1
2
3
4
5
6
7
8
9
10
11

Analysis of Local Variability and Allostery in Macromolecular Assemblies using Cryo-EM and Focused Classification

Cheng Zhang[1‡], William Cantara[2‡], Youngmin Jeon[1], Karin Musier-Forsyth[2], Nikolaus Grigorieff[3*], and Dmitry Lyumkis[1*]

1. Salk Institute for Biological Studies, 10010 N Torrey Pines Rd, La Jolla, CA 92037, USA.

2. Department of Chemistry and Biochemistry, Ohio State University, Columbus, OH 43210, USA.

3. Janelia Research Campus, Howard Hughes Medical Institute, 19700 Helix Drive, Ashburn, VA, USA.

‡ equal contribution

* Correspondence: niko@grigorieff.org or dlyumkis@salk.edu

Abstract

Single-particle electron cryo-microscopy and computational image classification can be used to analyze structural variability in macromolecules and their assemblies. In some cases, a particle may contain different regions that each display a range of distinct conformations. We have developed strategies, implemented within the Frealign and *cis*TEM image processing packages, to focus classify on specific regions of a particle and detect potential covariance. The strategies are based on masking the region of interest using either a 2-D mask applied to reference projections and particle images, or a 3-D mask applied to the 3-D volume. We show that focused classification approaches can be used to study structural allostery, a concept that is likely to gain more importance as datasets grow in size, allowing the distinction of more structural states and smaller differences between states. Finally, we apply the approaches to an experimental dataset containing the HIV-1 Transactivation Response (TAR) element RNA fused into the large bacterial ribosomal subunit, to deconvolve structural mobility within localized regions of interest.

Highlights

- Description of different image classification strategies in single-particle cryo-EM
- Quantitative evaluation of two classification methods using simulated data
- Application of the two classification methods to an experimental dataset

Keywords

Single-particle cryo-EM, cisTEM, classification, heterogeneity, ribosome

2

1. Introduction

Single-particle electron cryo-microscopy (cryo-EM) enables the visualization of macromolecules and their assemblies under near-native conditions [1]. In recent years, the technique has gained popularity, in part due to its ability to determine macromolecular structures at near-atomic resolution and without the need for crystallization [2]. While advances in resolution [3,4] have expanded the scope of the technique over the last five years, the ability to decipher structural heterogeneity is an ongoing area of development in the field [5,6]. Given that macromolecules, and especially their assemblies, are dynamic, image classification opens up the possibility to address novel types of questions pertaining to the molecular mechanisms underlying their function.

Structural heterogeneity can be either compositional or conformational in nature. Compositional heterogeneity means that the stoichiometry of subunits within an assembly varies within the dataset, such as particles containing or missing an additional, loosely associated protein factor. Conformational heterogeneity assumes that particles are uniform in composition, but the constituent components within each object can be flexible and can adopt one of several structurally different states. Conformational heterogeneity can be further subdivided into either discrete or continuous conformational heterogeneity. In the former case, the macromolecule would adopt one of several distinct structural states, each represented by a local minimum within the energy landscape describing all possible states. In the latter case, no distinct local energy minima exist, and the flexible regions can move in a mostly random manner. Finally, a fourth case can be defined as containing a combination of the above scenarios.

To understand structural heterogeneity within a single-particle experiment, the particle images are subject to a classification procedure, which assigns each particle to one of potentially many different classes. In the simplest scenario, a global classification strategy assigns each particle into a specific class on the basis of variability across the entire image. Different classification approaches have been developed, including supervised and unsupervised techniques, and numerous variations have been implemented to analyze structural heterogeneity [6-11]. Global 3-D classification does not require specific knowledge about the type and location of the heterogeneity, making it an integral part of today's processing workflow of virtually all single-particle software packages. Given that macromolecular assemblies can be highly dynamic, and because every subdivision leads to fewer particles within each class (and thus lower signal and loss of resolution), the fundamental disadvantage of a global classification strategy is the limited number of well-defined classes that can be recovered from a dataset of a given size. This is particularly true when one wants to resolve variability in small, heterogeneous regions that may easily be lost during a global

3

97   classification procedure. In contrast to a global classification strategy, "focused classification" zooms in
98   on a region or feature of interest, in order to understand structural heterogeneity in a localized manner [12-
99   15]. Focused classification can overcome the potential particle number limit associated with global
100   classification by reducing the number of classes needed to represent the local variability and (in principle)
101   excluding other regions of the particle from the analysis. This approach is particularly advantageous when
102   regions outside of the area of interest are themselves dominated by structural heterogeneity. For example,
103   minor domain movements within an otherwise dynamic macromolecular assembly might be difficult to
104   resolve using global classification techniques alone because the majority of the signal guiding the
105   classification procedure is dominated by regions outside of the area of interest. In another example, two
106   large regions can exhibit independent variability, and a global classification may not converge on a solution
107   that represents all possible states, or the number of states required leaves too few particles in the
108   corresponding reconstructions, limiting their resolution. In general, focused classification provides an
109   alternative means to deconstruct highly dynamic and/or heterogeneous datasets, reducing the analysis to a
110   more tractable problem. Numerous successful applications of focused classification have been used to
111   understand the independent movements of regions of large macromolecular complexes, such as the
112   spliceosome and the ribosome [16-19].

113

114   Focused classification requires selecting a region of interest within the particle and excluding the remaining
115   density. In the simplest implementation, a 3-D mask is applied to the reconstructed densities after each
116   iteration to select the area of interest, and standard global classification is then performed using the masked
117   reconstructions as references. A typical example of this is the classification of membrane proteins that
118   contain detergent micelles: the 3-D mask is used to exclude the heterogeneous micelle while focusing on
119   the protein [20]. The primary disadvantage of this "3-D masking" approach is that a projection of the
120   density, which *only* contains the masked region, is compared with the particle image, which contains the
121   masked region *in addition to* all other overlapping density, and this additional density can obscure the
122   features to be classified. To reduce the problem of density discrepancy, the density outside the mask could
123   be included in the reference after applying a low-pass filter [21,22]. The filter removes noise from the
124   disordered regions of the particle while maintaining valid low-resolution signal to minimize the mismatch
125   between reference and images. To further reduce density mismatch, another approach has been introduced,
126   whereby, in addition to masking the 3-D object, the density outside the mask is computationally subtracted
127   from the particle images [12,13,15]. This leaves a projection of the masked 3-D object and a density-
128   subtracted 2-D particle image, which contains comparable features that can be used for classification.
129   Another advantage of the "density subtraction" approach is that it can, in principle, be implemented in a
130   hierarchical fashion, in order to subtract increasingly finer features in a step-wise manner. The (non-

4

131   hierarchical) density subtraction approach has been used to improve heterogeneous regions of numerous
132   macromolecular complexes that could not be improved using a global classification approach alone
133   [12,13,15,19,23]. However, there are also disadvantages to this method. First, density subtraction requires
134   an accurate measure of the signal in each particle image to properly subtract the desired density. Especially
135   when looking at small regions and subtracting density corresponding to larger volumes, the subtraction may
136   leave residual signal in the raw images, a problem that is exacerbated if the complex exhibits greater
137   heterogeneity than is accounted for in the references used for density subtraction. The residual signal from
138   the incomplete density subtraction can interfere with subsequent classification and obscure the variability
139   in smaller regions (especially if applied in a hierarchical context). We and others have introduced another
140   approach, where focused classification is performed in 2-D, with masks imposed on both the projection
141   images and the experimental data [14,22]. In this alternative approach, a 3-D mask is defined for a region
142   of interest, projected along the view determined for each particle and applied as a 2-D mask to the particle
143   images and reference projections. Such an approach has been described in the context of bootstrap
144   resampling and using the cross-correlation function to find the optimal solution [14] and has now been
145   implemented within a likelihood-based framework in Frealign [8,22] and *cis*TEM [24]. The advantage of
146   the "2-D masking" approach with focused classification is that it does not require signal subtraction, while
147   constraining the classification to the area in the 2-D images that contain the region of interest and removing
148   noise outside this region.

