2019; Drulyte et al., 2018; Kampjut et al., 2021; Liu & Wang, 2023; Noble, Dandey, et al., 2018; Noble, Wei, et al., 2018). Along with protein biochemistry, the propensity of macromolecules to adopt preferential orientations is perhaps a common rate-limiting step in high-resolution structure determination using cryo-EM.

Depending on the instrumental setup, the time between blotting and freezing can vary, but it is typically around a second. During this period, the macromolecules are suspended in

an extremely thin film and constantly tumble in a solution, undergoing Brownian motion. An intact macromolecule undergoing diffusion can encounter and get trapped in specific orientations at the freshly formed AWI. This effect is protein-dependent and more severe in some cases than in others (Naydenova & Russo, 2017).

When and how does having preferred orientations become a problem? There are cases where the preferred orientation exists but does not pose a problem, such as in proteasomes and viruses, because of their highly symmetric architecture (Campbell et al., 2015; Vogel et al., 1986). A characteristic feature of maps reconstructed from particles with orientation bias is the stretching of density in the cryo-EM map in one direction (anisotropy) (Figure 1), which makes the maps uninterpretable (Sorzano et al., 2022). In such cases, the resolution estimate based on the comparison of the half-maps is unrealistic, and the features in the map do not justify the resolution. Two examples of macromolecules exhibiting orientation bias are shown in Figure 1. It should be noted that 2D projections of different views of a macromolecule contribute differently to the final reconstruction, and this is an important criterion for obtaining a reconstruction with isotropic resolution, as shown by (Tan et al., 2018) and is further demonstrated using catalase as an example in Figure S1.

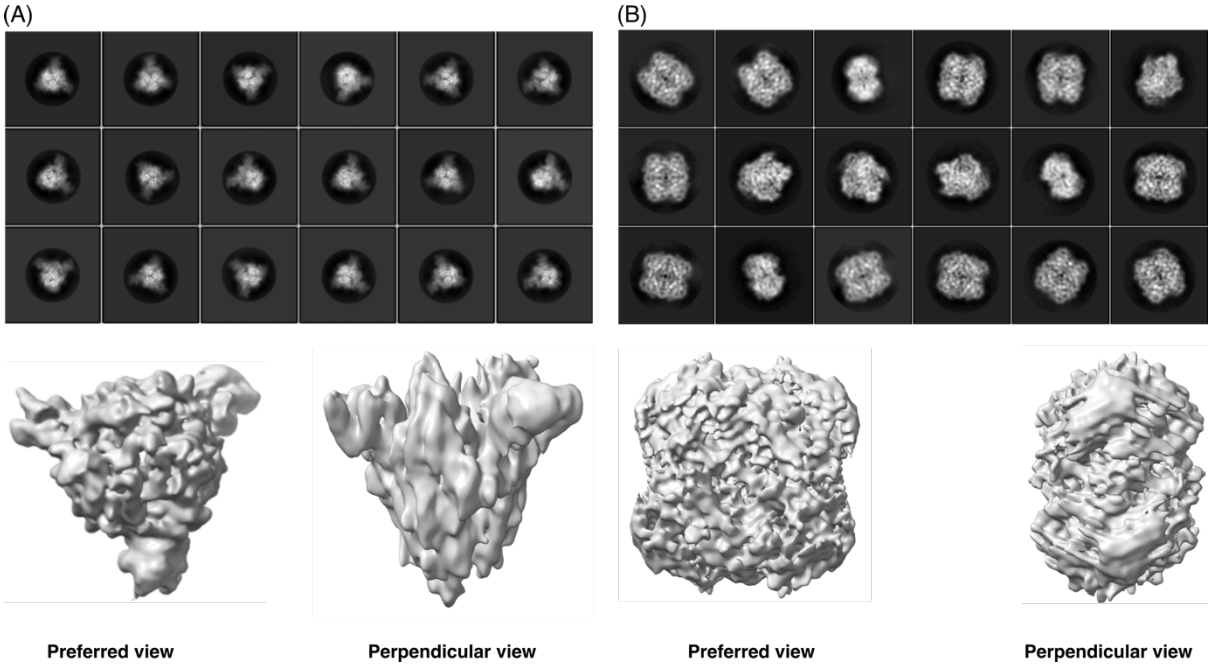


Figure 1: Examples of anisotropic cryo-EM maps resulting from orientation bias. The upper panel shows the reference-free 2D class averages of (A) SARS-CoV-2 spike protein and (B) human erythrocyte catalase. For the spike protein, the preferred bottom views are observed. In the case of catalase, preference for the top/bottom view is evident. In the lower panel, 3D maps with anisotropic

features are shown with the preferred and perpendicular views as labelled. The symmetries applied during reconstruction were C1 and D2 for the spike protein and catalase, respectively.

Several approaches have been proposed to address the preferred orientation problem (Drulyte et al., 2018). One such approach is to change the sample preparation process by skipping the blotting with filter papers altogether and minimising the time between thin film creation and freezing to as little as 100 ms. This method uses pL to nL quantities of the sample, which is sprayed onto a self-wicking grid that contains nanowires to absorb the excess liquid, creating a thin film, followed by rapid freezing of the grid. This protocol has been shown to introduce views in some proteins; however, even in these cases, the timescale of 100 ms has been large enough to allow the formation of denatured protein films on the surface (Dandey et al., 2020; Jain et al., 2012; Wei et al., 2018). Other options include using support films, such as carbon and functionalized graphene, or tilting the stage to obtain alternate views (Aiyer et al., 2023; Lu et al., 2022; Noble, Wei, et al., 2018; Tan et al., 2018; Xu et al., 2021). Tilting the stage or use of carbon support film suffers from a decrease in signal/noise ratio, especially if the protein is small, as the contrast is decreased due to either the support film or thick ice while imaging at higher tilt angles and tilting is less commonly used. While carbon support films are more commonly used in samples such as ribosomes, viruses, and some ion channels (Baker et al., 2021; Hesketh et al., 2018; Tobiasson et al., 2022), graphene/graphene oxide is suitable for low-molecular-weight specimens. These techniques have been successful in alleviating orientation bias in a few cases but are not generally applicable. Additionally, there is now some image processing software that deals with the orientation bias in the data better than others because of the better weighing of the different projections during reconstruction, but the problem still persists (Ramírez-Aportela et al., 2022; Sorzano et al., 2021).

The most widely used approach in decreasing the preferred orientation of macromolecules is to add surfactants to the sample buffer prior to grid preparation (Chen et al., 2019, 2022; B. Li et al., 2021). This method is popular because it is simple and does not require any additional steps, techniques, or instruments. A large number of surfactants are available and can be used, but care must be taken to avoid disrupting the protein's native structure in the process. Often, extensive trial and error is involved when surfactants are screened to overcome the preferential orientation problem. EMDB has at least 83 different soluble protein structures (retrieved in 2022), in which a detergent is added to the buffer. The most popular choice of detergent is non-ionic or zwitterionic. Recently, cationic or anionic

118 detergents have been found to work for a few samples but their use is not generalized yet (B.
119 Li et al., 2021). We also encountered the preferred orientation problem in many of our
120 projects. We asked if an informed decision on the choice of grid freezing conditions can be
121 made based on the properties of the protein, thus minimising the time spent screening many
122 surfactants and specimen preparation conditions.

123 To achieve this goal, we tested some commonly used surfactants with different
124 properties on a set of five proteins, CRP-pentamer, CRP-decamer, Catalase, PaaZ and Spike.
125 In addition, we explored the effect of the presence of the histidine tag on Spike and β-
126 galactosidase and physical factors such as the temperature during the sample application step
127 on catalase and PaaZ. We also serendipitously observed the effect of grid hole dimensions of
128 the holey carbon grid, on the orientation distribution of catalase and have discussed this
129 briefly. Through this analysis, we identified factors that affect and determine the
130 macromolecule behaviour on grids before freezing and studied their effects with a focus on
131 the preferred orientation problem. This account highlights the factors that contribute to
132 orientation bias and provides valuable information that can assist in achieving the optimal
133 freezing conditions for any given macromolecule.

134 **2. Results**

135 **2.1. Analysis of preferred views of selected macromolecules**

136 A set of well-characterized proteins with varying molecular weights, shapes,
137 symmetries, surface charge distributions, and known experimental structures were selected,
138 to rule out any artefacts that may arise owing to the changing conditions of sample
139 preparation. The test proteins used for the analysis include C-reactive protein (CRP)
140 pentamer, CRP decamer, human erythrocyte catalase, SARS-CoV-2 spike protein, *E. coli.*
141 PaaZ and *E. coli* β-galactosidase. Micrographs and 2D class averages of these datasets
142 collected at the start of the study are shown in Figure 2.

143 CRP exists in a concentration-dependent equilibrium of pentamer and decamer of a 25
144 kDa polypeptide under physiological conditions (Okemefuna et al., 2010). At the
145 concentration range used for cryo-EM experiments, both pentamer and decamer populations
146 were observed in the same micrograph and analyzed separately (Figure 2A, B). The effect
147 that both these populations can have on each other's behaviour in the thin film is an
148 interesting concept that has not been analysed in the current study but is discussed briefly. C5
149 symmetry was applied to the smaller pentameric molecule, and no symmetry was applied to

150 the decamer during image processing. The preferred view in the CRP pentamer is the
151 top/bottom view, where the symmetric arrangement of the monomers can be seen. Some side
152 views are observed, but no tilted side views are seen in the 2D class averages (Figure 2A). In
153 contrast, in the case of the CRP decamer side views are predominantly observed and no
154 top/bottom views are seen (Figure 2B). Catalase exists as a dimer of tetramers and possesses
155 D2 symmetry (Ko et al., 2000). It adopts a preferred top/bottom view on the grids when the
156 grids are prepared at 4°C using Quantifoil 0.6/1 holey carbon grids (Figure 2C). This enzyme
157 has previously been noted for its peculiar behaviour during cryo-EM grid preparation and has
158 been used as a test specimen to overcome orientation bias, and we used it as a control sample
159 and to probe additional parameters (S. Chen et al., 2022; Fan et al., 2021; Vinothkumar &
160 Henderson, 2016). The SARS-CoV-2 spike protein is now a well-studied trimeric viral
161 membrane protein whose symmetry can vary depending on the conformation of the receptor-
162 binding domain (Huang et al., 2020; Wrapp et al., 2020). Spike trimer has three receptor
163 binding domains (RBD), one per monomer, and they can adopt different conformational
164 states and thus dictate the symmetry of the protein. Soluble spike protein shows a preference
165 for a bottom view on grids in the absence of an additive (Figure 2D). In this data set, we
166 observed that there was heterogeneity in the RBD domain, and hence C1 symmetry was
167 applied for the rest of the analysis.

168 PaaZ is a bifunctional enzyme and consists of six polypeptides, where a domain-
169 swapped dimer assembles to form a trimeric structure. In an earlier study, when frozen in ice,
170 the enzyme was found to form clumps, and aggregates, which was overcome by the use of
171 graphene oxide and a low concentration of protein (Sathyanarayanan et al., 2019). PaaZ
172 adopts a preferred side view in the absence of an additive when frozen in ice at 4°C, with a
173 tendency to clump or cluster (Figure 2E). Nevertheless, the isolated particles were sufficient
174 to obtain reasonable maps, and here we were interested in screening additional parameters to
175 check if some of the aggregation and clumping could be overcome to enhance the quality of
176 data of the enzyme in ice. β-galactosidase from *E. coli* is one of the well-studied enzymes as
177 well as a test specimen in cryo-EM (Bartesaghi et al., 2018; Juers et al., 2012). It is a
178 glycoside hydrolase enzyme, similar to catalase and exists as a tetramer with D2 symmetry.
179 When purified with a tag at its N-terminus, the enzyme adopts a preferred side view and
180 tends to form clumps or aggregates in ice (Figure 2F). Among all the standard proteins tested,
181 spike, PaaZ and β-galactosidase were obtained by overexpression and had additional poly-
182 histidine affinity tags at either the N or C termini of their monomers.