149

150   A major advantage of any focused classification approach is its ability to selectively classify features of
151   interest within a distinct region of a cryo-EM map, which opens up numerous potential directions. First, it
152   enables classification of pseudo-symmetric features in a particle that are related by a symmetry operator
153   but not strictly symmetric due to independently dynamic mobility [15,25,26]. For example, surface-
154   exposed regions of macromolecules may not obey the strict symmetry that may apply to the particle core,
155   leading to loss of resolution in the surface regions of otherwise symmetric particles such as icosahedral
156   viruses (reviewed in [27]). To classify pseudo-symmetric regions of a particle, the images are first aligned
157   according to a common reference frame compatible with the pseudo-symmetry. The symmetry is then
158   dropped, and multiple alignments for each particle image are determined, corresponding to all possible
159   symmetry-related views, and an asymmetric reconstruction is calculated using each particle image multiple
160   times to include all symmetry-related alignments. This effectively multiplies the number of particles in a
161   dataset by the number of different possible symmetry operations and enables classification of different
162   views into different classes, thereby resolving the heterogeneity in the pseudo-symmetric regions. This
163   approach can, therefore, improve the resolution of density that would otherwise be an average of multiple
164   structural states due to symmetrization. The approach has been applied, for example, to resolve density

5

165 detail that was not visible after global classification alone [26], and to reveal genome structures within viral

166 particles [28] (for other examples, see [27]). Second, selectively focusing on discrete asymmetric units can

167 reveal covariant heterogeneity within the data. For example, two different regions located on opposite sides

168 of a particle might be structurally coupled with each other. If the variability of two regions is random, there

169 should be no correlation in the assignment of these regions to different classes during pseudo-symmetric

170 classification. However, if correlation is present, this indicates covariance in the two regions. In the

171 simplest case, counting of the number of matching asymmetric units within the same class, and comparison

172 with a random distribution, would provide evidence for structural allostery. This phenomenon represents

173 an area of development that may facilitate understanding global structural landscapes of dynamic

174 macromolecular machines.

175

176 In this manuscript, we explore several different focused classification strategies with both synthetic and

177 experimental data. We show the advantages and disadvantages of the "2-D masking" and "3-D masking"

178 approaches, and additionally explore their ability to discover density covariances within otherwise distinct

179 regions of a reconstruction. Finally, we show how focused classification can be applicable to heterogeneous

180 experimental datasets, highlighting a particular test case that is relevant to visualizing mounted targets on

181 scaffolds using single-particle cryo-EM.

182

183 **2. Materials and methods**

184

185 *2.1 Generation of synthetic humanoid datasets.* Synthetic datasets were generated as previously described

186 [8]. Briefly, we randomly shifted and rotated projection images of humanoid structures, added noise, a CTF

187 (to have CTF-modulated noise components), envelope function, and a final layer of noise. To reduce

188 spurious correlations associated with the CTF for covariance analysis, we used a 640-pixel box size for

189 projecting the data, and prior to the addition of noise and the CTF. 28 distinct datasets were made,

190 corresponding to the different structural combinations of arms, hands, and feet (Figure 1). Combined

191 datasets corresponding to the three distinct scenarios were then generated from the individual 28 datasets.

192 Each combined dataset contained 10,000 particles (pixel size 5.24, box size 80 after Fourier resampling)

193 with each of the 28 sub-datasets selected randomly.

194

195 *2.2 Particle assignment during focused classification*. To facilitate quantitative assessment, we made the

196 assumption that each classified particle belongs to the class with the highest probability (occupancy in

197 Frealign/*cis*TEM). At higher SNRs, this was an insignificant assumption, as most occupancies were close

198 to 1; however, at lower SNRs, particles are represented by lower occupancies in multiple classes with slight

6

199 differences between them. By assuming that each asymmetric unit corresponds to the class with the highest

200 occupancy, we could simplify the calculation of κ coefficients and other analyses.

201

202 *2.3 Measures for evaluating the accuracy of classification*. To evaluate the accuracy of each classification

203 trajectory, we define the following measures. For each asymmetric unit in each class:

204 - TP (true positive) — starting occupancy 100, ending marginal occupancy greater than all other

205 classes.

206 - FP (false positive) — starting occupancy 0, ending marginal occupancy greater than all other

207 classes.

208 - TN (true negative) — starting occupancy 0, ending occupancy less than the class with greatest

209 marginal occupancy

210 - FN (false negative) — starting occupancy 100, ending occupancy less than the class with greatest

211 marginal occupancy

212 - N: number of observations — TP+FP+TN+FN

213

214 Using the definitions above, the following metrics are defined:

215 Accuracy (the relative observed agreement among raters, or *Po*) = (TP + TN) / N

216 Sensitivity = True Positive Rate (TPR) = TP / (TP + FN)

217 Specificity = True Negative Rate (TNR) = TN / (TN + FP)

218 Kappa:

219 $$\frac{Po - Pe}{1 - Pe} = 1 - \frac{1 - Po}{1 - Pe}$$

220

221 where *Po* is the accuracy, above, and *Pe* is the probability of chance agreement.

222 Youden's Index (J Statistic) = TPR + TNR − 1.

223

224 *2.4 Merging cryo-EM difference maps.* Merging of the difference maps in Figure 4 was performed

225 according to the following procedure. A merge volume was generated with 0s for the pixel values.

226 Subsequently, for each pairwise difference map, and for each voxel, if the value of the voxel is greater than

227 the value of this voxel in the merged map, set this as the value in the verged map.

228

229 *2.5 Covariance analysis of separate regions of cryo-EM density maps.* To determine whether different

230 regions correlate with one another, normalized covariances were computed comparing fractional density

231 occupancies of distinct components. An identical procedure was used for both scenarios 2 and 3. First, we

7

232 performed 3-D focused classification, with the requested number of classes, k, identical to the expected

233 number of non-degenerate asymmetric units. Binary masks were created for each region of interest (ROI),

234 namely the hand in each of two positions, the near foot, and the far foot. The masks encompassed the ROI,

235 with minimal incursion into neighboring density. A soft edge was not employed, because the mask was

236 solely used for the purpose of computing fractional density occupancy values. For each of the k resulting

237 maps, and for each ROI, the mask was used to extract the resulting density. Subsequently, the approximate

238 mass in the ROI was calculated using the "volume" command implemented within the EMAN1 processing

239 suite [29]. The resulting mass was optionally normalized to the true mass arising from a perfect

240 classification to judge the quality of the classification, although this step is not strictly necessary for

241 normalized covariance analysis. Finally, the normalized covariance matrix R_{ij} was computed as:

242
$$R_{ij} = \frac{C_{ij}}{\sqrt{C_{ii} * C_{jj}}}$$

243 where C_{ij} refers to the covariance between two components i and j. To make sure that there was adequate

244 sampling, the resulting volumes represent an average of 3 independent runs, using random starting class

245 occupancy values for initiating each classification.

246

247 **2.6 Ribosome preparation.** The 57-nt HIV-1 TAR element was appended inserted into twelve different

248 helices (H9, H12, H19, H24, H25, H31, H45, H46, H59, H63, H68 and H98) by replacement of the loop

249 residues to screen for optimal attachment sites. These twelve were chosen based on their location on the

250 periphery of the ribosome and lack of tertiary contacts. All insertions resulted in viable bacterial growth

251 (albeit much slower in some cases). H45 qualitatively yielded the most complete density with the least

252 apparent mobility of the attached RNA (data not shown). Uniformly labeled ribosomes were prepared in

253 the same way for all insertions. To ensure that all ribosomes contain the appended construct, a well-

254 established protocol for introducing and characterizing site-specific mutations into *Escherichia coli*

255 ribosomes was used [30,31]. Briefly, a *Δ7 prrn E. coli* strain SQZ10 [32], which has a genomic deletion of

256 all rRNA genes, was used. The rRNA genes are supplied by a plasmid that also contains the *levansucrase*

257 gene and confers kanamycin resistance (Plasmid 1, pHK-rrnC-sacB). Levansucrase expression is lethal to

258 *E. coli* when grown on sucrose-containing media [33]. An additional ampicillin-resistant plasmid

259 containing the rRNA genes with the RNA construct of interest inserted (Plasmid 2, p278) was then

260 transformed and grown in liquid culture. Cells were plated on media containing ampicillin and 5% sucrose

261 to select for those that had lost Plasmid 1 but retain Plasmid 2. To confirm the selection, colonies were

262 plated on Kan media to ensure that they cannot grow.

263

264 Insertion of TAR into helix 45 of p278 was carried out using site-directed ligase-independent mutagenesis

265 [34]. Mutant plasmids were then transformed into SQZ10 cells and selected using the strategy described

266 above. Mutant ribosomes were purified by first growing to mid-logarithmic phase (OD_{550} = 0.3-0.5) in 500

267 mL Luria Broth while shaking at 37 °C then chilled on ice for 30 minutes and pelleted by centrifugation.

268 The cell pellet was then resuspended in 20 mL Resuspension Buffer (20 mM Tris-HCl, pH 7.5, 10 mM

269 $MgCl_2$, 100 mM NH_4Cl, 0.5 mM EDTA, 2 mM $CaCl_2$, 6 mM β-mercaptoethanol). The resulting

270 resuspension was lysed through a French Press three times, filtered through a 0.45 µm syringe filter and

271 clarified by centrifugation at 18,000g for 30 minutes twice. The supernatant was concentrated to ~500 uL

272 using a 50K MWCO filter (Amicon) and layered onto 36 mL 10-40% sucrose gradient in Gradient Buffer

273 (50 mM Tris-HCl, pH 7.5, 10 mM $MgCl_2$, 100 mM NH_4Cl, 6 mM β-mercaptoethanol) and ultracentrifuged

274 in SW-32Ti rotor at 16,700g for 18.5 hours at 4 °C. 70S ribosomes fractions were collected, buffer

275 exchanged into Storage Buffer (20 mM Tris-HCl, pH 7.5, 10 mM $MgCl_2$, 100 mM NH_4Cl, 6 mM β-

276 mercaptoethanol), aliquoted and stored at 4 °C until ready for grids.