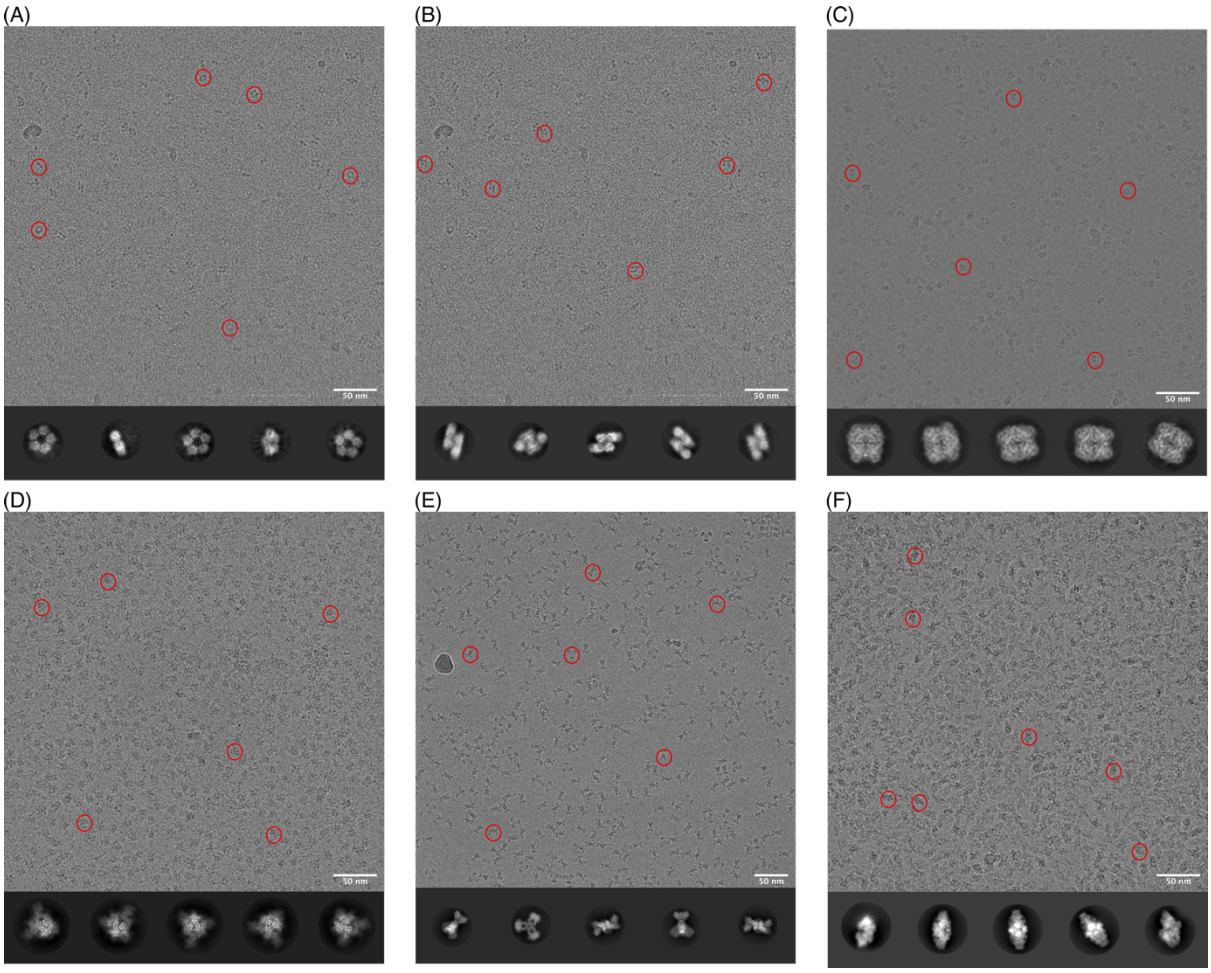


Figure 2: Representative micrographs with few select particles indicated with red circles and 2D class averages of test proteins used in this study. (A) C-reactive protein (CRP) pentamer adopts a preferred bottom view, which shows the pentameric arrangement of the monomers (B) C-reactive protein decamer adopts a preferred side view, which shows the staggered arrangement of two CRP pentamers stacked on top of each other. Same micrograph is used in A and B. (C) catalase adopts a preferred top view as seen in the micrograph and 2D class averages (D) SARS-CoV2-spike adopts a preferred bottom view showing the trimeric arrangement (E) PaaZ adopts a preferred side view as seen in the 2D class averages and the micrograph shows occasional clumping of hexamers on the grids (F) β-galactosidase with an N-terminal his tag adopts a preferred side view as seen in the 2D class averages and the micrograph shows aggregation on grids. Among the above datasets catalase and PaaZ grids were prepared at 4°C and all other grids were prepared at 20°C.

2.2. Surfactants affect macromolecule orientation distributions in a charge-dependent manner

We tested a specific set of surface-active molecules with varying head group charges (cationic, anionic, and non-ionic), chain length packing (unsaturated vs. saturated alkyl chain), differing in their critical micelle concentration (CMC) and concentration on the set of

200 proteins described above. All the surfactants were used at concentrations lower than their
201 respective CMCs. The properties of the surfactants are listed in Table 1.

202 **Table 1** Properties of surfactants used in this study

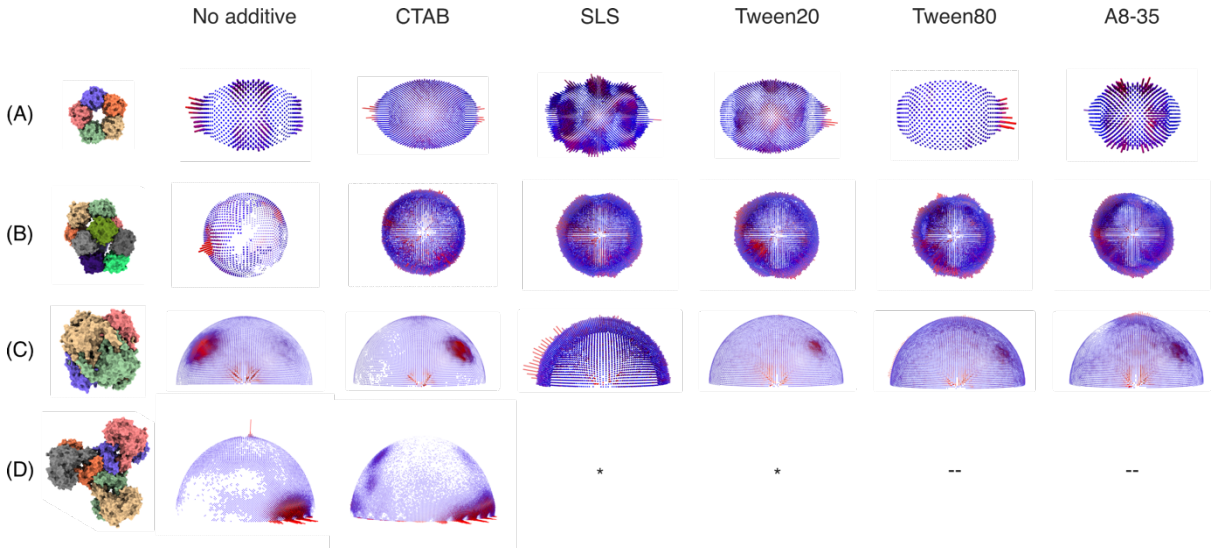
Additive	Charge	Ionic or non-ionic	CMC*	Concentration used	Aggregation number	Molecular weight*	Alkyl chain length	Saturation in alkyl chain
CTAB	Positive ++	Ionic	1mM (0.04%)	0.054mM (0.002%)	170	364	16	Saturated
SLS	Negative --	Ionic	14.4mM (0.4%)	1.37 mM (0.04%)	-	293	12	Saturated
Tween20	Neutral	Non-ionic	0.06mM (0.007%)	0.04mM (0.005%)	80	1225	12	Saturated
Tween80	Neutral	Non-ionic	0.012mM (0.002 %)	0.038mM (0.005%)	58	1310	18	Unsaturated
A8-35	Negative -	Ionic	NA	0.01%	NA	NA	NA	NA

203 * - CMC and Molecular weight values are from the Anatrace and Sigma Aldrich webpages.

204

205 Upon the addition of cationic surfactant cetyltrimethylammonium bromide (CTAB), a
206 change in the orientation distribution was observed for all the samples tested (Figure 3). An
207 improvement in the orientation distribution of CRP pentamer and decamer was observed
208 resulting in isotropic 3D reconstructions (Figure 3 A, B & Table 2). For catalase, the initial
209 no-additive dataset was collected from a specimen prepared at 4°C on a Quantifoil 0.6/1 grid,
210 which has a severe preference for the top/bottom views (Figure 2C). However, as the grids
211 for all the additive datasets were made at 20°C, for an appropriate comparison, we prepared
212 catalase grids at 20°C with no additive also. This dataset shows a preference for a tilted side
213 view, annotated as side view A (Figure S2). The changes in orientation distributions of
214 catalase are a result of variations in the incubation temperature and grid hole geometry, which
215 we did not anticipate to have such a large effect on the quality of the reconstruction and were
216 explored further in this study (discussed below). Upon addition of CTAB to catalase grids,
217 the preference for the tilted side view A was lost, instead the particles adopted a preference
218 for the other tilted side view B, and a map of similar quality to the no additive dataset was
219 obtained (Figure 3C & Table2). In the case of PaaZ, the preference for the side view was
220 maintained even in the presence of CTAB, but additional side-tilted views were sampled,
221 which led to a better resolution and higher-quality map (Figure 3D & Table 2). The addition
222 of a negatively charged detergent sodium lauryl sarcosine (SLS), resulted in evenly

223 distributed orientation distributions, with no preference for any particular view for all cases
224 tested except PaaZ (Figure 3). The resulting maps with SLS as additive in all the cases are of
225 lower resolution, which may be due to a lower number of particles or the ice thickness and
226 possible effect of the anionic headgroup (Table 2).



227
228 **Figure 3:** Orientation distribution plots from Relion (Scheres, 2012) of proteins upon the addition of
229 surfactants with varying properties to the sample buffer before grid preparation. The reference structures
230 of the respective proteins are generated by creating a surface representation in ChimeraX from models
231 from PDB with IDs 7pkb, 7pk9, 1dgf, and 6jql. (A) CRP-pentamer orientation distribution change upon
232 the addition of any surfactant. The distributions are distinct from each other except Tween 20 and Tween
233 80, which have similar distributions (B) CRP-decamer orientation distribution change upon surfactant
234 addition, all surfactants lead to a similar even orientation distribution (C) catalase orientation
235 distribution changes upon surfactant addition where the charged surfactants have distinct distributions
236 (CTAB, SLS) and the neutral surfactants (Tween 20/80) and A8-35 show similar distributions (D) PaaZ
237 orientation distributions change upon the addition of cationic surfactant CTAB. * The effect of SLS and
238 Tween 20 on PaaZ was also tested, but visual inspection of the micrographs (Figure S3) showed no
239 improvement and no data was collected and thus not included. -- The effect of Tween 80 and A8-35
240 was not tested on PaaZ.
241

242 The addition of non-ionic detergents Tween 20 and Tween 80 to CRP pentamer,
243 decamer, and catalase led to evenly distributed orientations (Figure 3), but PaaZ showed
244 aggregation on grids (Figure S3). Further, Tween 80 was tested as it is known to form denser
245 layers at AWI compared to Tween 20 due to differences in the alkyl chain (length and
246 unsaturation) (Szymczyk et al., 2018). Additionally, it has been shown to differ in its ability
247 to separate protein films from the AWI as compared to Tween 20 (Rabe et al., 2020). In the

macromolecules tested in our study, no significant difference was observed between the orientation distributions obtained from either of the Tween surfactant datasets (Figure 3). However, we note that with CRP pentamer, the orientation distributions look similar with both Tween 20 and 80 but only the addition of Tween 20 led to an isotropic map (R=3.3 Å), while the addition of Tween 80 led to a low-resolution map (R=7.5 Å). These differences can be a result of variations in ice thickness or due to differential interaction of Tween 80 with CRP pentamer and more experiments need to be performed to determine the cause of this discrepancy. Amphipol A8-35, a surface-active ionic polymer that does not form typical micelles and is thus different from other surfactants was also tested to see its effect on orientations. The orientation distributions obtained in this case were similar to those of the Tween 20 and Tween 80 datasets (Figure 3). The observed results with the given test samples indicate that the charge on the detergent headgroup is an important parameter, whereas the chain length and saturation of the hydrophobic chain have very little effect in modulating the orientation distribution of the macromolecules.

To understand the importance of the discrete views that were sampled and their contribution to the quality of the map, the efficiency of the orientation distribution (E_{od}) using CryoEF (Naydenova& Russo, 2017) and the sphericity of the final maps using the 3DFSC (Tan et al., 2018) server were calculated (Table 2). E_{od} provides an estimate of the coverage of the 3D Fourier space from the 2D projections in the data, whereas sphericity provides a measure of anisotropy in the final maps by estimating the directional resolution. Among all the datasets, the CTAB datasets stand out and result in higher-resolution maps. We think this could be a result of optimum ice thickness, along with a sampling of the orientations that contribute more to the reconstruction, which is also demonstrated in Figure S1 by using a subset of particles from catalase data as an example.

2.3. The presence of solvent-exposed poly-histidine tag affects protein orientations in thin films

SARS-CoV-2 spike protein and *E.coli.* β-galactosidase are two well-studied samples by cryo-EM and multiple structures have been reported in the past (Bartesaghi et al., 2018; Bodakuntla et al., 2023; Esfahani et al., 2024; Hardenbrook & Zhang, 2022; Harvey et al., 2021; Juers et al., 2012; Wrobel, 2023). However, we observed orientation bias with these two recombinantly purified proteins, with differing severity. The spike protein showed severe

280 orientation bias, which resulted in anisotropic maps of poor quality that could not be
281 improved by surfactant addition (data not shown). β-galactosidase showed comparatively less
282 orientation bias and led to a reconstruction of moderate quality but the map was still
283 anisotropic (Figure 4B). Given that both of these proteins feature poly-histidine affinity tags,
284 and considering that the preferred view in both cases may involve an exposed tag, we
285 hypothesized that this tag could be inducing a bias in orientation (Figure 4A).

286 **Table 2.** Comparison of parameters of the no additive and surfactant additive datasets.