277

278 *2.7 Cryo-EM grid preparation and data acquisition.* 2.5 µl of purified ribosomes after sucrose

279 fractionation were diluted to a concentration of 4 mg/ml with Storage Buffer and placed on UltrAuFoil

280 R1.2/1.3 300-mesh grids (Quantifoil) that were plasma-cleaned (75% argon/25% oxygen atmosphere, 15

281 W for 7 s using a Gatan Solarus). After 1 min incubation under >80% humidity at 4 °C, grids were blotted

282 manually with a filter paper (Whatman No. 1) before being plunged into liquid ethane cooled by liquid

283 nitrogen using a manual plunger. Leginon was used for automated EM image acquisition [35]. Grids were

284 imaged on a Titan Krios microscope (FEI) operating at 300kV and equipped with a K2 Summit direct

285 electron detector (Gatan). A nominal magnification of 22,500x was used for data collection, giving a pixel

286 size of 1.31 Å at the specimen level, with the defocus range of -0.5 μm to -2.5 μm. Movies were recorded

287 in counting mode with an accumulated total dose ~50 electrons/$Å^2$ fractionated into 60 frames with an

288 exposure rate of ~7 electrons/pixel/s.

289

290 *2.8 Image processing and model generation.* All pre-processing was performed within the Appion suite

291 [36]. Motion correction was carried out by using the program MotionCor2 [37] and exposure-filtered in

292 accordance with the relevant radiation damage curves [38]. The CTF for each micrograph was estimated

293 using CTFFind4 [39] during data collection. 70S ribosomes served as a template for automatic particle

294 picking using FindEM [40]. 346K particles were selected and subjected to per-particle CTF estimation

295 using the program GCTF [41]. After 2D and 3D classification in GPU-enabled Relion [42,43], selected

296 classes containing 232K particles were combined to a single stack and imported to Frealign for global

297 refinement with 8 classes. Every ten cycles of refinement/classification, the reconstructed maps of all 8

298 classes were aligned to a common 50S scaffold using custom scripts implemented for performing a 3-D

9

299   alignment within the Chimera package [44] while running Frealign/*cis*TEM, in order to maintain a common

300   reference-frame for subsequent focused classification. A total of 50 cycles of global

301   refinement/classification were performed. Subsequently, the best orientations were combined into a single

302   parameter file for focused classification. Focused classification was performed for 500 cycles, and without

303   further alterations to the orientations, by defining a spherical mask of 30 Å, centered on the expected region

304   of TAR. Global resolution for the final map was estimated using the Fourier shell correlation (FSC [45]) at

305   0.143 and directional resolution anisotropy was evaluated by the 3D FSC server [46]. Local resolution

306   estimation was performed using sxlocres.py implemented within Sparx [47].

307

308   The model of TAR attached to H45 of the 23S ribosome was prepared by first removing the loop residues

309   of H45 from a recent 2.9 A structure, PDB ID 5AFI [48], and removing the polyA nucleotides from a model

310   of TAR based on small-angle X-ray scattering data. The terminal backbone atoms were docked and aligned

311   in UCSF Chimera [44]. The TAR region was then rigid-body refined into the cryo-EM density in Coot

312   [49].

313

314   **3. Results**

315

316   *3.1 Quantitative characterization of focused classification with 2-D and 3-D masking*

317

318   3-D classification with different masking options, including the 3-D masking and 2-D masking, have been

319   described and implemented within Frealign [8,22] and *cis*TEM [24]. In the present study, we quantitatively

320   characterize the performance of these different options using simulated data, highlighting strengths and

321   weaknesses of each approach. We generated multiple synthetic datasets that are characterized by various

322   degrees of heterogeneity. Figure 1 shows the distinct components of a "humanoid" reconstruction, with

323   the legs, body, neck, and head positioned identically, and representing the constant, homogeneous regions

324   of a particle, characterized by twofold rotational symmetry. In contrast, the arms can belong to one of two

325   conformations, and are therefore characterized by pseudo-symmetry. Lastly, the hands and feet, which

326   represent small features of a map that might be lost during global classification, can be either present or

327   absent. We generated maps representing all possible combinations of these features and created multiple

328   synthetic datasets containing random translations and rotations, a contrast transfer function (CTF), an

329   envelope function, and multiple levels of noise, bringing the final CTF-modulated SNR down to 0.100,

330   0,050, 0.025, 0.013, or 0.006, as previously described (Supplementary Figure 1 and [8,50]). Below, we

331   describe three scenarios, which serve to demonstrate different aspects of focused classification.

332   Importantly, in all described cases, focused classification is performed on an asymmetric subunit basis,

333   which allows one to break down and constrain the heterogeneity problem [27] and reveal discrete

334   movements within a more complex landscape of heterogeneity.

335

336   *First scenario – the base, pseudo-symmetric case*: In the base scenario, only the arms/hands are mobile and

337   can adopt one of two distinct positions within an asymmetric unit, and the hand always remains co-occupied

338   with an arm (Figure 1A). This case represents a common problem with pseudo-symmetric experimental

339   datasets, whereby most of the molecule is homogeneous and characterized by symmetry (here, twofold),

340   but one feature does not obey symmetry constraints (here, the arms/hands). There are four combinatorial

341   possibilities, three of which would be expected to be recovered using a global classification strategy

342   (structures A2 and A3 are degenerate and are related by 180° rotation). However, in an asymmetric focused

343   classification centered on one side of the humanoid, one would expect to find only two non-degenerate

344   possibilities, because the arm/hand can reside in only one of two structural states.

345

346   *Second scenario – identifying small densities*: In the second scenario, we use focused classification to

347   recover finer features within a more complex structural landscape. In addition to the arms occupying one

348   of two distinct positions, the hands can be either present or absent, and their occupancy is completely

349   randomized (Figure 1B). Thus, for each of the four structural states described in the base scenario, one

350   would see four additional structural states represented by the presence or absence of each of two hands. In

351   sum, there are 16 different combinatorial possibilities, global classification would be expected to uncover

352   10 non-degenerate classes, but only four classes should be resolved using asymmetric classification.

353

354   *Third scenario – identifying small densities and covariances*: The third scenario is identical to the second

355   scenario, except that a hand on each asymmetric unit is always co-associated with its corresponding foot

356   (Figure 1C). For example, if the left hand is present, so is the left foot, and if it is absent, the foot too is

357   absent; the same applies to the opposite asymmetric unit. One can then classify on the hand only, but look

358   at both the hand and foot areas in the resulting maps and count the number of times that density for the hand

359   co-occurs with density for the foot. In doing so, one can begin to decipher patterns and relationships within

360   distinct components.

361

362   *3.2 Focused classification on an asymmetric subunit of a synthetic humanoid*

363

364   For each of the three cases described above, and for all five levels of noise, we performed focused

365   classifications on a single asymmetric unit, with a mask around the region encompassing an arm and hand

366   (Figure 2A). For these experiments, the particle alignment parameters were set to the correct parameters

11

367   used to generate the data and were kept fixed during classification. To quantitatively evaluate the accuracy

368   of classification, we used the κ coefficient as a statistical measure, which captures the performance of a

369   diagnostic test, while taking into account the possibility of occurrence by chance [51]. We also used the

370   Youden's J statistic (informedness, [52]), but found that the results largely paralleled those of κ (data not

371   shown). The κ coefficient evaluates the agreement of raters for classifying N items into mutually exclusive

372   classes and relies on the precise knowledge of the number of false positives (FP), false negatives (FN), true

373   positive (TP), and true negatives (TN), which we can obtain from the data (see Methods). Importantly, κ

374   estimates the probability of an "informed" decision by taking into account random chance and returns 0

375   when classification is random (chance) and 1 when perfect classification is achieved. Qualitatively, it is

376   simple to visually assess how "clean" the classification is, and whether or not the particles were correctly

377   partitioned, by looking at the separation of the arms in our data. Supplementary Figure 2 shows how the

378   results look when classification is nearly perfect (Supplementary Figure 2A), when classification is

379   completely random (Supplementary Figure 2D), and two intermediate cases (Supplementary Figure 2B-C).

380   A correct classification partitions the arms within a single asymmetric unit (and not its counterpart) into

381   two distinct classes, with no signs of contaminating density (κ close to 1); as more errors are introduced,

382   the two classes become progressively more mixed, up to a point where one cannot distinguish between the

383   two volumes within or outside the asymmetric unit (κ close to 0, Supplementary Figure 2). In this manner,

384   we could also determine which parameters provide optimal classification results (e.g. mask size, soft edge

385   drop-off, etc., as demonstrated in Supplementary Figure 3), which we determined prior to evaluating the

386   test cases.

387

388   Table 1 shows the result of focused classification for all three scenarios, using both a 2-D masking approach

389   and a 3-D masking approach, as implemented in Frealign and evaluated using the κ coefficient. The

390   resulting numbers indicate the following general trends. First, for all three cases and for virtually all SNRs,

391   the 2-D masking approach was superior to the 3-D masking approach. Such a result is not surprising

392   because, as indicated in the introduction, the disadvantage of the 3-D masking approach, in the absence of

393   density subtraction, is that the experimental projection images contain overlapping density along the path

394   of the projection, as compared to a projection of the masked region from the reference map. The second

395   general trend is that, with more mobile components within a dataset, and the smaller the desired features

396   for detection, the lower the κ value and the more challenging it is to correctly classify the data. We observe

397   major differences in accuracy between case 1 and either 2 or 3, because the latter contain more moving

398   parts. However, the accuracies between cases 2 and 3 are roughly similar, likely because only small

399   structural differences characterize the two datasets. Third, a lower SNR makes it more challenging to

400   correctly classify the data, which is not surprising. However, it was surprising that, for the base scenario,

401   even at the lowest SNRs and given how small of a feature we were trying to detect, we could still recover

402   meaningful information and reasonably clean classes using the 2-D masking approach in particular, and to

403   a lesser extent using the 3-D masking approach. In scenarios 2-3, higher SNRs were required to recover

404   the correct classes (0.025 compared to 0.006, or ~4 times as high).

405

406   Our experiments reveal that the 2-D masking approach, in its implementation within the likelihood-based

407   framework of Frealign/*cis*TEM, does not completely isolate the area of interest from its surrounding

408   density. While the 2-D masking approach produces more accurate results in the cases analyzed, its primary

409   disadvantage is that projection images can contain additional density along the direction of the projection;

410   if this density is homogeneous, it should be neutral in terms of classification, but if it is itself heterogeneous,

411   it can bias the classification results. To account for this and to quantify the bias, we went back to the base

412   scenario, where only the arm/hand combinations can move, but applied the mask onto an area of a leg and

413   classified in that region (Figure 2B). We thus asked whether we can recover density for the arms, despite

414   the mask being situated in a different location. As before, the number of correctly assigned particles was

415   judged based on the arm/hand classes. If the arms completely determine the classification results, we would

416   expect to see a κ coefficient of 1, whereas in the absence of crosstalk between arms and legs, the arms/hands

417   would be randomly assigned and the κ coefficient would be 0. Table 2 shows that only at the highest SNRs

418   does the heterogeneity outside of the area of interest influence the classification, and with a maximum κ

419   coefficient of 0.23, the bias is not very severe. For SNR values of 0.025 and below, the results are

420   effectively random. For the same dataset, a κ coefficient of 0.87 is obtained for an SNR of 0.025 when the

421   mask is in its correct position around an arm. In contrast to the 2-D mask, when a 3-D mask is applied to

422   the same location, the results are completely random at all SNRs. This is exactly what we would expect,

423   because density outside this mask should not be introduced into a projection image after application of a 3-

424   D mask. The above results indicate that bias generated by heterogeneity outside the area of interest is

425   present but minor when using the 2-D masking approach, and absent in the 3-D masking approach.

426

427   *3.3 Focused classification can identify covariant components in distinct regions of a map*

428

429   Each individual object within a heterogeneous single-particle cryo-EM experiment can contain a unique

430   combination of dynamic elements residing in distinct structural states. When multiple components are

431   dynamic, and/or if they bind (or dissociate) in different regions, the conformational/compositional states of

432   the components can be linked. Using focused classification, one can treat two distinct regions separately,

433   and then ask whether there is any inter-dependence by calculating covariances within masked regions.

434

435   To evaluate covariance between distinct regions of a map, we used the datasets prepared for scenarios 2-3.

436   In scenario 2, the presence of either hand, or either foot, are random and are not related to one another. In

437   contrast, in scenario 3, the presence of a hand on one side of the humanoid is always correlated to the

438   presence of a foot on that same side, whereas the opposite foot is randomly occupied and is not correlated

439   to anything. Thus, one can apply a mask around the hands (encompassing both conformations), focus-

440   classify the data, and then look for the presence or absence of a foot, which has not been subjected to

441   focused classification. Quantitatively, once the dataset is classified and subdivided into groups, one would

442   simply calculate the fractional density occupied by each component within the class (e.g. hand in position

443   1, hand in position 2, near foot, and far foot) normalized to its expected value, and compute a normalized

444   covariance matrix (also known as a correlation coefficient matrix, see Methods) between the components.

445   Since the presence of a foot is always correlated with the hand on the same side of the humanoid,

446   irrespective of the conformation of the arm/hand, we further simplify the analysis by grouping both

447   mutually exclusive hand positions into, more generally, a "near hand". Thus, there are three regions for

448   which fractional occupancies are computed – a "near hand" (blue in Figure 3), where the mask is applied

449   for classification, a "near foot" (purple in Figure 3) on the same side of the humanoid, and a "far foot" (pink

450   in Figure 3) on the opposite side of the humanoid. Given the nature of the mask, everything except for the

451   hands is excluded from the classification. Since the mask is applied on an asymmetric-unit basis, the region

452   that would otherwise constitute the "far hands" is not separated, and both mixed conformations are

453   observed.

454

455   For scenario 2, whereby no covariance is expected, the volumes captured through focused classification on

456   an asymmetric-unit basis, and representing the four non-degenerate classes, are displayed in Figure 3A. As

457   expected, they differ in the presence, absence, and overall conformation of the hands. For example, classes

458   1,2 or classes 3,4 differ by the presence or absence of a single hand; classes 1,3 or classes 2,4 either do or

459   don't have hands, respectively, but differ in the conformation of the arms; finally, classes 1,4 or classes 2,3

460   differ in both hand occupancy and arm conformation. Other than the hand/arm differences, no other regions

461   of the maps have any apparent variability. Quantitatively, this is summarized by a normalized covariance

462   matrix that describes the relative interdependence between the different components (Figure 3B). A value

463   of 1 means that the pairwise occupancies of any two components are perfectly correlated, whereas a value

464   of 0 means that they are completely random (a value of -1 means that they are anti-correlated). Identical

465   components, related by the diagonal, are perfectly correlated, by definition. Otherwise, it is apparent that

466   no two regions of the map are correlated to one another. This situation is different for scenario 3, however,

467   which was designed to have the nearby hand and foot co-vary. The volumes captured through focused

14

468 classification again represent the expected non-degenerate classes, and the hands/arms are related to one

469 another in an identical manner as before. However, this time, it is clear that classes 2 and 4 are missing the

470 nearby foot, whereas classes 1 and 3 maintain full occupancy. The normalized covariance matrix now

471 shows that the hand is always co-associated with the nearby foot. The occupancy of the far foot, on the

472 other hand, remains random, and is accordingly associated with a low normalized covariance value. The

473 same experiment can be performed for more complicated combinations of hands and feet, but the principle

474 is the same – that assessing the inter-dependence of density occupancies within distinct regions of a

475 macromolecular complex can provide insight into hidden allostery within the data.

476

477 *3.4 Focused classification facilitates deconvolving heterogeneous regions within an experimental*

478 *dataset*

479

480 The techniques described here have been used to decipher both conformational and compositional

481 heterogeneity within biological samples (for example, [16,26,53]). In addition to the published results, one

482 area where they will be particularly useful is to deconvolve conformational heterogeneity when using

483 scaffolds for the purpose of structure determination. Several groups have shown that larger protein and/or

484 nucleic-acid scaffolds can be used to aid in the determination of smaller structures, which by themselves

485 would be too challenging to analyze [54,55]. However, the problem with all current approaches is that the

486 particles of interest are not necessarily rigidly bound. Thus, the regions closer to the site of attachment will

487 be characterized by less heterogeneity (and a lower B-factor), whereas the regions further from the site of

488 attachment will exhibit more heterogeneity (and a higher B-factor). To demonstrate this, we used a bacterial

489 70S ribosome as a scaffold, and engineered in a fusion RNA representing the HIV-1 Transactivation

490 Response (TAR) element. Subsequently, we performed either global classifications on the entire dataset

491 or focused classifications on the region around TAR.