Protein	Condition	No. of particles	Resolution (Å) (half-map FSC @0.143)	Efficiency of Fourier space coverage	Sphericity
CRP pentamer	No additive	14,601	18	0.78	NA
	CTAB	36,353	3.3	0.80	0.98
	SLS	31,699	4.2	0.80	0.97
	Tween 20	25,674	3.3	0.80	0.86
	Tween 80	32,330	7.5	0.69	NA
	A8-35	26,737	10	0.78	NA
CRP decamer	No additive	9,419	20	0.52	NA
	CTAB	25,992	3.5	0.85	0.98
	SLS	59,211	3.7	0.79	0.97
	Tween 20	51,784	4.0	0.78	0.92
	Tween 80	36,870	4.2	0.79	0.76
	A8-35	104,369	3.5	0.78	0.98
Catalase	No additive	138,000	2.7	0.72	0.96
	CTAB	153,336	2.8	0.76	0.97
	SLS	33,241	3.7	0.80	0.98
	Tween 20	88,395	2.9	0.78	0.98
	Tween 80	92,163	2.9	0.80	0.96
	A8-35	122,000	3.1	0.77	0.98
PaaZ	No additive	51,393	4.0	0.76	0.80
	CTAB	89,454	2.3	0.75	0.98

287

288 Thus, the affinity tag from the spike protein and β-galactosidase were removed by
289 protease cleavage, which led to a change in orientation distributions in both cases (Figure 4).
290 In the case of spike, although the orientation bias remained, the preference shifted from the
291 bottom view (where the histidine tag is attached) to the top view (Figure 4B). Subsequently,
292 CTAB was added to the sample, imaging was performed on thin ice, and side-tilted views
293 were obtained, which resulted in a reasonable isotropic map. The final reconstruction was
294 obtained with 261,703 particles to an overall resolution of 3 Å (Figure 4B). At this resolution,
295 amino acid side chains and glycosylation on the protein surface were visible in the ordered
296 regions but the regions of RBD were poorly resolved substantiated by the local resolution
297 plot (Figure S4). Further, we used the spike data set to evaluate the effect of different post-
298 processing methods and the fit of the model to the map (He et al., 2023; Pintilie et al., 2020;
299 Sanchez-Garcia et al., 2021)(summarized in Figure S4). In comparison to the other datasets,

the increase in E_{od} and sphericity is evident in this case (Figure 4B). To verify that the improvement in map quality is a result of new orientations and not an increase in the total number of particles, all the top/bottom views were removed for 3D refinement, and the reconstruction using only the side/tilted views corresponding to ~94,000 particles was performed. This resulted in a map with a sphericity of 0.88 and a resolution of 3.4 Å (Figure S5) similar to that with all the particles, illustrating that the top/bottom views contribute little information to the final reconstruction of spike protein.

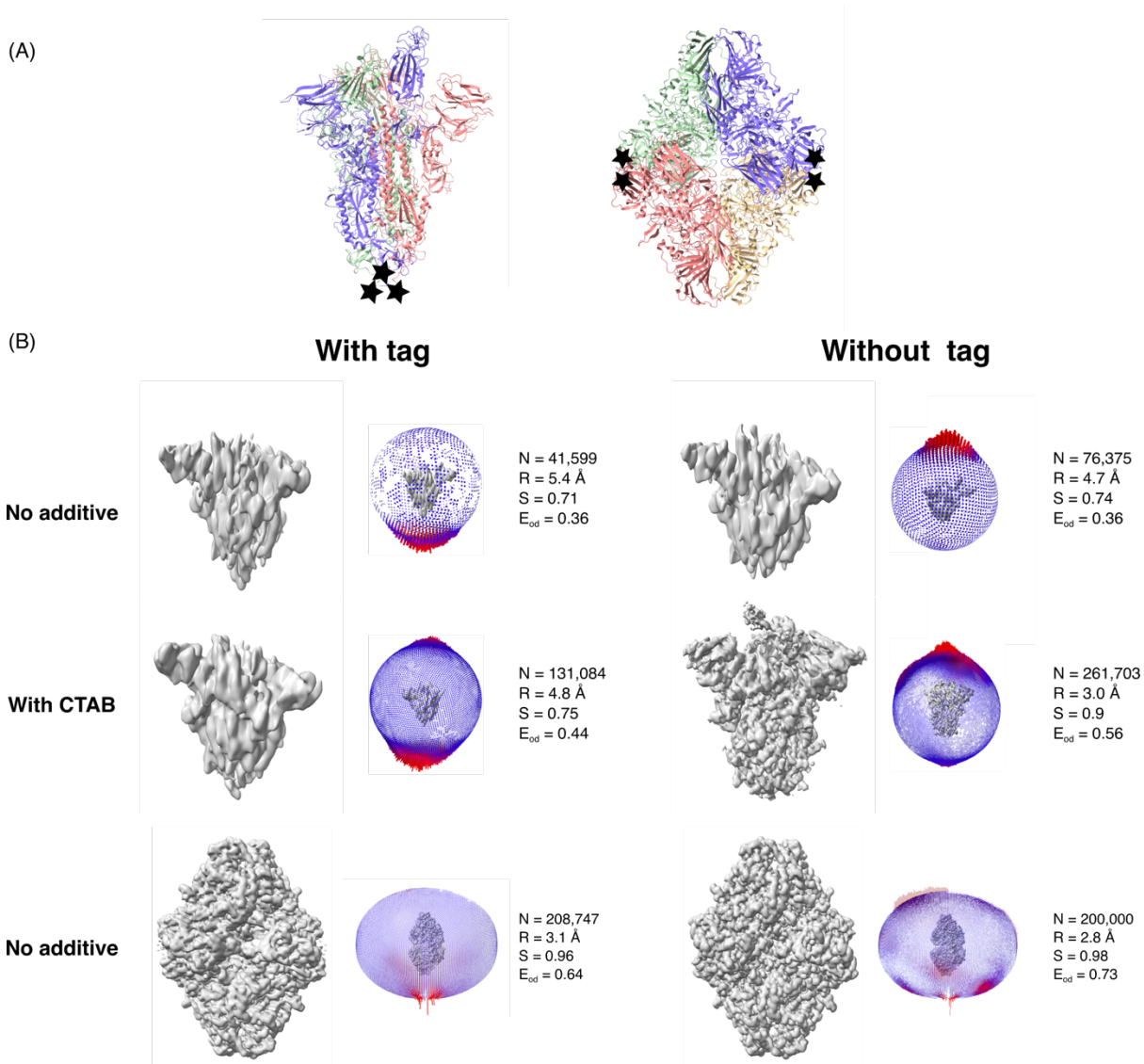


Figure 4: Effect of poly-histidine affinity tag on SARS-Cov2 spike protein and β-galactosidase orientation distribution. Different parameters that are used to analyse the quality of the maps are shown next to the orientation plots. N indicates the number of particles used for reconstruction, R indicates the final resolution of the map, S indicates the sphericity and E_{od} indicates the efficiency of Fourier space coverage. (A) The location of the tags on the protein models is indicated by black stars. The models used as reference are PDBs 8h3d and 6cvm for the spike protein and β-galactosidase respectively. (B)

Orientation distribution plots of spike protein change upon removal of the affinity tag but the change is not enough to obtain an isotropic map. The addition of cationic surfactant CTAB further alters orientations of the spike protein without tag and leads to a more isotropic map. β-galactosidase enzyme orientations change upon removal of the affinity tag and lead to an isotropic high-resolution map without any additive. The unsharpened final combined maps are shown in grey in panel B.

In the case of β-galactosidase, the orientation bias was reduced and alternate views were sampled as indicated by the orientation plot and improvement of the E_{od} value from 0.64 to 0.73 when the tag is cleaved. The resulting reconstruction also shows improvement as indicated by visual inspection of the map (Figure 4B and Figure S6) and the improvement in resolution and sphericity. Zoomed-in areas of the map to highlight the difference in map quality and comparison of the half map and model vs. map FSCs of the β-galactosidase are shown in Figure S6.

2.4. The temperature of the incubation chamber during freezing affects protein orientations

The temperature at which the grids are held and blotted is critical for determining ice thickness, protein stability, and dynamics. Hence, grids were prepared at different temperatures to determine the effect of temperature on the orientation distributions. Typically, lower temperatures are used during sample application to grids to keep the protein stable (except in some cases, such as microtubules, where a higher temperature is a prerequisite), and the temperature as a parameter is not very often screened to overcome the preferred orientations. For catalase, the orientation distributions varied significantly when grids were prepared at different temperatures. The coverage of the 3D Fourier space (E_{od}) improved from 0.64 at 4°C to 0.72 at 20°C to 0.77 at 37°C (Figure 5A). However, the resolution of the final reconstruction is best for the 4°C dataset, which could be due to the higher number of particles used for reconstruction, that is, 478,656 compared to 138,031 for the 20°C dataset and 75,312 for the 37°C dataset. The respective B factors (from Relion post-process) for the reconstructions were -107 $Å^2$, -110 $Å^2$, and -126 $Å^2$.

In the case of PaaZ, when the grids were prepared at 4°C, 20°C and 37°C, increasing higher-order structures were observed, particularly at 37°C. At 4°C, higher-order structures were minimal, but a preference for the side view was observed. Data collected at 4°C and 20°C have similar E_{od}, and the orientation distribution plots show a slight increase in the

sampling of additional side-tilted views at 20°C, indicated by the reduction in empty spaces in the orientation plots obtained from Relion (Figure 5B). However, clumping of the particles was still observed at 20 °C. Hence, CTAB was added to the sample and the grids were prepared at 4°C, which led to less clumping and consequently resulted in a high-resolution map at 2.3 Å, described further in the next section.

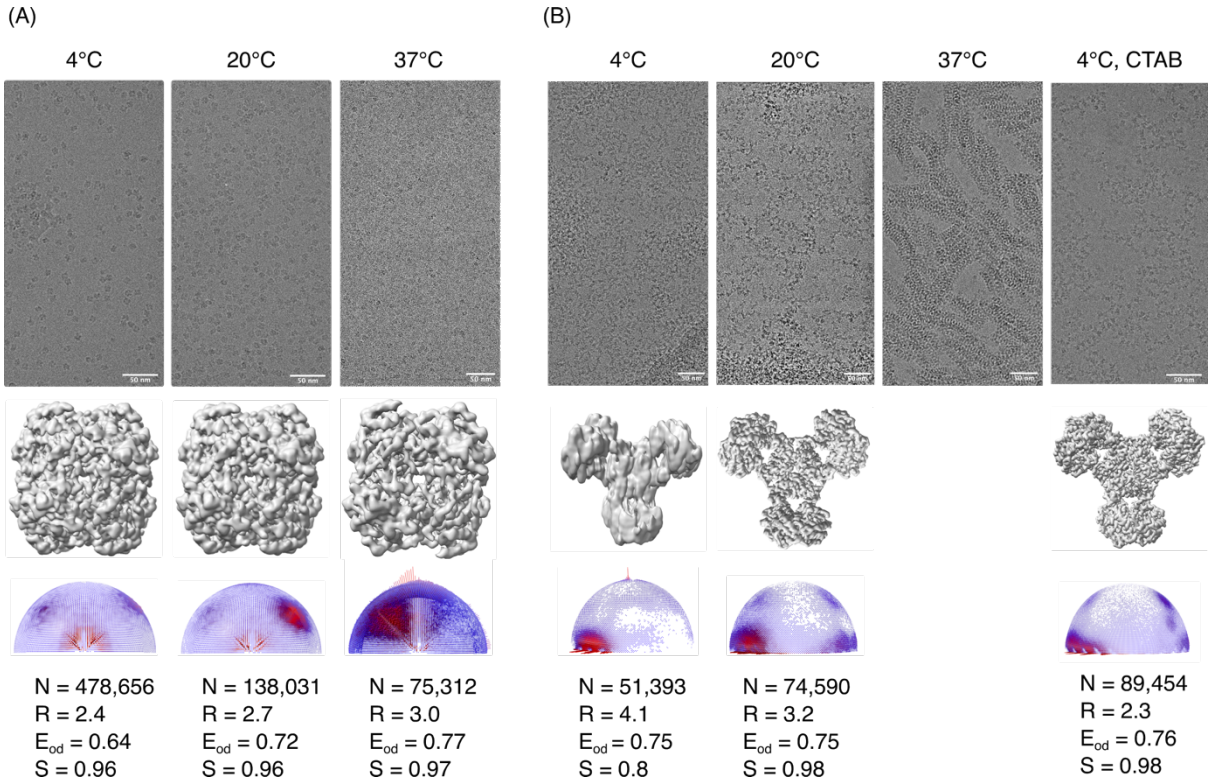
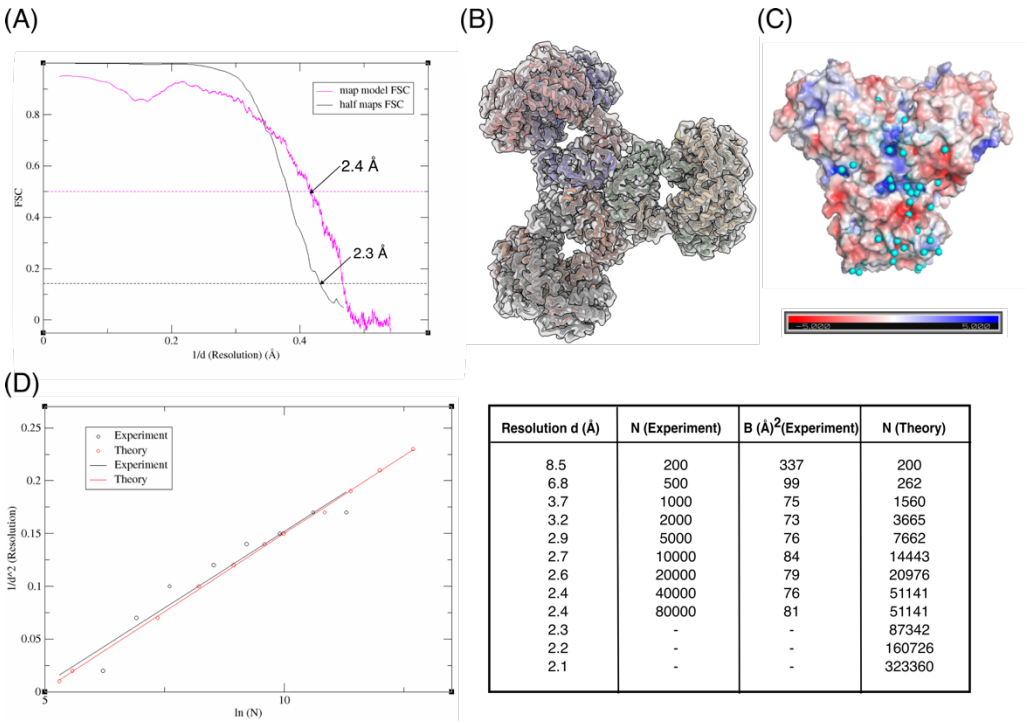


Figure 5: Effect of temperature during cryo-EM sample preparation of catalase and PaaZ. Micrographs, maps, orientation distribution plots and different parameters that are used to analyse the quality of the maps are shown. N indicates the number of particles used for reconstruction, R indicates the final resolution of the map, S indicates the sphericity and E_{od} indicates the efficiency of Fourier space coverage. (A) Catalase orientation distributions change significantly when grids are blotted at different temperatures in the absence of any additive (B) PaaZ orientation distributions change slightly when grids are held and blotted at different temperatures in the absence of any additive. In the case of PaaZ, the condition with grids prepared at 4°C with CTAB additive is included for comparison as this combination led to a high-resolution isotropic map.