492

493 The HIV-1 TAR element was uniformly inserted into Helix 45 of the E. coli large 23S ribosomal RNA.

494 Ribosomes containing the TAR knock-in were selectively purified (see Methods) and subjected to single-

495 particle cryo-EM analysis. We collected 929 micrographs, providing 346,851 particles in the dataset

496 (Supplementary Figure 4A). A single-model refinement, in the absence of any classification, showed high-

497 resolution in the ribosome core, and lower resolution in the regions characterized by structural heterogeneity

498 (Supplementary Figure 4B-C). Due to a large amount of mobility, the site of TAR fusion was only partially

499 visible at the normal thresholds used for displaying the coulombic potential map. We then performed a

500 global classification of the data, using a soft-edge spherical mask. This procedure resulted in distinct

501 classes, separated according to the expected heterogeneity associated with purified bacterial ribosomes [56]

15

502   (Supplementary Figure 4D). The combined differences are summarized with a merged map, demonstrating

503   the full extent of heterogeneity for the global classification case (Figure 4A); notably, the resolved

504   heterogeneity did not improve the density at the site of fusion. Subsequently, we performed a focused

505   classification of the data using 2-D masks, applying the mask to the area where TAR has been inserted. As

506   expected, the resulting maps were able to clearly separate out some of the different conformations of TAR

507   (Supplementary Figure 4E). However, the majority of the normal ribosomal heterogeneity was largely

508   ignored, as summarized by the merged difference maps (Figure 4B) and an overlay of the reconstructed

509   classes (Figure 4C). In terms of characterizing classification performance, this result is important for

510   several reasons. First, even though the area of interest is small, the focused classification approach using

511   2-D masks can partially deconvolve the density. Second, despite the extensive "normal" structural

512   heterogeneity present on bacterial ribosomes (e.g. Figure 4A), which may confound the 2-D focused

513   classification approach (e.g. Figure 2 and Table 2), we do not observe this in our results. We also performed

514   focused classifications using 3-D masks, but the quality of the reconstructed TAR region was noticeably

515   poorer (data not shown), consistent with the poorer performance of the 3-D masking approach using

516   synthetic data (e.g. Table 1). These experimental results further demonstrate the ability of the 2-D masking

517   approach to separate out local structural variabilities in the context of otherwise extensive global structural

518   differences.

519

520   The best reconstruction of HIV-1 TAR showed a clearly defined RNA helix, a marked improvement over

521   a global classification strategy alone (Figure 4D). The density was characterized by progressively poorer

522   resolution, as a function of distance from the site of attachment. For a largely A-form HIV-1 TAR RNA

523   helix, the behavior of the fusion can be thought of as a lever pivoting around a fulcrum; the further out from

524   the point of attachment, the more inherent mobility, and thus the lower the resolution. A similar behavior

525   has been observed with other scaffolding strategies, whereby the peripheral regions are characterized by

526   lower resolution [54,55]. In addition to providing novel biological insight, focused classifications can

527   broadly facilitate scaffolding approaches for solving structures of small proteins and RNAs.

528

529   **4. Discussion**

530

531   Using a synthetic dataset, we describe a quantitative assessment for several focused classification

532   implementations within the Frealign/*cis*TEM processing packages. The algorithms have been used to

533   classify features in several experimental studies [16,26,53], and we further demonstrate the applicability of

534   the approaches for deconvolving heterogeneous regions within small scaffolded RNAs to facilitate the

535   development of substrate supports for cryo-EM [54,55].

536

537 The present study will help users decide which strategy to use in a particular case. Focused classification

538 using 2-D masks can be applied to individual asymmetric features (also known as symmetry expansion

539 [27]), and, as implemented within Frealign/*cis*TEM, have generally been found to perform better than 3-D

540 masking approaches, due to density mismatch between particles images and reference projections after 3-

541 D masking. A possible disadvantage of the 2-D masking approach arises from the projection nature of the

542 data. Any area within a 2-D projection image will not only contain density relevant to the region of interest,

543 but also residual density along the projection path. If the residual density is itself heterogeneous, it can

544 potentially confuse or bias the classification procedure (especially if the variability within the region of

545 interest is significantly smaller compared to variability elsewhere). In Table 2, we demonstrate that this

546 effect is real, at least with high SNR data. However, in practice this problem appears to be small, based on

547 the results obtained with the synthetic data (compare Tables 1 and 2), and in an experimental setting in the

548 context of large-scale global heterogeneity in the current work (Figure 4A-B), and in previous biological

549 studies [16,18]. Conflating heterogeneity along the projection path would be treated as noise, in a manner

550 that is perhaps analogous to incomplete density subtraction.

551

552 Our tests with the synthetic dataset demonstrate that additional questions, such as those pertaining to

553 structural allostery, can be addressed in single-particle experiments. We showed how classifying variability

554 in a region of a density map can reveal covariance with a secondary region, in this case between a hand and

555 a foot. With synthetic data, such analyses are predicated upon having knowledge of the real density; in an

556 experimental setting, an analogous approach would mask out regions corresponding to, for example, known

557 components prior to analyzing the resulting normalized covariance matrices, as has been previously shown

558 in one simplified example with ribosome-associated factors [57]. In general, the ability to classify

559 independently on separate regions of a map provides opportunities to inter-relate distinct regions of an

560 object beyond simply recovering densities, a form of computational identification of allostery within a

561 system. Some cautions should be taken in the analyses of covariance. First, to avoid under-sampling, it is

562 advisable to compute an equal or greater number of classes than expected. Second, and related to the

563 previous point, classifications should be run multiple times, starting from different random particle seeds.

564 Both of these precautions will ensure that sufficient pairwise occupancies have been calculated to reach

565 statistical significance and avoid spurious correlations. Third, some caution should be taken in the

566 interpretations of results using 2-D masks (due to the possibility of "leaky" biases during classification),

567 although our experimental observations suggest that the biases should be minimal (Figure 4B). Finally,

568 global classifications can also be used for the purpose of covariance analysis, and they can have specific

569 advantages, as they would recover non-degenerate differences that are lost during classification on an

17

570   individual asymmetric unit (which is easily seen with the experimental setup of the humanoid, as the

571   number of non-degenerate structures (globally) far outnumbers the number of distinct asymmetric units).

572   Whereas focused classifications help constrain the number of different classes and can simplify the analysis,

573   the results should ideally relate to the global context of heterogeneity. In the future, more elaborate methods

574   could be devised for broader applicability beyond pairwise covariances.

575

576   Our results using HIV-1 TAR fused to bacterial large ribosomal subunits show how focused classifications

577   can help computationally deconvolve highly mobile features within experimental cryo-EM datasets. These

578   data are particularly applicable for the development of structural scaffolds for the analysis of small proteins

579   and RNAs [54,55]. The TAR fusions are universally mobile about a central fulcrum point, which

580   corresponds approximately to the site of attachment, and the density is lost in the absence of proper

581   classification. However, careful application of masks during focused classification enables partial recovery

582   of some of the structural elements within the TAR fusion, visualizing most of the A-form RNA helix.

583   Scaffolding approaches are gaining popularity in single-particle analysis, because small proteins may not

584   have sufficient signal for accurate assignment of Eulerian orientations. Focused classification can help

585   ameliorate problems associated with structural mobility and bring out the most of the structure of interest.

586

18

Acknowledgements

We thank Dr. Kurt Fredrick for assistance and helpful discussions regarding the ribosome purifications, Bill Anderson for help with cryo-EM data collections. We acknowledge the support of NIH grants RO1 GM065056 (to KMF) and P50 GM103368 (HIVE Center, to KMF and DL). DL also acknowledges the support of DP5 OD021396. WAC was supported by a Pelotonia Postdoctoral Fellowship from the Ohio State University. NG is an Investigator of the Howard Hughes Medical Institute.

Conflict of Interest

The authors declare no competing financial interests.

Author Contributions

WAC, KMF, NG and DL designed the study. NG is the primary developer of Frealign and *cis*TEM. DL prepared and performed calculations with the simulated data. CZ, WAC and YJ performed the calculations with experimental data. CZ, WAC, NG and DL analyzed the data. DL wrote the paper, with help from all authors.

Code availability

Frealign and *cis*TEM are open-source and distributed under the Janelia Research Campus Software License (http://license.janelia.org/license/janelia_license_1_2.html). All scripts and datasets for these studies will be made available from the Lyumkis laboratory upon request.

Figure and table legends

609
610
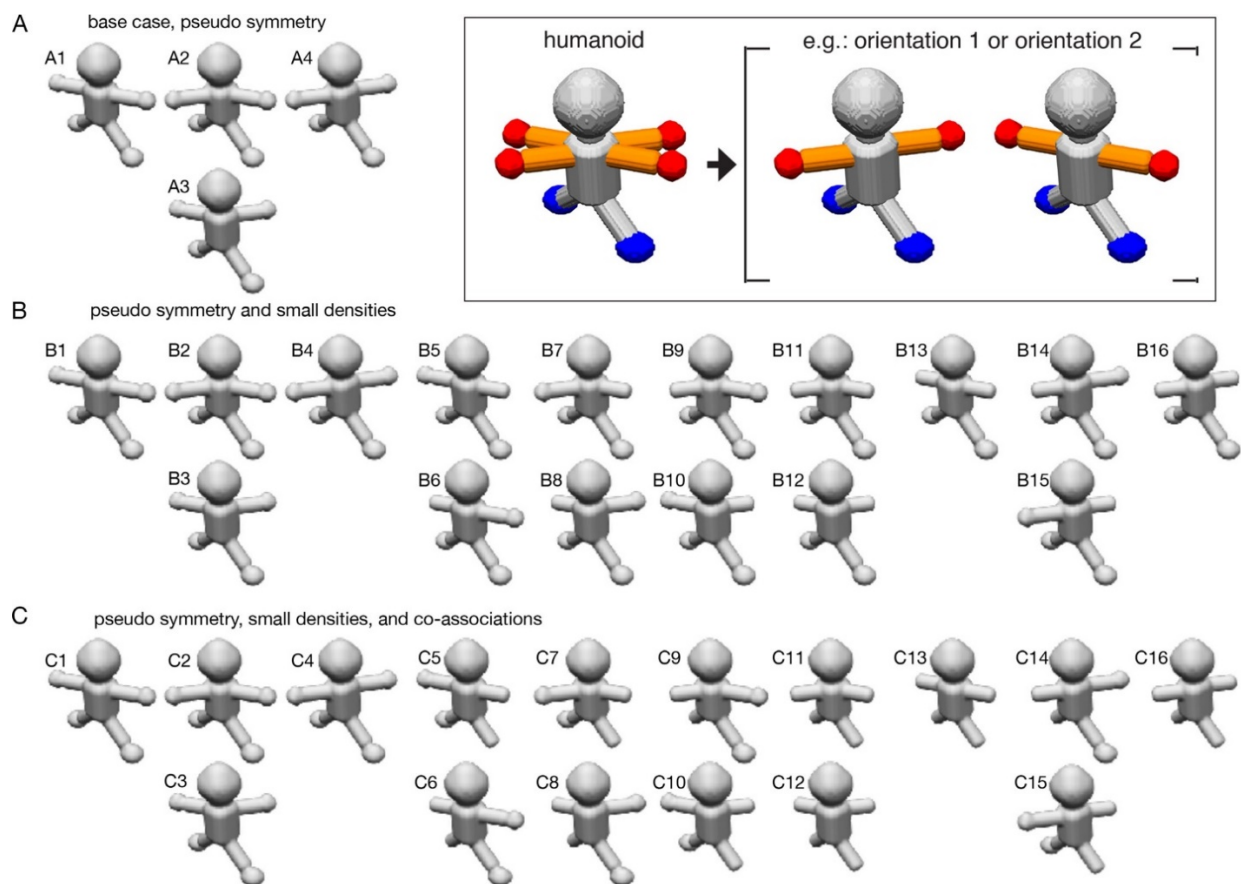
611
612

613 **Figure 1 – humanoid datasets and distinct scenarios used to assess focused classification**. Different
614 maps used to generate synthetic datasets described by the three scenarios are displayed. In each panel, A-
615 C, two maps which are degenerate and related to one another by a 180° rotation are positioned vertically
616 with respect to one another. The components used to generate the datasets are displayed in the inset, with
617 the heterogeneous elements colored (arms, orange; hands, red; feet, blue). (A) For the base scenario, only
618 the arms/hands are conformationally mobile. Four different combinations of maps lead to a dataset
619 characterized by two different asymmetric units. Maps A2/A3 are related by a 180° rotation. (B) For the
620 second scenario, in addition to the conformational mobility of the arms, the hands can be either present or
621 absent. 16 different combinations of maps lead to a dataset characterized by four different asymmetric
622 units. Maps B2/B3, B5/B6, B7/B8, B9/B10, B11/B12, and B14/B15 are related by a 180° rotation. (C)
623 For the third scenario, in addition to the conformational mobility of the arms, the hands can be either present
624 or absent, but their occupancy is *always* co-associated with a nearby foot. 16 different combinations of
625 maps lead to a dataset characterized by four different asymmetric units. Maps C2/C3, C5/C6, C7/C8,
626 C9/C10, C11/C12, and C14/C15 are related by a 180° rotation.
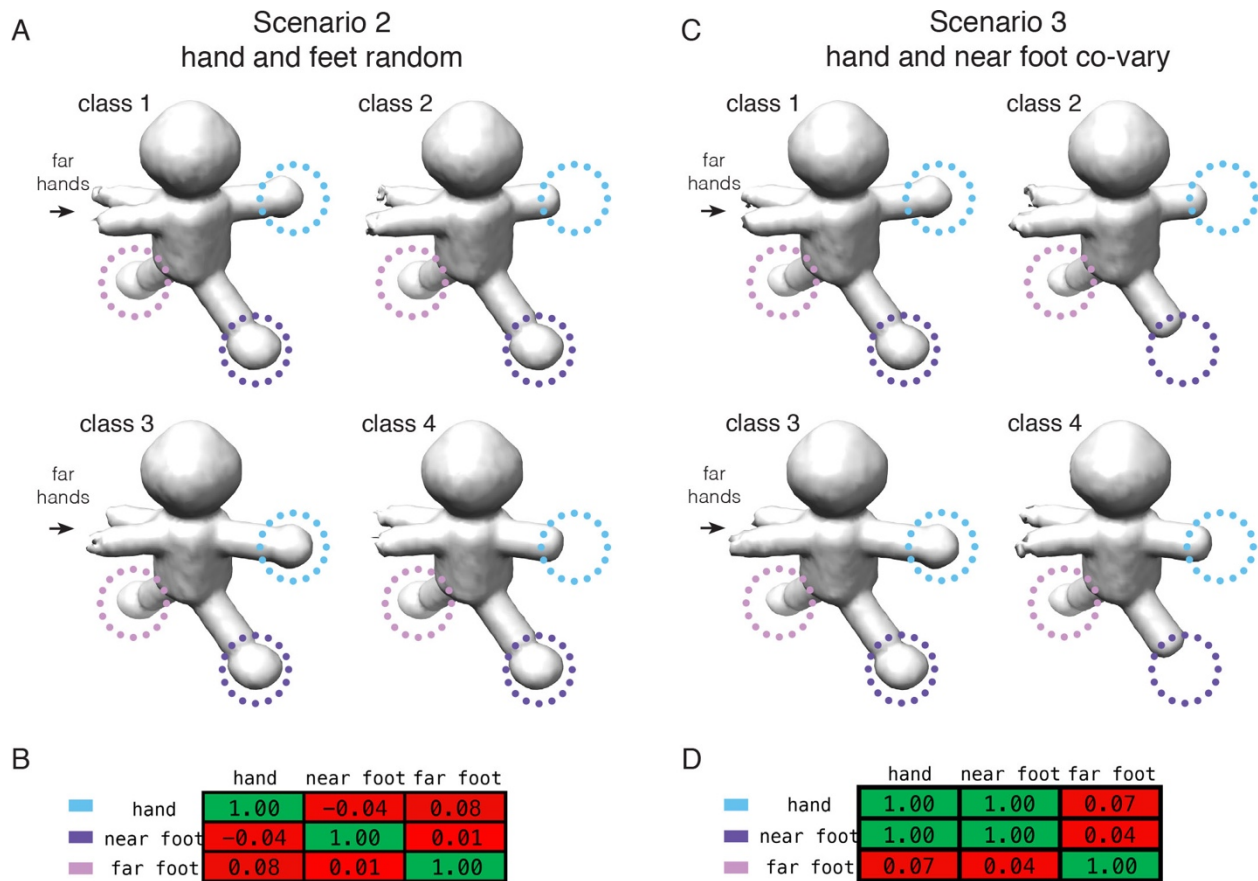
20

627
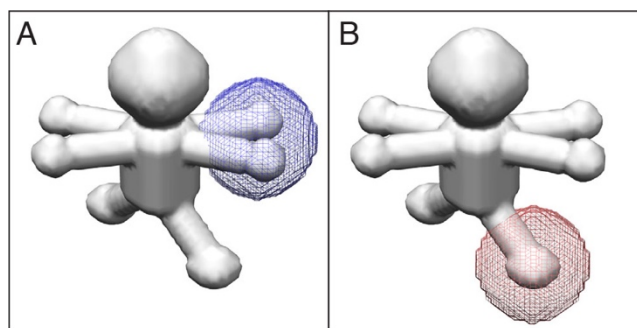
628

629

630   **Figure 2 – Application of masks onto regions of an asymmetric unit**. Masks were applied either (A)

631   onto the arm/hand region (blue) or (B) the leg region (red) prior to focused classification. Both types of

632   asymmetric units are displayed, showing both orientations of the arm/hand combinations.