It is evident from these observations that physical factors, such as the grid preparation temperature can affect protein behaviour and should be considered as an important screening condition when dealing with orientation bias along with surfactants.

366 **2.5. High-resolution map of *E. coli* PaaZ in ice**

367 To understand the mechanism of action of a macromolecule, high-resolution maps are
368 required to build an accurate model and more importantly, to model the solvent molecules
369 and other ligands of interest. We previously reported the structure of PaaZ, a bifunctional
370 enzyme from *E. coli* that catalyzes two steps of the phenylacetic acid degradation (paa)
371 pathway. In the previous study, the structure of PaaZ was determined at ~2.9 Å using
372 graphene-oxide support grids (Sathyanarayanan et al., 2019). During that study, it was
373 observed that the enzyme tended to clump when frozen in ice in the absence of any additive,
374 and the use of a graphene oxide support and a low concentration of enzyme yielded a very
375 good distribution of particles. In the current study, we obtained a high-resolution map using
376 CTAB as an additive during grid preparation, with D3 symmetry and higher-order aberration
377 correction during image processing. The resolution estimated by comparison of the half maps
378 FSC (at 0.143 threshold) is 2.3 Å, and a comparable resolution of 2.4 Å is obtained with the
379 comparison of the map vs. model FSC (at 0.5 threshold) (Figure 6A). The model-fitted PaaZ
380 structure is shown in Figure 6B. Several water molecules could be modelled at this resolution
381 and an electrostatic potential map of the dimer with water molecules modelled as spheres and
382 coloured in cyan is shown (Figure 6C). To further analyse the quality of the data, the Reslog
383 plot was calculated by obtaining a 3D reconstruction using different numbers of particles and
384 plotting the resolution ($1/d^2$) obtained with the number of particles (lnN) used (Figure 6D)
385 (Rosenthal & Henderson, 2003). From the experimental data, it can be seen that as the
386 resolution approaches the Nyquist limit of 2.1 Å, the addition of more particles from 40,000
387 to 80,000 does not lead to an improvement in resolution. The slope of this graph was used to
388 calculate a B-factor of 69.4 $Å^2$, which is slightly lower than the B-factor estimated from the
389 post-processing option in Relion, which was used for map sharpening (-75 $Å^2$). A 200-
390 particle dataset was used as a base, and the particles required to obtain a certain resolution
391 were calculated theoretically. These theoretical estimates are in agreement with the
392 experimental values.



Resolution d (Å)	N (Experiment)	B (Å)^2(Experiment)	N (Theory)
8.5	200	337	200
6.8	500	99	262
3.7	1000	75	1560
3.2	2000	73	3665
2.9	5000	76	7662
2.7	10000	84	14443
2.6	20000	79	20976
2.4	40000	76	51141
2.4	80000	81	51141
2.3	-	-	87342
2.2	-	-	160726
2.1	-	-	323360

Figure 6: High-resolution Cryo-EM map of PaaZ grids prepared at 4°C with CTAB additive (A) Comparison of the half maps and map versus model FSCs of the PaaZ dataset. (B) The six polypeptides coloured individually and in cartoon representation fit into the cryo-EM map (transparent grey) of PaaZ. (C) Electrostatic potential surface representation of the domain-swapped PaaZ dimer with waters modelled and shown as cyan spheres. (D) Reslog plot of PaaZ with the experimental and theoretical numbers of particles required to reach a particular resolution. N indicates the number of particles used for reconstruction, d is the resolution, and B indicates the B factor, as estimated by the Relion post-processing. D3 symmetry was applied for the reconstruction, and the Reslog plot indicates the number of particles used, not the asymmetric units averaged.

3. Discussion

In this era of artificial intelligence-based protein structure prediction (Baek et al., 2021; Jumper et al., 2021; Schauperl & Denny, 2022), there is still no substitute for the joy one feels when the molecule of interest is imaged and observed for the first time after hard work of expression and purification. However, the first data collection for a new sample often does not result in a high-resolution structure, as there can be one or more challenges that need to be overcome. Of these, the preferred orientation of macromolecules and the possibility of denaturation due to exposure to AWI is a serious issue and one of the major bottlenecks. It is now known that the tendency of macromolecules to adopt preferential orientation when frozen in ice using cryo-EM grids is a consequence of protein-AWI interaction (J. Chen et al.,

414 2019; Noble, Wei, et al., 2018), but the nature of this interaction remains unclear. On the
415 other hand, amphipathic molecules have been used in cryo-EM grid preparation since the
416 1980s (Frederik et al., 1989). Their use in solving the preferred orientation of proteins by
417 modulating the air-water interface has recently been appreciated in cryo-EM and is now
418 being used more routinely for high-resolution structure determination (J. Chen et al., 2019; S.
419 Chen et al., 2022; Li et al., 2021). At a fundamental level, several open questions remain:
420 what causes proteins to adapt to certain preferred views? How do surfactants affect this
421 behaviour? What is the role of other physical properties such as ice thickness, surface tension,
422 and temperature in the behaviour of macromolecules during freezing? To address some of
423 these questions, we systematically tested grid freezing conditions on a set of well-
424 characterised proteins that adapt preferred orientations. The selected proteins had molecular
425 weights ranging from 125 kDa to 466 kDa and varying symmetries (Figure 2). The degree of
426 orientation bias also varied significantly. The CRP decamer and SARS-CoV-2 spike had a
427 more severe orientation bias, with more than 90% abundance for the preferred view, whereas
428 in the case of the CRP pentamer, the abundance was 50%, and in catalase, it was near 60%.
429 Additionally, Spike, PaaZ and β-galactosidase were recombinantly overexpressed and had
430 poly-histidine as the affinity tag at the termini of each monomer, whereas CRP and catalase,
431 which were purified from native sources and did not contain any affinity tags.

432 The air-water interface exhibits an affinity for OH^- ions from water, resulting in a
433 negative charge (Chaplin, 2009; Drzymala et al., 1999). Extensive research has been
434 conducted on the behaviour of proteins and surfactants at the air-water interface and in thin
435 films, spanning the fields of cryo-EM and surface chemistry (Li et al., 2021; Rabe et al., 2020
436 ;Samanta & Ghosh, 2011; Szymczyk et al., 2018). Based on these investigations and an
437 examination of the surface charge distributions of macromolecules in our study, we
438 postulated that electrostatic interactions may play a pivotal role in dictating protein
439 orientations. To visualise the surface charge distributions of the proteins, electrostatic
440 potential maps were calculated by assigning charges using the pdb2pqr server (Jurrus et al.,
441 2018) at a given pH and visualised using ChimeraX (Pettersen et al., 2021). It was observed
442 that the charge on the preferred view was positive in the case of catalase and neutral in CRP-
443 pentamer, CRP-decamer, and PaaZ and negative and neutral in β-galactosidase (Figure S7).
444 The preferred bottom view of the spike is negatively charged, but this charge is likely masked
445 by the presence of a 56 amino acid-long affinity tag from each monomer. Further, the charge

446 on the alternate view of β-galactosidase with tag, which is not seen in the 2D classes is more
447 negative compared to the preferred view (Figure S7C).

448 An interesting sample in our study is the C-reactive protein (CRP) that exists as both
449 pentamer and decamer populations dependent on concentration and, the impact of these
450 populations on each other's orientation distribution is intriguing. The particle numbers of
451 each population of CRP and their relative abundance on the micrographs (after 2D
452 classification) are summarized in Table S1. It's plausible that one population exhibits a
453 greater affinity for the AWI, monopolizing its free space and shielding the other from
454 interaction. However, this scenario doesn't seem applicable to CRP, as both pentamer and
455 decamer populations adopt preferred orientations in the absence of additives indicating their
456 interaction with the AWI. Furthermore, interactions and potential denaturation of a specific
457 population at the AWI can alter the solution's effective concentration, potentially causing a
458 shift in the pentamer-decamer equilibrium. These factors are important in determining the
459 behaviour of the protein on the grids but for this study, we have not considered these effects
460 and analysed both the pentamer and decamer populations independently. The negatively
461 charged B-face of CRP in the absence of additive is (shown in Figure S7B) less sampled in
462 the case of the CRP pentamer and not sampled at all in the CRP decamer. When the surface
463 charge at the AWI was changed to positive by the addition of cationic surfactant CTAB, the
464 orientations changed in almost all cases, with a slight preference for certain alternate views
465 (Figure 3). Hence, electrostatic interactions at the AWI are an important factor in determining
466 protein orientations in the absence of any additive. When non-ionic surfactants Tween 20 and
467 Tween 80 are added to the sample buffer, the observed orientation distributions are uniform
468 (Figure 3). Among the non-ionic surfactants, the orientation plots are similar, as they all
469 make the AWI charge neutral by occupying the free surface and work by diminishing the
470 charge-based interactions between the protein and the AWI.

471 Although CTAB addition resulted in isotropic maps, the coverage of the 3D Fourier
472 space was the least when compared with other surfactants (Table 2). The CTAB datasets
473 showed a preference for one or more views, but in the macromolecules tested, these
474 preferences were desirable. Upon addition of SLS, an even orientation distribution was
475 observed for the CRP pentamer, CRP decamer, and catalase. Tween 20 and Tween 80 had
476 similar effects on orientation, except in the case of the CRP pentamer. A8-35 had an
477 intermediate orientation distribution, which led to isotropic maps (Figure 3). In the order of
478 homogeneity of the orientation distributions, the surfactants can be ranked as SLS > Tween

479 20 = Tween 80 > A8-35 = CTAB. The map vs. model FSCs were calculated for the datasets
480 to obtain an independent resolution estimate, and they closely match the resolution estimated
481 by the comparison of the half-maps in most datasets (Figure S8 and Table S2). Further, the
482 use of different surfactants or sample preparation at different temperatures does not affect the
483 structure of catalase as shown by the density of heme in catalase (Figure S9).

484 Poly-histidine tag-dependent changes in orientation have been reported recently by
485 Bromberg et al., 2022 but the mechanism by which this occurs remains unclear. The addition
486 of poly-histidine tag is a convenient method for expressing and isolating proteins (Hochuli et
487 al., 1988; Merchant et al., 1998). Unlike crystallography, it is intuitive to think that the
488 flexibility of histidine tags does not pose a problem in structure determination by cryo-EM as
489 the molecules are imaged individually, and the tag is averaged out during the reconstruction
490 process if it is present in random conformations. At pH 8, which is more often used as the
491 sample buffer, histidine should be deprotonated and exist as a hydrophilic polar group, and
492 why this should cause severe preference in orientations is not clear to us. The
493 microenvironment at the AWI is complex and could be interacting with the histidine tag
494 favourably and stabilizing it in that orientation.

495 Among the samples studied here, spike, PaaZ and β-galactosidase had a poly-histidine
496 tag at the termini of the monomers. While the presence of the tag had no detrimental effect on
497 the orientation distribution and reconstruction of PaaZ, it affected the orientations of the
498 spike and β-galactosidase proteins significantly, and the addition of surfactants was not
499 sufficient to mitigate the bias in the case of spike. In the case of PaaZ, the presence of a poly-
500 histidine tag also causes a preference for the view that contains the affinity tag but the side
501 view of the protein which contributes maximally to the reconstruction and hence even with
502 very little sampling of alternate views the reconstruction is not as anisotropic as spike. We
503 note that the his-tag might cause the clumping observed in PaaZ and this remains to be tested.
504 Further, the use of D3 symmetry in PaaZ and D2 symmetry in β-galactosidase as compared to
505 C1 symmetry in the spike contribute to the quality of the final reconstruction. As the spike
506 protein and β-galactosidase have been extensively studied and many structures have been
507 reported we wondered what could be the reason that we see such severe orientation bias and
508 others do not. From the published structural studies on the SARS-CoV2 spike it was realised
509 that there have been variations in the reported specimens used for cryo-EM in terms of the
510 expression system, detergents added to the sample, cleavage of tags, buffers, etc (Bangaru et
511 al., 2020; Bodakuntla et al., 2023; H. Cheng et al., 2023; Hardenbrook & Zhang, 2022;

Wrapp et al., 2020; Wrobel, 2023). The variability in the spike protein's behaviour is outlined in the 2022 review authored by Chua et al., 2022. They provide instances of the spike protein showing a preference for very thick ice (100 nm), displaying unfavourable behaviour on gold grids, and exhibiting notably low particle concentrations on grids in the absence of a detergent (Chua et al., 2022). Other parameters that might affect the orientation distribution include the glow discharge settings for grids that change the hydrophilicity of the grid surface, grid type (for example, Quantifoil with carbon support or Ultrafoil with gold support), grid hole diameter (0.6/1, 1.2/1.3, or 2/2), freezing temperature, and blotting duration. In our case, the removal of the affinity tag and CTAB addition to the sample buffer was required to sample other views of spike protein, which resulted in a map of reasonable resolution and improved isotropy (Figure 5B). In the case of β-galactosidase, the removal of the affinity tag was sufficient to obtain a high-resolution isotropic without the need for any additive (Figure 5B). We note that β-galactosidase with a C-terminal tag also showed preferred orientation in ice as recently reported by Esfahani et al., 2024. Thus, the effect of the poly-histidine tag is a key factor to be tested when faced with orientation bias.

Furthermore, physical factors, such as the temperature of grid freezing and grid type, can affect the mechanics of thin film formation and evaporation/blotting and, in turn, influence protein behaviour. We show that the temperature of grid freezing can change the orientation distributions of proteins by affecting the ice thickness and protein-protein interactions (Figure 5). Additionally, varying only the hole diameter in a holey grid and keeping other factors constant led to a change in the preference of views in the case of catalase (Figure S2). In Quantifoil 1.2/1.3 grids, the approximate ratio of the area occupied by the hole to carbon is 38:62%, whereas, in 0.6/1 grids, the ratio decreases to 30:70%. This change can affect protein distribution in the holes, carbon, and AWI. It is now evident that this also significantly affects the orientation distribution of macromolecules in the thin films formed in the holes. We note two caveats in this study: a) all the grids used were from Quantifoil with amorphous carbon as the support film, and it is unclear if similar behaviour will be observed with C-flat grids from Protochips that are manufactured differently or other grids such as UltraFoil or NiTi alloy (Fan et al., 2021; Quispe et al., 2007; Russo & Passmore, 2016; Schürmann et al., 2017). b) only one sample (catalase) was tested and further extensive study with more samples is required to generalise the effect of the holey grid geometry.

544 Previous studies have used surfactants to determine the structures of the proteins
545 examined in this study. A comparison of these results with our observations is presented
546 below. Catalase has been used as a standard test sample in the context of the preferred
547 orientation problem in two independent studies, which take different approaches to solve the
548 problem. In an earlier study, several high-CMC detergents were screened, among which
549 CHAPSO resulted in the best orientation distribution and map quality, although a high
550 protein concentration of 30-40 mg/ml was used to obtain a good distribution of particles and
551 a small data set was sufficient to obtain a 2.2 Å reconstruction by averaging 119,000 particles
552 with an E_{od} of 0.76 (S.Chen et al., 2022). Furthermore, the effects of ionic strength and pH
553 were tested and were observed to have little or no effect on the protein orientation (S.Chen et
554 al., 2022). Our results, in extension to this study, provide an understanding of the effects of
555 surfactant headgroup charge, temperature of grid freezing, and hole carbon geometry on the
556 orientations of catalase (Figure 3, 5 & S2). In another study, nickel/titanium grids coated with
557 2D crystals of the hydrophobic protein HFBI (which shields the protein from its exposure to
558 AWI) and 2.3 mg/ml of catalase were used for preparing specimen and subsequent data
559 collection and structure determination (Fan et al., 2021). The resulting catalase map has a
560 resolution of 2.3 Å from 169,897 particles with an E_{od} of 0.80, which is comparable to our
561 datasets (Table 2).

562 Similarly, multiple structures of CRP pentamer and decamer were reported in a study
563 by (Noone et al., 2021), which focused on the effects of pH and ligand addition on the
564 complement binding properties of CRP. One of the grid freezing conditions reported was
565 CRP apoprotein at pH 7.5 with 0.05% Tween 20 added and the grid held at 4 °C and 65%
566 humidity. These freezing conditions are similar to those described here, except for the
567 concentration of Tween 20 used and the temperature and humidity during freezing. The
568 orientations observed in this case look similar, and the resulting reconstructions from 256,289
569 and 204,354 particles for the pentamer and decamer (C5 and C1 symmetry) resulted in
570 resolutions of 3.2 Å and 2.8 Å, respectively. In comparison, we have obtained a 3.3 Å map
571 for the pentamer and a 4 Å map for the decamer by averaging 25,000 and 51,000 particles,
572 respectively (Table 2). In the field of cryo-EM, the biggest variation is observed in sample
573 preparation, sometimes using the same instrument. Therefore, it is reassuring to observe that
574 other independent investigations (along with different grid freezing procedures) of catalase
575 and CRP yield 3D reconstructions of comparable quality.

The air–water interface is a complex environment; the properties of the bulk and the surface vary drastically, and it is an active area of research in surface chemistry and aerosol fields (Martins-Costa & Ruiz-López, 2023; Nguyen et al., 2020; Zhong et al., 2019). Only the effects that occur at the surface can be observed, and trends can be analysed to improve the sample preparation methods. The explanation as to why they occur cannot be understood using standard cryo-EM experiments alone and requires complementary approaches such as infrared spectroscopy or dynamic surface tension measurement studies (Carter-Fenk et al., 2021; F. Tang et al., 2020). Alternative approaches to sample preparation have been proposed in recent years to overcome the preferred orientation issue (Esfahani et al., 2024 ;Drulyte et al., 2018; Glaeser & Han, 2017; Jain et al., 2012). One such approach was recently reported by Huber et al., where the sample was filled with nanosized capillaries made of silicon-rich nitride membranes embedded on a chip that physically controls the ice thickness. No blotting is involved in this process, and some test samples have demonstrated the potential of this method (Huber et al., 2022). However, the air-water interface is replaced by the solid-water interface, which may contribute to some effects of its own. Further use and standardisation of such methods are required to understand their general applicability.

4. Conclusion and outlook

One of the goals of single particle cryo-EM is to collect a smaller data set and average minimum number of particles to obtain a high-resolution reconstruction (Henderson, 1995). However, in reality, many factors affect this including the sample heterogeneity, detectors, beam-induced motion etc., and the recently observed effect of AWI and preferred orientation can be added to this list. We performed a comprehensive examination of how macromolecule orientations respond to alterations in physical factors, such as freezing temperature, and chemical factors, such as the addition of surfactants or the presence of affinity tags, during grid freezing for a standard set of proteins. This analysis provides insights into the behaviour of proteins on grids and can be utilised to address the preferred orientation problem systemically for any given macromolecule. When using surfactants, it is crucial to carefully assess protein stability using other techniques such as native gel electrophoresis, differential scanning fluorimetry, or size-exclusion chromatography before data collection to save time and resources. Furthermore, our findings highlight the necessity to innovate and create small-molecule additives that are inert to biological samples and can effectively occupy the AWI and alleviate its deleterious effects on proteins.

608 **5. Materials and Methods**

609 **Source of proteins:**

610 Human C-reactive protein (Cat. No. C4063) and Human erythrocyte catalase (Cat. No.
611 C3556) were obtained from Sigma Aldrich chemicals. The protein samples were either
612 concentrated using an Amicon 100kDa concentrator or diluted in respective buffers for grid
613 freezing. All the detergent stocks were made in ultrapure water and dilutions were made and
614 used on the day of the experiment.

615 PaaZ was expressed and purified as described in (Sathyanarayanan et al., 2019).

616 SARS-CoV-2 S plasmid was a kind gift from the Krammer laboratory at Icahn School of
617 Medicine at Mount Sinai. The spike gene was amplified from the plasmid and subcloned in
618 the BacMam vector with a C-terminus HRV-3C cleavage tag followed by a 7x poly-histidine
619 and twin strep tag. Bacmid DNA and virus were prepared as described in the Invitrogen bac-
620 to-bac manual. After two generations of amplification in Sf9 cells, the V2 virus was used for
621 transfection of HEK293F cells at 2 million/ml density. Sodium butyrate (4 mM) was added to
622 enhance the production of protein 8 hours post-infection. The media supernatant containing
623 the secreted spike protein was harvested on day 3 by centrifuging the cells at 150 g for 10
624 minutes. The media was incubated with pre-equilibrated Ni-NTA beads (Qiagen) at room
625 temperature for 1-2 hours (1 ml beads/200 ml media). Ni-NTA beads were washed with PBS
626 buffer containing 20 mM imidazole followed by elution with 280 mM imidazole in PBS
627 buffer. The eluted protein was run on SDS-PAGE to assess purity and further concentrated
628 and injected on a 24 ml column Superdex 200 size exclusion to exchange the buffer to 50
629 mM Tris pH 8, 200 mM NaCl,1 mM DTT. To cleave the tag, the eluted fractions from Ni-
630 NTA chromatography were diluted with 50 mM Tris pH 8, 200 mM NaCl, and 1 mM DTT
631 buffer and incubated with HRV3C protease overnight at 4°C, followed by reverse IMAC to
632 obtain the spike protein without tag in the flow through. The flow-through was concentrated
633 using an Amicon 100 kDa concentrator and flash-frozen using liquid nitrogen and stored at -
634 80°C until further use.

635 β-galactosidase with N-terminal 6X histidine tag followed by a thrombin cleavage tag was a
636 gift from Prof. Doug Juers, which was cloned in the pET15b vector and was transformed in
637 *E. coli.* Strain JM109 (Juers et al., 2000). The cells from glycerol stock were patched on an
638 LB agar with ampicillin and allowed to grow overnight at 37°C. A single colony was picked
639 and allowed to grow in LB with ampicillin overnight. The next day 4 litres of LB with 100

640 μg/ml ampicillin was inoculated and OD_{600} was monitored every 30 minutes. The expression
641 of protein was induced by adding 1 mM IPTG when the OD_{600} of the cells reached 0.6 and
642 the cells were allowed to grow for 4-5 hours at 37°C. Subsequently, cells were harvested by
643 centrifugation at 4000g for 20 minutes and the cell pellet was flash-frozen and stored at -
644 80°C until further use. On the day of purification, the cells were resuspended in ~50 ml of
645 resuspension buffer containing 20 mM Tris pH 8, 500 mM NaCl, 5 mM Imidazole, 1 mM β-
646 mercaptoethanol (β-ME), 1 mM $MgCl_2$ and sonicated for 5 minutes (5 secs on 10 secs off) at
647 40% amplitude. The lysate was clarified by centrifugation at 18000 g for 45 minutes and the
648 supernatant was collected. A 5 ml Ni-NTA column was equilibrated with 5 CV resuspension
649 buffer followed by sample application of the supernatant at a flow rate of 1 ml/min. After the
650 binding of the protein, the column was washed with 50 CV of resuspension buffer followed
651 by a linear gradient elution with buffer B that consisted of resuspension buffer with 500 mM
652 imidazole and 1.8 ml fractions were collected. The fractions from Ni-NTA elution were
653 analyzed using SDS-PAGE and the fractions from the last half of the peak were pooled and
654 dialyzed with 2x4L of dialysis buffer containing 25 mM Tris pH 8, 125 mM NaCl, 2.5 mM
655 $CaCl_2$, 2 mM β-ME. The protein was concentrated using a 100 kDa cutoff Amicon
656 concentrator, aliquoted and stored at -80° C. An aliquot of this protein was thawed and
657 injected into a Superdex 200 column and the buffer was exchanged to 100 mM Tris pH8, 200
658 mM NaCl, 2.5 mM $MgCl_2$, 5 mM $CaCl_2$, 2 mM β-ME. The cleavage of the tag was carried
659 out by thrombin protease at room temperature for 4 hours followed by reverse IMAC to
660 collect the protein without tag in flow-through. The protein without tag was concentrated and
661 grids were prepared.

662 **Grid preparation:**

663 6.3 μl of the protein was thawed on ice and 0.7 μl of 10 x additive (surfactant) stock was
664 added to get a final concentration of 1 x. This sample was incubated on ice for 2-5 mins and
665 then centrifuged at 21,000g for 20 mins. Meanwhile, Vitrobot Mark IV (ThermoFisher
666 Scientific) chamber was equilibrated at 20° C (unless stated otherwise) and 100% humidity.
667 Quantifoil 1.2/1.3 or Quantifoil 0.6/1 grids were glow discharged in a reduced air
668 environment using a PELCO EasiGlow chamber using a standard setting of 25 mA current
669 for 1 minute. The grid was mounted on the Vitrobot Mark IV (ThermoFisher Scientific) and 3
670 μl of the sample was applied to the grid. A blotting time of 3-4 seconds, a wait time of 10
671 seconds and a blot force of 0 were used to obtain a thin film of the specimen. For the datasets
672 where grids were prepared at different temperatures, the protein was incubated on a thermal

673 block at the required temperature for 3-7 minutes before applying it to the grid. The Vitrobot
674 chamber was maintained at the required temperature and 100% humidity. The blot time, blot
675 force and wait time were kept constant.