633

634

635

636

Figure 3 – Evaluation of covariance within two different regions of a reconstructed object. Focused classifications using 2-D masks, applied to an arm/hand region (to the right of body in figure and encompassing both arm/hand conformations), were performed using either (A-B) the dataset for the 2nd scenario or (C-D) the dataset for the 3rd scenario, both at an SNR of 0.100. In all cases, four classes were recovered for the different asymmetric units (arms in two positions, each with and without a hand), and are displayed in panels A and C. (A) Volumes recovered from focused classification in the 2nd scenario, where all components are randomly occupied (control). (B) Normalized covariance matrix describing the relationships between the near hand, near foot, and far foot. (C) Volumes recovered from focused classification in the 3rd scenario, where the near hand is always co-associated with the near foot. (D) Normalized covariance matrix describing the relationships between the components. Near hand, where focused classification is performed, is circled in blue, near foot is circled in purple, and far foot is circled in pink. In the tables, the values are colored using a gradient: -1 (green, anti-correlated) < 0 (red, not correlated) < 1 (green, correlated).

650

22
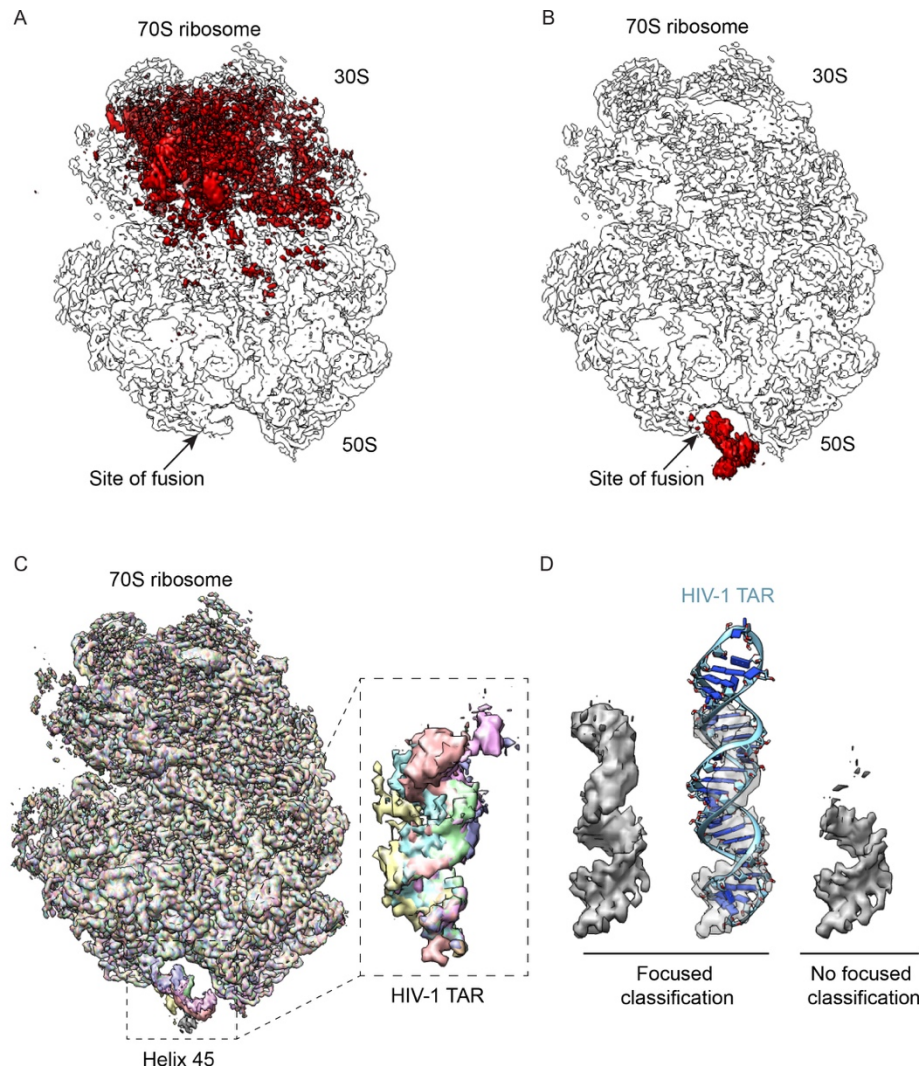
651
652
653
654 **Figure 4 – Experimental reconstructions highlighting the use of focused classification to analyze**
655 **highly heterogeneous datasets**. Bacterial 70S ribosomes containing an HIV-1 TAR element fused into
656 Helix 45 (H45) were used to analyze different classification approaches. (A) Combinatorial pairwise
657 differences between all 8 classes from global classification merged into a single volume to highlight the
658 overall variability. (B) Same as A, but from the result of focused classification using 2-D masks, applied
659 on the region of TAR fusion into H45. In both A-B, arrows denote the site of fusion. (C) Overlaid
660 reconstructions after focused classification, highlighting the differences within the TAR element, but not in
661 the rest of the ribosome. (D) Close-up of TAR reconstruction after deconvolving its mobility through
662 focused classifications (left), shown also with a rigid-body docking of the TAR element into density
663 (middle). A control reconstruction, without focused classification but using the same number of particles,
664 is displayed alongside (right).

23

SNR	Scenario 1 2-D mask	Scenario 1 3-D mask	Scenario 2 2-D mask	Scenario 2 3-D mask	Scenario 3 2-D mask	Scenario 3 3-D mask
0.100	0.99	0.91	0.85	0.70	0.87	0.71
0.050	0.96	0.85	0.73	0.56	0.75	0.61
0.025	0.87	0.74	0.42	0.41	0.46	0.39
0.013	0.72	0.57	0.21	0.17	0.21	0.17
0.006	0.47	0.36	0.09	0.08	0.08	0.09

665

666

667 **Table 1 – Results of focused classification on an asymmetric unit for the three different scenarios**.

668 Five different SNRs are evaluated, and the κ coefficient is displayed for the 2-D masking and 3-D masking

669 case for each of three scenarios.

670

SNR	2-D mask	3-D mask
0.100	0.23	-0.01
0.050	0.11	0.00
0.025	0.01	0.01
0.013	0.00	0.01
0.006	0.00	0.00
pure noise	-0.01	0.00

671

672

673 **Table 2 – Results of focused classification on an asymmetric unit when the mask is applied on the**

674 **wrong region.** Classification was performed after application of a 2-D mask or 3-D mask onto a leg (see

675 Figure 2B), while the heterogeneity was characterized by the mobility in the arms/hands (scenario 1), and

676 the κ coefficient was evaluated for the five SNRs and for each mask. Whereas the 2-D masking displayed

677 some "leakiness" at the highest SNRs, the 3-D masking showed completely random classification.

678
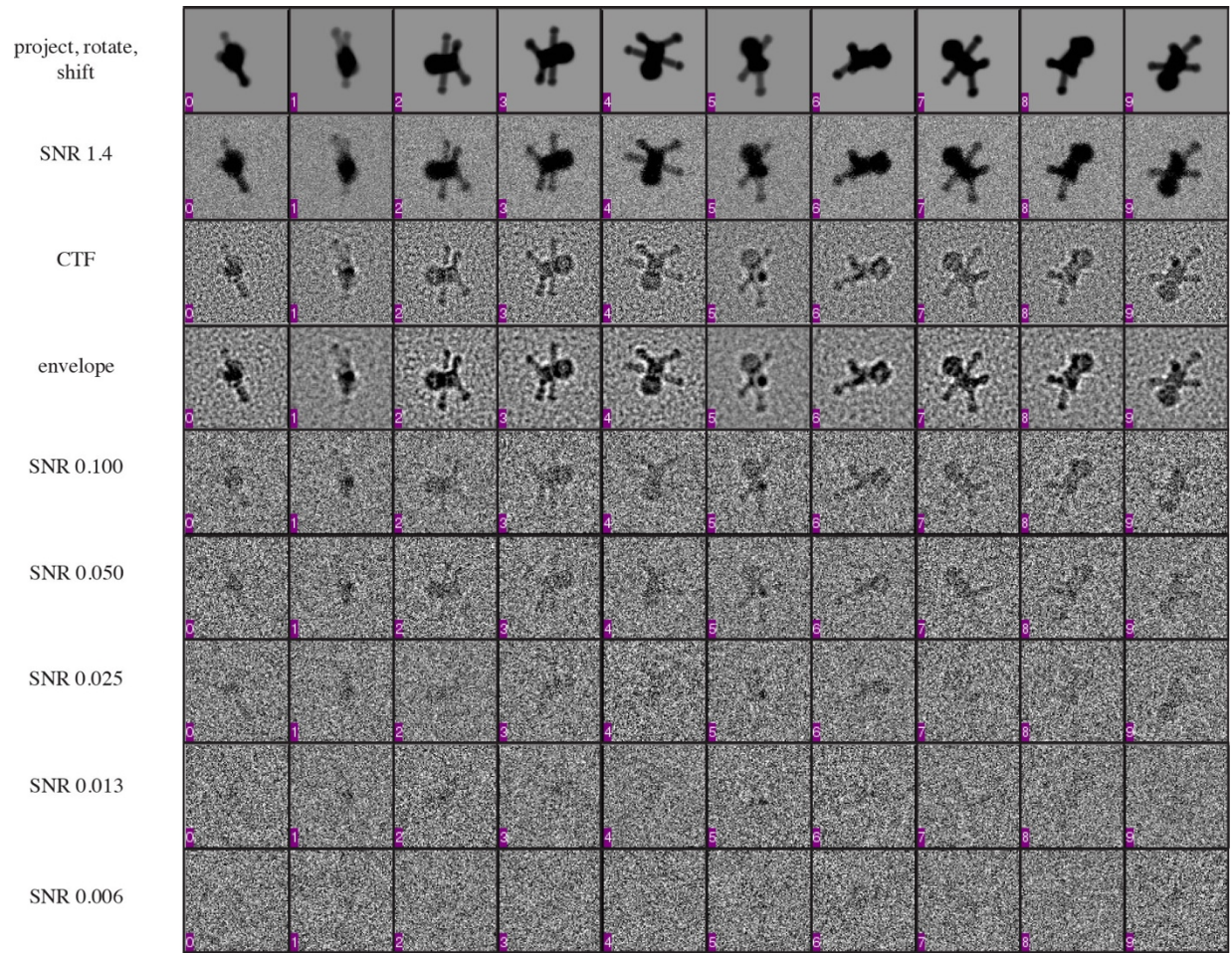
679
680
681 **Supplementary Figure 1 – synthetic data generated from the humanoid volumes**. Each volume was
682 randomly projected, rotated, and shifted. Noise was then applied to the projection images, followed by a
683 CTF and envelope function, and lastly the level of noise was brought down to one of five different levels
684 (0.100, 0.050, 0.025, 0.013, and 0.006), as previously described [8]. The different projections were then
685 randomly inserted into a 10,000-particle dataset for focused classification experiments.
686