676 **Grid screening and data collection:**

677 Grids were screened on Titan Krios microscope operating at 300 kV using standard low-dose
678 settings and the automated data collection was set up either on the Falcon 3 or Gatan K2
679 detectors in counting mode with EPU software (ThermoFisher Scientific). Magnification of
680 59,000 x was used only for the catalase 37 °C dataset with a pixel size of 1.38 Å (and PaaZ
681 enzyme incubated at 37°C with no additive but no data was collected). For all the other
682 datasets collected on Falcon3, a magnification of 75,000 x corresponding to a pixel size of
683 1.07 Å and a dose of approximately 1 $e^-/Å^2$/frame were used and movies of 24-25 frames
684 were collected. The datasets on the K2 detector (Gatan) were collected in EFTEM mode with
685 a slit width of 20 eV and at a nominal magnification of 130,000 x corresponding to a pixel
686 size of 1.08 Å. The dose was 5 e^-/p/s and an exposure of 8 seconds was used. Each movie was
687 fractionated into 32 frames with a total dose of ~36.28 $e^-/Å^2$. The grid preparation condition
688 and data parameters are summarized in Table 3.

689 **Table 3.** Summary of parameters of datasets collected under different conditions. All the imaging
690 was performed in counting mode

Protein	Condition	Grid type (Quantifoil)	Buffer composition	Protein concentration	Detector	Box size (pixels)	Pixel size
CRP pentamer and decamer	No additive	1.2/1.3	20 mM tris pH 8, 280 mM NaCl, 5 mM $CaCl_2$, 0.03% NaN_3	2.1 mg/ml	Falcon3	256	1.07
	CTAB	1.2/1.3		2.6 mg/ml	Falcon3	256	1.07
	SLS	1.2/1.3		2.6 mg/ml	K2	320	1.08
	Tween 20	1.2/1.3		3.6 mg/ml	K2	320	1.08
	Tween 80	1.2/1.3		6.8 mg/ml	K2	320	1.08
	A8-35	1.2/1.3		3.6 mg/ml	Falcon3	320	1.07
Catalase	No additive, 20 °C	1.2/1.3	50 mM Tris pH 8	0.625 mg/ml	Falcon3	256	1.07
	CTAB	1.2/1.3		3.4 mg/ml	K2	320	1.08
	SLS	1.2/1.3		3.4 mg/ml	Falcon3	256	1.07
	Tween 20	1.2/1.3		3.4 mg/ml	K2	320	1.08
	Tween 80	1.2/1.3		4.1 mg/ml	Falcon3	320	1.07
	A8-35	1.2/1.3		3.4 mg/ml	Falcon3	320	1.07
	4°C	1.2/1.3		0.625 mg/ml	Falcon3	256	1.07
	37°C	1.2/1.3		0.625 mg/ml	Falcon3	320	1.38
	4°C	0.6/1		0.625 mg/ml	Falcon3	320	1.07
	20°C	0.6/1		0.625 mg/ml	Falcon3	256	1.07
PaaZ	No additive, 4°C	0.6/1	25 mM HEPES pH 7.4, 50 mM NaCl	0.8 mg/ml	Falcon3	320	1.07
	No additive, 20°C	0.6/1		0.8 mg/ml	Falcon3	256	1.07
	No additive, 37°C[1]	0.6/1		0.8 mg/ml	Falcon3	-	1.38
	CTAB, 4°C	0.6/1		0.8 mg/ml	Falcon3	256	1.07
Spike	With tag no additive	0.6/1	50 mM Tris pH 8, 200 mM NaCl, 1 mM DTT	1 mg/ml	Falcon3	256	1.07
	With tag with CTAB	0.6/1		1.3 mg/ml	Falcon3	320	1.07
	Without tag no additive	0.6/1		2 mg/ml	Falcon3	256	1.07
	Without tag with CTAB	0.6/1		2 mg/ml	Falcon3	256	1.07
β-galactosidase	With tag no additive	0.6/1	100 mM Tris pH 8, 200 mM NaCl, 5 mM $CaCl_2$, 2.5 mM $MgCl_2$, 2 mM β-ME	5 mg/ml	Falcon3	320	1.07
	Without tag no additive	0.6/1		5 mg/ml	Falcon3	320	1.07

691 [1] – no data was collected due to higher order structures of particles

692 **Data processing and model refinement:**

693 RELION 3.1. (Scheres, 2012, Zivanov et al., 2018) was used to process all the data with a
694 standard workflow which is summarized as follows. The multi-frame movies were summed
695 and corrected for beam-induced motion using Relion's in-built algorithm Motioncorr. The
696 resulting summed micrograph images were used to estimate the contrast transfer function
697 (CTF) using GCTF (Zhang, 2016). The micrographs after CTF correction were used to pick
698 particles using either reference-free methods such as the LoG-based picking in Relion or
699 Gautomatch (https://www2.mrc-lmb.cam.ac.uk/download/gautomatch-053/) or using
700 previously obtained 2D templates. After picking, the particles were boxed and extracted using
701 the box sizes of 256 or 320 pixels (Table 3). These particles were then subjected to multiple
702 rounds of iterative 2D classifications to obtain good-quality, high-resolution classes. A low-
703 resolution initial model was either generated from the same dataset or a model from pre-
704 existing dataset was used as a reference for 3D refinement. The selected good-quality
705 particles were then used for either 3D classification or 3D refinement based on the data
706 (Scheres, 2012). If the resolution at this stage was below 4.5 Å, CTF refinement and
707 Bayesian polishing were performed on the final particles to correct for beam tilt, anisotropic
708 magnification, per-particle defocus and the effects of beam-induced motion respectively
709 (Zivanov et al., 2018, 2019). Another round of 3D refinement was performed after this step.
710 Further, postprocessing was done to sharpen the final map and estimate the final global
711 resolution (Rosenthal & Henderson, 2003).
712 PaaZ-CTAB and Spike-CTAB datasets were subjected to multiple rounds of 3D
713 classification. For most datasets (all except CRP pentamer), as a part of standard workflow
714 the resulting map after reconstruction had the wrong hand and the maps were flipped to
715 obtain the correct hand before the model was fit and refinement was performed. The 3D FSC
716 server was used to calculate the sphericity of the final maps (Tan et al., 2018). The E_{od} was
717 calculated using the cryoEF by using a subset of 1000 angles and the appropriate symmetries
718 for respective molecules (Naydenova & Russo, 2017). ChimeraX (Pettersen et al., 2004) was
719 used for generating the electrostatic potential maps of the structures. The Pymol APBS plugin
720 was used to generate the PaaZ electrostatic potential map (Schrödinger, LLC, 2015). EMAN2
721 (G. Tang et al., 2007)and Phenix (Liebschner et al., 2019) were used to generate the map vs.
722 model FSCs. The final model refinement was done in real space for the PaaZ and CRP
723 datasets using Phenix (Afonine et al., 2018) and in reciprocal space using Refmac Servalcat
724 (Murshudov et al., 2011; Yamashita et al., 2021) for catalase and spike datasets. Initial water
725 picking in the PaaZ dataset was performed with Coot (Emsley & Cowtan, 2004) and
726 manually inspected to remove those that were not hydrogen bonded to amino acid residues,

also using the fofc map from servalcat (Yamashita et al., 2021) and a conservative approach was used so that noise was not modelled. Figures were made with Chimera and Pymol (Pettersen et al., 2004; Schrödinger, LLC, 2015).

Acknowledgements: We acknowledge the National Cryo-EM facility, Bangalore, for data collection, which was supported by the Department of Biotechnology, DBT/PR12422/MED/31/287/2014, and the computing facility in the Bangalore Life Science Cluster. We thank Prof. Ramaswamy S and all the lab members for critical reading of the manuscript. SY acknowledges the graduate fellowship from TIFR/NCBS. KRV acknowledges the support of the Department of Atomic Energy, Government of India, Government of India, under Project Identification No. RTI4006. KRV is part of the EMBO Global Investigator Network. KRV acknowledges the discussion with Drs Pamela Williams and Judith Reeks, Astex, UK on β-galactosidase and the effect of tag.

Conflict of interest: The authors declare no conflict of interest.

Data Availability: The cryo-EM maps and the coordinates have been deposited in the EMDB and PDB respectively with the following accession codes.

CRP Pentamer with CTAB – EMD-37864 and PDB 8WV4

CRP decamer with CTAB – EMD-37865 and PDB ID 8WV5

PaaZ with CTAB at 4 °C – EMD-37866 and PDB ID 8WV6

Spike with CTAB – EMD-37953 and PDB ID 8WZI

Catalase with CTAB – EMD-37956 and PDB ID 8WZM

Catalase with SLS – EMD-37952 and PDB ID 8WZH

Catalase at 4 °C – EMD-37955 and PDB ID 8WZK

Catalase at 20 °C – EMD-37954 and PDB ID 8WZJ

β-galactosidase, with tag – EMD- 39809

β-galactosidase, with no tag – EMD- 39808

References

Adrian, M., Dubochet, J., Lepault, J., & McDowall, A. W. (1984). Cryo-electron microscopy of viruses. *Nature*, *308*(5954), 32–36. https://doi.org/10.1038/308032a0

Afonine, P. V., Poon, B. K., Read, R. J., Sobolev, O. V., Terwilliger, T. C., Urzhumtsev, A., & Adams, P. D. (2018). Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallographica Section D: Structural Biology*, *74*(6), 531–544. https://doi.org/10.1107/S2059798318006551

Aiyer, S., Baldwin, P. R., Tan, S. M., Shan, Z., Oh, J., Mehrani, A., Bowman, M. E., Louie, G., Passos, D. O., Selena, Đ., Hull, J. A., Hoshika, S., Barad, B. A., Grotjahn, D. A., Agbandje-mckenna, M., Benner, S. A., & Noel, J. A. P. (2023). *Overcoming Resolution Attenuation During Tilted Cryo-EM Data Collection*.

Baek, M., DiMaio, F., Anishchenko, I., Dauparas, J., Ovchinnikov, S., Lee, G. R., Wang, J., Cong, Q., Kinch, L. N., Dustin Schaeffer, R., Millán, C., Park, H., Adams, C., Glassman, C. R., DeGiovanni, A., Pereira, J. H., Rodrigues, A. V., Van Dijk, A. A., Ebrecht, A. C., … Baker, D. (2021). Accurate prediction of protein structures and interactions using a three-track neural network. *Science*, *373*(6557), 871–876. https://doi.org/10.1126/science.abj8754

Bai, X. (2021). Seeing Atoms by Single-Particle Cryo-EM. *Trends in Biochemical Sciences*, *46*(4), 253–254. https://doi.org/https://doi.org/10.1016/j.tibs.2021.01.001

Baker, M. R., Fan, G., Seryshev, A. B., Agosto, M. A., Baker, M. L., & Serysheva, I. I. (2021). Cryo-EM structure of type 1 IP3R channel in a lipid bilayer. *Communications Biology*, *4*(1). https://doi.org/10.1038/s42003-021-02156-4

Bangaru, S., Ozorowski, G., Turner, H. L., Antanasijevic, A., Huang, D., Wang, X., Torres, J. L., Diedrich, J. K., Tian, J. H., Portnoff, A. D., Patel, N., Massare, M. J., Yates, J. R., Nemazee, D., Paulson, J. C., Glenn, G., Smith, G., & Ward, A. B. (2020). Structural analysis of full-length SARS-CoV-2 spike protein from an advanced vaccine candidate. *Science*, *370*(6520), 1089–1094. https://doi.org/10.1126/science.abe1502

Bartesaghi, A., Aguerrebere, C., Falconieri, V., Banerjee, S., Earl, L. A., Zhu, X., Grigorieff, N., Milne, J. L. S., Sapiro, G., Wu, X., & Subramaniam, S. (2018). Atomic Resolution Cryo-EM Structure of β-Galactosidase. *Structure*, *26*(6), 848-856.e3. https://doi.org/10.1016/j.str.2018.04.004

Bodakuntla, S., Kuhn, C. C., Biertümpfel, C., & Mizuno, N. (2023). Cryo-electron microscopy in the fight against COVID-19—mechanism of virus entry. *Frontiers in Molecular Biosciences*, *10*(October), 1–11. https://doi.org/10.3389/fmolb.2023.1252529

Bromberg, R., Cai, K., Guo, Y., Plymire, D., Emde, T., Puzio, M., Borek, D., & Otwinowski, Z. (2022). The His-tag as a decoy modulating preferred orientation in cryoEM. *Frontiers in Molecular Biosciences*, *9*(October), 1–13. https://doi.org/10.3389/fmolb.2022.912072

Campbell, M. G., Veesler, D., Cheng, A., Potter, C. S., & Carragher, B. (2015). 2.8 Å Resolution Reconstruction of the Thermoplasma Acidophilum 20 S Proteasome Using Cryo-Electron Microscopy. *ELife*, *2015*(4), 1–22. https://doi.org/10.7554/eLife.06380

Carter-Fenk, K. A., Carter-Fenk, K., Fiamingo, M. E., Allen, H. C., & Herbert, J. M. (2021). Vibrational exciton delocalization precludes the use of infrared intensities as proxies for surfactant accumulation on aqueous surfaces. *Chem. Sci.*, *12*(24), 8320–8332. https://doi.org/10.1039/D1SC01276B

Chaplin, M. (2009). Theory vs Experiment : What is the Surface Charge of Water ? *Water*, *1*(July), 1–28.

Chen, J., Noble, A. J., Kang, J. Y., & Darst, S. A. (2019). Eliminating effects of particle adsorption to the air/water interface in single-particle cryo-electron microscopy: Bacterial RNA polymerase and CHAPSO. *Journal of Structural Biology: X*, *1*, 100005. https://doi.org/https://doi.org/10.1016/j.yjsbx.2019.100005

Chen, S., Li, J., Vinothkumar, K. R., & Henderson, R. (2022). Interaction of human erythrocyte catalase with air-water interface in cryoEM. *Microscopy*, *71*(August 2021), I51–I59. https://doi.org/10.1093/jmicro/dfab037

Cheng, H., Zheng, L., Liu, N., Huang, C., Xu, J., Lu, Y., Cui, X., Xu, K., Hou, Y., Tang, J., Zhang, Z., Li, J., Ni, X., Chen, Y., Peng, H., & Wang, H. W. (2023). Dual-Affinity Graphene Sheets for High-Resolution Cryo-Electron Microscopy. *Journal of the American Chemical Society*, *145*(14), 8073–8081. https://doi.org/10.1021/jacs.3c00659

Cheng, Y. (2015). Single-Particle Cryo-EM at Crystallographic Resolution. *Cell*, *161*(3), 450–457. https://doi.org/10.1016/j.cell.2015.03.049

Chua, E. Y. D., Mendez, J. H., Rapp, M., Ilca, S. L., Tan, Y. Z., Maruthi, K., Kuang, H., Zimanyi, C. M., Cheng, A., Eng, E. T., Noble, A. J., Potter, C. S., & Carragher, B. (2022). *Better , Faster , Cheaper : Recent Advances in Cryo – Electron Microscopy*. 1–32.

Dandey, V. P., Budell, W. C., Wei, H., Bobe, D., Maruthi, K., Kopylov, M., Eng, E. T., Kahn, P. A., Hinshaw, J. E., Kundu, N., Nimigean, C. M., Fan, C., Sukomon, N., Darst, S. A., Saecker, R. M., Chen, J., Malone, B., Potter, C. S., & Carragher, B. (2020). Time-resolved cryo-EM using Spotiton. *Nature Methods*, *17*(9), 897–900. https://doi.org/10.1038/s41592-020-0925-6

D'Imprima, E., Floris, D., Joppe, M., Sánchez, R., Grininger, M., & Kühlbrandt, W. (2019). Protein denaturation at the air-water interface and how to prevent it. *ELife*, *8*, 1–18. https://doi.org/10.7554/eLife.42747

Drulyte, I., Johnson, R. M., Hesketh, E. L., Hurdiss, D. L., Scarff, C. A., Porav, S. A., Ranson, N. A., Muench, S. P., & Thompson, R. F. (2018). Approaches to altering particle distributions in cryo-electron microscopy sample preparation. *Acta Crystallographica Section D: Structural Biology*, *74*(6), 560–571. https://doi.org/10.1107/S2059798318006496

Drzymala, J., Sadowski, Z., Holysz, L., & Chibowski, E. (1999). Ice/water interface: Zeta potential, point of zero charge, and hydrophobicity. *Journal of Colloid and Interface Science*, *220*(2), 229–234. https://doi.org/10.1006/jcis.1999.6528

Dubochet, J., Adrian, M., Chang, J.-J., Homo, J.-C., Lepault, J., McDowall, A. W., & Schultz, P. (1988). Cryo-electron microscopy of vitrified specimens. *Quarterly Reviews of Biophysics*, *21*(2), 129–228. https://doi.org/DOI: 10.1017/S0033583500004297

Dubochet, J., Adrian, M., Lepault, J., & McDowall, A. W. (1985). Emerging techniques: Cryo-electron microscopy of vitrified biological specimens. *Trends in Biochemical Sciences*, *10*(4), 143–146. https://doi.org/10.1016/0968-0004(85)90150-1

Dubochet, J., Lepault, J., Freeman, R., Berriman, J. A., & Homo, J. -C. (1982). Electron microscopy of frozen water and aqueous solutions. *Journal of Microscopy*, *128*(3), 219–237. https://doi.org/10.1111/j.1365-2818.1982.tb04625.x

Dubochet, J., & McDowall, A. W. (1981). Vitrification of Pure Water for Electron Microscopy. *Journal of Microscopy*, *124*(3), 3–4. https://doi.org/10.1111/j.1365-2818.1981.tb02483.x

Emsley, P., & Cowtan, K. (2004). {\it Coot}: model-building tools for molecular graphics. *Acta Crystallographica Section D*, *60*(12 Part 1), 2126–2132. https://doi.org/10.1107/S0907444904019158

Esfahani, B. G., Randolph, P. S., Peng, R., Grant, T., Stroupe, M. E., & Stagg, S. M. (2024). SPOT-RASTR - a cryo-EM specimen preparation technique that overcomes problems with preferred orientation and the air/water interface. *BioRxiv*. https://doi.org/10.1101/2024.01.24.577038

Fan, H., Wang, B., Zhang, Y., Zhu, Y., Song, B., Xu, H., Zhai, Y., Qiao, M., & Sun, F. (2021). A cryo-electron microscopy support film formed by 2D crystals of hydrophobin

HFBI. *Nature Communications*, *12*(1), 1–13. https://doi.org/10.1038/s41467-021-27596-8

Frederik, P. M., Stuart, M. C. A., Bomans, P. H. H., & Busing, W. M. (1989). Phospholipid, Nature's own slide and cover slip for cryo-electron microscopy. *Journal of Microscopy*, *153*(1), 81–92. https://doi.org/10.1111/j.1365-2818.1989.tb01469.x

Glaeser, R. M., & Han, B.-G. (2017). Opinion: hazards faced by macromolecules when confined to thin aqueous films. *Biophysics Reports*, *3*(1–3), 1–7. https://doi.org/10.1007/s41048-016-0026-3

Hardenbrook, N. J., & Zhang, P. (2022). A structural view of the SARS-CoV-2 virus and its assembly. *Current Opinion in Virology*, *52*, 123–134. https://doi.org/https://doi.org/10.1016/j.coviro.2021.11.011

Harvey, W. T., Carabelli, A. M., Jackson, B., Gupta, R. K., Thomson, E. C., Harrison, E. M., Ludden, C., Reeve, R., Rambaut, A., Peacock, S. J., Robertson, D. L., & Consortium, C.-19 G. U. K. (COG-U. (2021). SARS-CoV-2 variants, spike mutations and immune escape. *Nature Reviews Microbiology*, *19*(7), 409–424. https://doi.org/10.1038/s41579-021-00573-0

He, J., Li, T., & Huang, S.-Y. (2023). Improvement of cryo-EM maps by simultaneous local and non-local deep learning. *Nature Communications*, *14*(1), 3217. https://doi.org/10.1038/s41467-023-39031-1

Henderson, R. (1995). The potential and limitations of neutrons, electrons and X-rays for atomic resolution microscopy of unstained biological molecules. *Quarterly Reviews of Biophysics*, *28*(2), 171–193. https://doi.org/10.1017/S003358350000305X

Hesketh, E. L., Saunders, K., Fisher, C., Potze, J., Stanley, J., Lomonossoff, G. P., & Ranson, N. A. (2018). The 3.3 Å structure of a plant geminivirus using cryo-EM. *Nature Communications*, *9*(1), 1–10. https://doi.org/10.1038/s41467-018-04793-6

Hochuli, E., Bannwarth, W., Döbeli, H., Gentz, R., & Stüber, D. (1988). Genetic Approach to Facilitate Purification of Recombinant Proteins with a Novel Metal Chelate Adsorbent. *Bio/Technology*, *6*(11), 1321–1325. https://doi.org/10.1038/nbt1188-1321

Huang, Y., Yang, C., Xu, X. feng, Xu, W., & Liu, S. wen. (2020). Structural and functional properties of SARS-CoV-2 spike protein: potential antivirus drug development for COVID-19. *Acta Pharmacologica Sinica*, *41*(9), 1141–1149. https://doi.org/10.1038/s41401-020-0485-4

Huber, S. T., Sarajlic, E., Huijink, R., Weis, F., Evers, W. H., & Jakobi, A. J. (2022). Nanofluidic chips for cryo-EM structure determination from picoliter sample volumes. *ELife*, *11*, 1–25. https://doi.org/10.7554/eLife.72629

Jain, T., Sheehan, P., Crum, J., Carragher, B., & Potter, C. S. (2012). Spotiton: A prototype for an integrated inkjet dispense and vitrification system for cryo-TEM. *Journal of Structural Biology*, *179*(1), 68–75. https://doi.org/10.1016/j.jsb.2012.04.020

Juers, D. H., Jacobson, R. H., Wigley, D., Zhang, X., Huber, R. E., Tronrud, D. E., & Matthews, B. W. (2000). High resolution refinement of β-galactosidase in a new crystal form reveals multiple metal-binding sites and provides a structural basis for α-complementation. *Protein Science*, *9*(9), 1685–1699. https://doi.org/10.1110/ps.9.9.1685

Juers, D. H., Matthews, B. W., & Huber, R. E. (2012). LacZ β-galactosidase: Structure and function of an enzyme of historical and molecular biological importance. *Protein Science*, *21*(12), 1792–1807. https://doi.org/10.1002/pro.2165

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., … Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, *596*(7873), 583–589. https://doi.org/10.1038/s41586-021-03819-2

Jurrus, E., Engel, D., Star, K., Monson, K., Brandi, J., Felberg, L. E., Brookes, D. H., Wilson, L., Chen, J., Liles, K., Chun, M., Li, P., Gohara, D. W., Dolinsky, T., Konecny, R., Koes, D. R., Nielsen, J. E., Head-Gordon, T., Geng, W., … Baker, N. A. (2018). Improvements to the APBS biomolecular solvation software suite. *Protein Science*, *27*(1), 112–128. https://doi.org/10.1002/pro.3280

Kampjut, D., Steiner, J., & Sazanov, L. A. (2021). Cryo-EM grid optimization for membrane proteins. *IScience*, *24*(3), 102139. https://doi.org/10.1016/j.isci.2021.102139

Ko, T.-P., Safo, M. K., Musayev, F. N., Di Salvo, M. L., Wang, C., Wu, S.-H., & Abraham, D. J. (2000). Structure of human erythrocyte catalase. *Acta Crystallographica Section D*, *56*(2), 241–245. https://doi.org/10.1107/S0907444999015930

Li, B., Zhu, D., Shi, H., & Zhang, X. (2021). Effect of charge on protein preferred orientation at the air–water interface in cryo-electron microscopy. *Journal of Structural Biology*, *213*(4), 107783. https://doi.org/10.1016/j.jsb.2021.107783

Liebschner, D., Afonine, P. V., Baker, M. L., Bunkoczi, G., Chen, V. B., Croll, T. I., Hintze, B., Hung, L. W., Jain, S., McCoy, A. J., Moriarty, N. W., Oeffner, R. D., Poon, B. K., Prisant, M. G., Read, R. J., Richardson, J. S., Richardson, D. C., Sammito, M. D., Sobolev, O. V., … Adams, P. D. (2019). Macromolecular structure determination using X-rays, neutrons and electrons: Recent developments in Phenix. *Acta Crystallographica Section D: Structural Biology*, *75*, 861–877. https://doi.org/10.1107/S2059798319011471

Liu, N., & Wang, H. W. (2023). Better Cryo-EM Specimen Preparation: How to Deal with the Air–Water Interface? *Journal of Molecular Biology*, *435*(9), 167926. https://doi.org/10.1016/j.jmb.2022.167926

Lu, Y., Liu, N., Liu, Y., Zheng, L., Yang, J., Wang, J., Jia, X., Zi, Q., Peng, H., Rao, Y., & Wang, H.-W. (2022). Functionalized graphene grids with various charges for single-particle cryo-EM. *Nature Communications*, *13*(1), 1–8. https://doi.org/10.1038/s41467-022-34579-w

Lyumkis, D. (2019). Challenges and opportunities in cryo-EM single-particle analysis. *Journal of Biological Chemistry*, *294*(13), 5181–5197. https://doi.org/10.1074/jbc.REV118.005602

Martins-Costa, M. T. C., & Ruiz-López, M. F. (2023). Electrostatics and Chemical Reactivity at the Air-Water Interface. *Journal of the American Chemical Society*, *145*(2), 1400–1406. https://doi.org/10.1021/jacs.2c12089

Merchant, a M., Zhu, Z., Yuan, J. Q., Goddard, a, Adams, C. W., Presta, L. G., & Carter, P. (1998). 1998 Nature Publishing Group http://www.nature.com/naturebiotechnology. *Group*, *16*(10), 291–294. https://doi.org/10.1038/nbt0898-773

Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F., & Vagin, A. A. (2011). REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallographica Section D: Biological Crystallography*, *67*(4), 355–367. https://doi.org/10.1107/S0907444911001314

Naydenova, K., & Russo, C. J. (2017). Measuring the effects of particle orientation to improve the efficiency of electron cryomicroscopy. *Nature Communications*, *8*(1), 8–12. https://doi.org/10.1038/s41467-017-00782-3

Nguyen, C. V., Nakahara, H., & Phan, C. M. (2020). Surface potential of the air/water interface. *Journal of Oleo Science*, *69*(6), 519–528. https://doi.org/10.5650/jos.ess20024

Noble, A. J., Dandey, V. P., Wei, H., Brasch, J., Chase, J., Acharya, P., Tan, Y. Z., Zhang, Z., Kim, L. Y., Scapin, G., Rapp, M., Eng, E. T., Rice, W. J., Cheng, A., Negro, C. J., Shapiro, L., Kwong, P. D., Jeruzalmi, D., Georges, A. des, … Carragher, B. (2018). Routine single particle CryoEM sample and grid characterization by tomography. *ELife*, *7*, 1–42. https://doi.org/10.7554/eLife.34257

Noble, A. J., Wei, H., Dandey, V. P., Zhang, Z., Tan, Y. Z., Potter, C. S., & Carragher, B. (2018). Reducing effects of particle adsorption to the air–water interface in cryo-EM. *Nature Methods*, *15*(10), 793–795. https://doi.org/10.1038/s41592-018-0139-3

Nogales, E. (2015). The development of cryo-EM into a mainstream structural biology technique. *Nature Methods*, *13*(1), 24–27. https://doi.org/10.1038/nmeth.3694

Noone, D. P., van der Velden, T. T., & Sharp, T. H. (2021). Cryo-Electron Microscopy and Biochemical Analysis Offer Insights Into the Effects of Acidic pH, Such as Occur During Acidosis, on the Complement Binding Properties of C-Reactive Protein. *Frontiers in Immunology*, *12*(December), 1–14. https://doi.org/10.3389/fimmu.2021.757633

Okemefuna, A. I., Stach, L., Rana, S., Zlai Buetas, A. J., Gor, J., & Perkins, S. J. (2010). C-reactive protein exists in an NaCl concentration-dependent pentamer-decamer equilibrium in physiological buffer. *Journal of Biological Chemistry*, *285*(2), 1041–1052. https://doi.org/10.1074/jbc.M109.044495

Patwardhan, A. (2017). Trends in the Electron Microscopy Data Bank (EMDB). *Acta Crystallographica Section D: Structural Biology*, *73*(6), 503–508. https://doi.org/10.1107/S2059798317004181

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., & Ferrin, T. E. (2004). UCSF Chimera - A visualization system for exploratory research and analysis. *Journal of Computational Chemistry*, *25*(13), 1605–1612. https://doi.org/10.1002/jcc.20084

Pettersen, E. F., Goddard, T. D., Huang, C. C., Meng, E. C., Couch, G. S., Croll, T. I., Morris, J. H., & Ferrin, T. E. (2021). UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Science*, *30*(1), 70–82. https://doi.org/https://doi.org/10.1002/pro.3943

Pintilie, G., Zhang, K., Su, Z., Li, S., Schmid, M. F., & Chiu, W. (2020). Measurement of atom resolvability in cryo-EM maps with Q-scores. *Nature Methods*, *17*(3), 328–334. https://doi.org/10.1038/s41592-020-0731-1

Quispe, J., Damiano, J., Mick, S. E., Nackashi, D. P., Fellmann, D., Ajero, T. G., Carragher, B., & Potter, C. S. (2007). An Improved Holey Carbon Film for Cryo-Electron Microscopy. *Microscopy and Microanalysis*, *13*(5), 365–371. https://doi.org/10.1017/S1431927607070791

Rabe, M., Kerth, A., Blume, A., & Garidel, P. (2020). Albumin displacement at the air–water interface by Tween (Polysorbate) surfactants. *European Biophysics Journal*, *49*(7), 533–547. https://doi.org/10.1007/s00249-020-01459-4

Ramírez-Aportela, E., Carazo, J. M., Sorzano, C. O. S., & Sun, F. (2022). Higher resolution in cryo-EM by the combination of macromolecular prior knowledge and image-processing tools. *IUCrJ*, *9*(Pt 5), 632–638. https://doi.org/10.1107/S2052252522006959

Rosenthal, P. B., & Henderson, R. (2003). Optimal Determination of Particle Orientation, Absolute Hand, and Contrast Loss in Single-particle Electron Cryomicroscopy. *Journal of Molecular Biology*, *333*(4), 721–745. https://doi.org/https://doi.org/10.1016/j.jmb.2003.07.013

Russo, C. J., & Passmore, L. A. (2016). Ultrastable gold substrates: Properties of a support for high-resolution electron cryomicroscopy of biological specimens. *Journal of Structural Biology*, *193*(1), 33–44. https://doi.org/https://doi.org/10.1016/j.jsb.2015.11.006

Samanta, S., & Ghosh, P. (2011). Coalescence of bubbles and stability of foams in aqueous solutions of Tween surfactants. *Chemical Engineering Research and Design*, *89*(11), 2344–2355. https://doi.org/10.1016/j.cherd.2011.04.006

Sanchez-Garcia, R., Gomez-Blanco, J., Cuervo, A., Carazo, J. M., Sorzano, C. O. S., & Vargas, J. (2021). DeepEMhancer: a deep learning solution for cryo-EM volume post-processing. *Communications Biology*, *4*(1), 1–8. https://doi.org/10.1038/s42003-021-02399-1

Sathyanarayanan, N., Cannone, G., Gakhar, L., Katagihallimath, N., Sowdhamini, R., Ramaswamy, S., & Vinothkumar, K. R. (2019). Molecular basis for metabolite channeling in a ring opening enzyme of the phenylacetate degradation pathway. *Nature Communications*, *10*(1). https://doi.org/10.1038/s41467-019-11931-1

Schauperl, M., & Denny, R. A. (2022). AI-Based Protein Structure Prediction in Drug Discovery: Impacts and Challenges. *Journal of Chemical Information and Modeling*, *62*(13), 3142–3156. https://doi.org/10.1021/acs.jcim.2c00026

Scheres, S. H. W. (2012). RELION: Implementation of a Bayesian approach to cryo-EM structure determination. *Journal of Structural Biology*, *180*(3), 519–530. https://doi.org/10.1016/j.jsb.2012.09.006

Schrödinger, LLC. (2015). *The {AxPyMOL} Molecular Graphics Plugin for {Microsoft PowerPoint}, Version~1.8*.

Schürmann, U., Chluba, C., Wolff, N., Smazna, D., Lima de Miranda, R., Junker, P., Adelung, R., Quandt, E., & Kienle, L. (2017). Functional NiTi grids for in situ straining in the TEM. *Ultramicroscopy*, *182*, 10–16. https://doi.org/https://doi.org/10.1016/j.ultramic.2017.06.003

Sorzano, C. O. S., Jimenez-Moreno, A., Maluenda, D., Martinez, M., Ramirez-Aportela, E., Krieger, J., Melero, R., Cuervo, A., Conesa, J., Filipovic, J., Conesa, P., Del Cano, L., Fonseca, Y. C., Jiménez-De La Morena, J., Losana, P., Sanchez-Garcia, R., Strelak, D., Fernandez-Gimenez, E., De Isidro-Gómez, F. P., … Carazo, J. M. (2022). On bias, variance, overfitting, gold standard and consensus in single-particle analysis by cryo-electron microscopy. *Acta Crystallographica Section D: Structural Biology*, *78*, 410–423. https://doi.org/10.1107/S2059798322001978

Sorzano, C. O. S., Semchonok, D., Lin, S. C., Lo, Y. C., Vilas, J. L., Jiménez-Moreno, A., Gragera, M., Vacca, S., Maluenda, D., Martínez, M., Ramírez-Aportela, E., Melero, R., Cuervo, A., Conesa, J. J., Conesa, P., Losana, P., Caño, L. del, de la Morena, J. J., Fonseca, Y. C., … Carazo, J. M. (2021). Algorithmic robustness to preferred orientations in single particle analysis by CryoEM. *Journal of Structural Biology*, *213*(1). https://doi.org/10.1016/j.jsb.2020.107695

Szymczyk, K., Zdziennicka, A., & Jańczuk, B. (2018). Adsorption and Aggregation Properties of Some Polysorbates at Different Temperatures. *Journal of Solution Chemistry*, *47*(11), 1824–1840. https://doi.org/10.1007/s10953-018-0823-z

Tan, Y. Z., Baldwin, P. R., Davis, J. H., Williamson, J. R., Clinton, S., Carragher, B., Lyumkis, D., Microscopy, S. E., York, N., Biology, S., Biophysics, M., Biology, C., Jolla, L., Studies, B., & Jolla, L. (2018). *Cryo-EM through Tilting*. *14*(8), 793–796. https://doi.org/10.1038/nmeth.4347.Addressing

Tang, F., Ohto, T., Sun, S., Rouxel, J. R., Imoto, S., Backus, E. H. G., Mukamel, S., Bonn, M., & Nagata, Y. (2020). Molecular Structure and Modeling of Water-Air and Ice-Air Interfaces Monitored by Sum-Frequency Generation. *Chemical Reviews*, *120*(8), 3633–3667. https://doi.org/10.1021/acs.chemrev.9b00512

Tang, G., Peng, L., Baldwin, P. R., Mann, D. S., Jiang, W., Rees, I., & Ludtke, S. J. (2007). EMAN2: An extensible image processing suite for electron microscopy. *Journal of Structural Biology*, *157*(1), 38–46. https://doi.org/10.1016/j.jsb.2006.05.009

Taylor, K. A., & Glaeser, R. M. (1974). Electron Diffraction of Frozen, Hydrated Protein Crystals. *Science*, *186*(4168), 1036–1037. https://doi.org/10.1126/science.186.4168.1036

Tobiasson, V., Berzina, I., & Amunts, A. (2022). Structure of a mitochondrial ribosome with fragmented rRNA in complex with membrane-targeting elements. *Nature Communications*, *13*(1), 1–9. https://doi.org/10.1038/s41467-022-33582-5

Vinothkumar, K. R., & Henderson, R. (2016). Single particle electron cryomicroscopy: trends, issues and future perspective. *Quarterly Reviews of Biophysics*, *49*. https://doi.org/10.1017/s0033583516000068

Vogel, R. H., Provencher, S. W., Von Bonsdorff, C. H., Adrian, M., & Dubochet, J. (1986). Envelope structure of semliki forest virus reconstructed from cryo-electron micrographs. *Nature*, *320*(6062), 533–535. https://doi.org/10.1038/320533a0

Wei, H., Dandey, V. P., Zhang, Z., Raczkowski, A., Rice, W. J., Carragher, B., & Potter, C. S. (2018). Optimizing "self-wicking" nanowire grids. *Journal of Structural Biology*, *202*(2), 170–174. https://doi.org/10.1016/j.jsb.2018.01.001

Wrapp, D., Wang, N., Corbett, K. S., Goldsmith, J. A., Hsieh, C. L., Abiona, O., Graham, B. S., & McLellan, J. S. (2020). Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science*, *367*(6483), 1260–1263. https://doi.org/10.1126/science.aax0902

Wrobel, A. G. (2023). Mechanism and evolution of human ACE2 binding by SARS-CoV-2 spike. *Current Opinion in Structural Biology*, *81*, 102619. https://doi.org/https://doi.org/10.1016/j.sbi.2023.102619

Xu, J., Cui, X., Liu, N., Chen, Y., & Wang, H. (2021). Structural engineering of graphene for high-resolution cryo-electron microscopy. *SmartMat*, *2*(2), 202–212. https://doi.org/10.1002/smm2.1045

Yamashita, K., Palmer, C. M., Burnley, T., & Murshudov, G. N. (2021). Cryo-EM single-particle structure refinement and map calculation using Servalcat. *Acta Crystallographica Section D: Structural Biology*, *77*, 1282–1291. https://doi.org/10.1107/S2059798321009475

Zhang, K. (2016). Gctf: Real-time CTF determination and correction. *Journal of Structural Biology*, *193*(1), 1–12. https://doi.org/10.1016/j.jsb.2015.11.003

Zheng, L., Liu, N., Gao, X., Zhu, W., Liu, K., Wu, C., Yan, R., Zhang, J., Gao, X., Yao, Y., Deng, B., Xu, J., Lu, Y., Liu, Z., Li, M., Wei, X., Wang, H.-W., & Peng, H. (2023). Uniform thin ice on ultraflat graphene for high-resolution cryo-EM. *Nature Methods*, *20*(1), 123–130. https://doi.org/10.1038/s41592-022-01693-y

Zhong, J., Kumar, M., Anglada, J. M., Martins-Costa, M. T. C., Ruiz-Lopez, M. F., Zeng, X. C., & Francisco, J. S. (2019). Atmospheric Spectroscopy and Photochemistry at Environmental Water Interfaces. *Annual Review of Physical Chemistry*, *70*, 45–69. https://doi.org/10.1146/annurev-physchem-042018-052311

Zivanov, J., Nakane, T., Forsberg, B. O., Kimanius, D., Hagen, W. J. H., Lindahl, E., & Scheres, S. H. W. (2018). New tools for automated high-resolution cryo-EM structure determination in RELION-3. *ELife*, *7*, e42166. https://doi.org/10.7554/eLife.42166

Zivanov, J., Nakane, T., & Scheres, S. H. W. (2019). A Bayesian approach to beam-induced motion correction in cryo-EM single-particle analysis. *IUCrJ*, *6*(1), 5–17. https://doi.org/10.1107/S205225251801463X