26
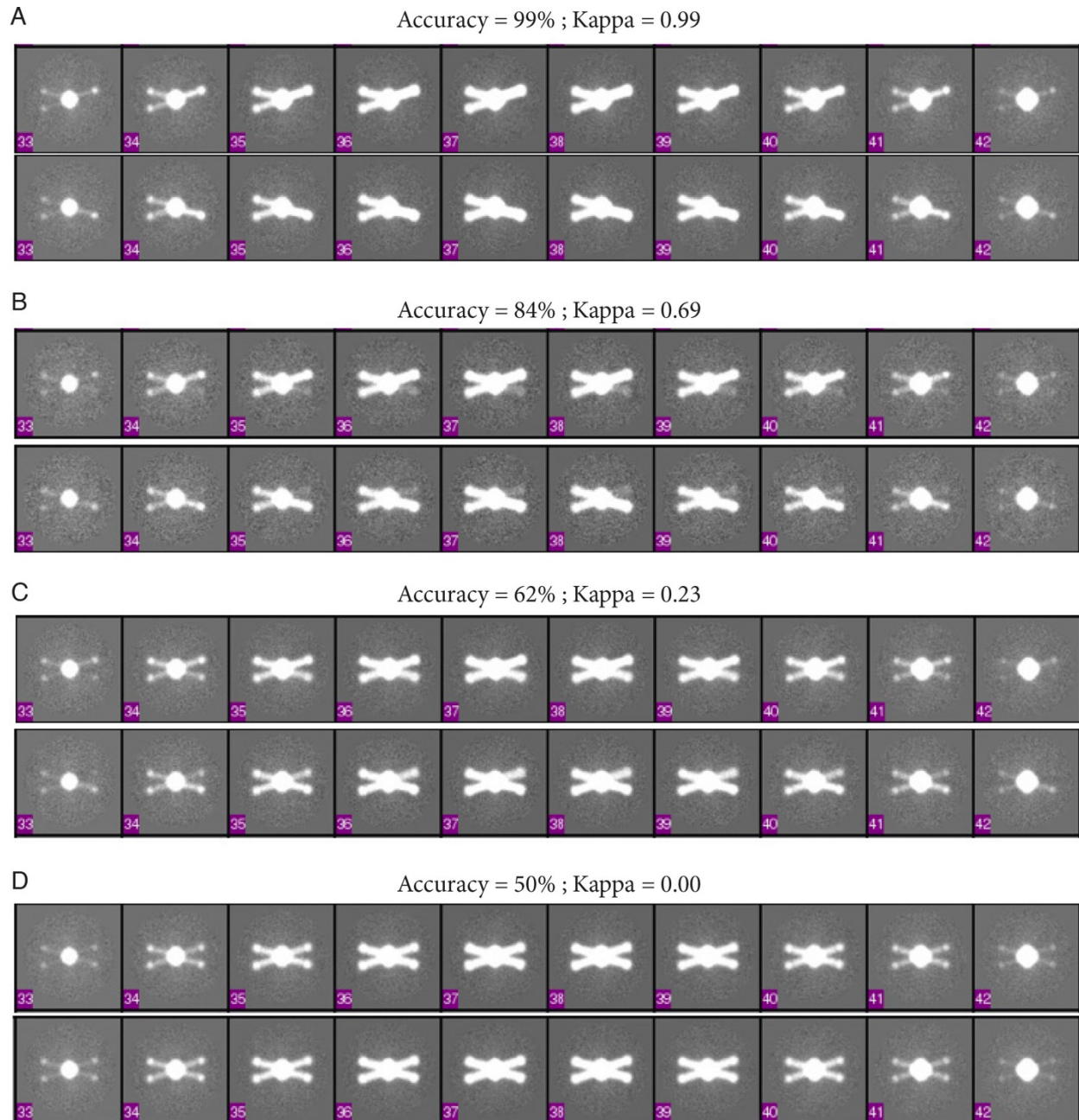
687
688
689 **Supplementary Figure 2 – visual demonstration of classification accuracy**. Slices through a
690 reconstruction are displayed for each panel (middle slices 33-42 within a 96-slice volume, for each of two
691 distinct classes [top and bottom]) around the Z-height of the arms. Classification was performed on the
692 right asymmetric unit and for the base dataset, where two different classes are expected. Ideally, only the
693 right arms would partition into one of several different classes. Classification was performed under four
694 different levels of noise, which resulted in distinct accuracies. Panels A-D demonstrate how the accuracies,
695 the associated κ coefficient, and the density varies with increasing errors. (A) Accuracy is nearly perfect,

27

696   κ is close to 1 and the two classes show complete distinction in the right arm region. (B) Accuracy is worse,

697   κ is has dropped to 0.69, and some contamination is evident in the opposing arm. (C) Accuracy has dropped

698   further, κ is close to 0, and the two volumes become virtually indistinguishable, although some differences

699   within the density amplitude point to residual heterogeneity. (D) Accuracy is completely random (50%

700   represents a coin toss when two possibilities are present), κ is correspondingly 0, and no difference in the

701   maps is evident.

702

A
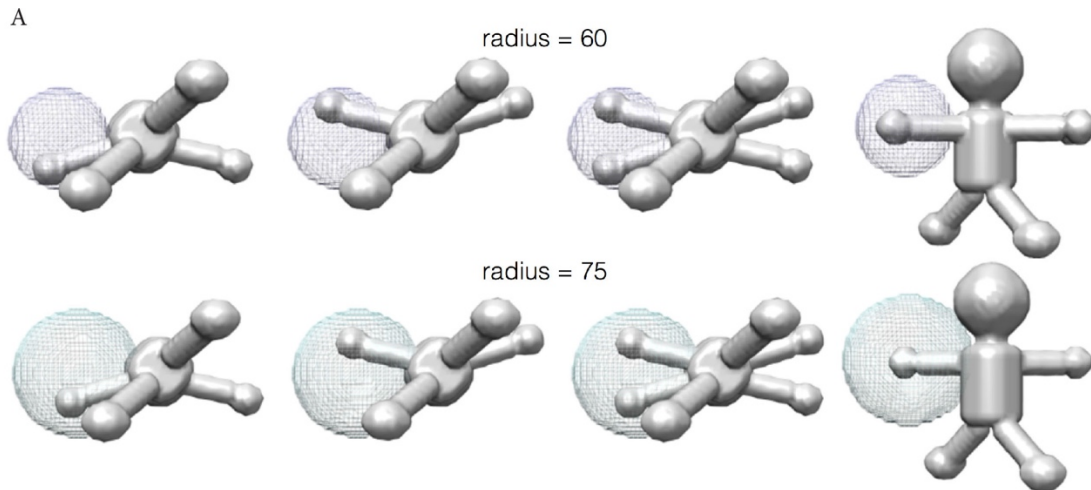
radius = 60

radius = 75

B

mask radius	Kappa coefficient
150	0.26
120	0.31
90	0.38
75	0.80
70	0.84
65	0.85
60	0.85
55	0.84
50	0.82

703

704

705 **Supplementary Figure 3 – titration of mask size used for focused classification**. Focused classification

706 parameters could be tuned for optimal performance with this particular dataset. Here, the mask size was
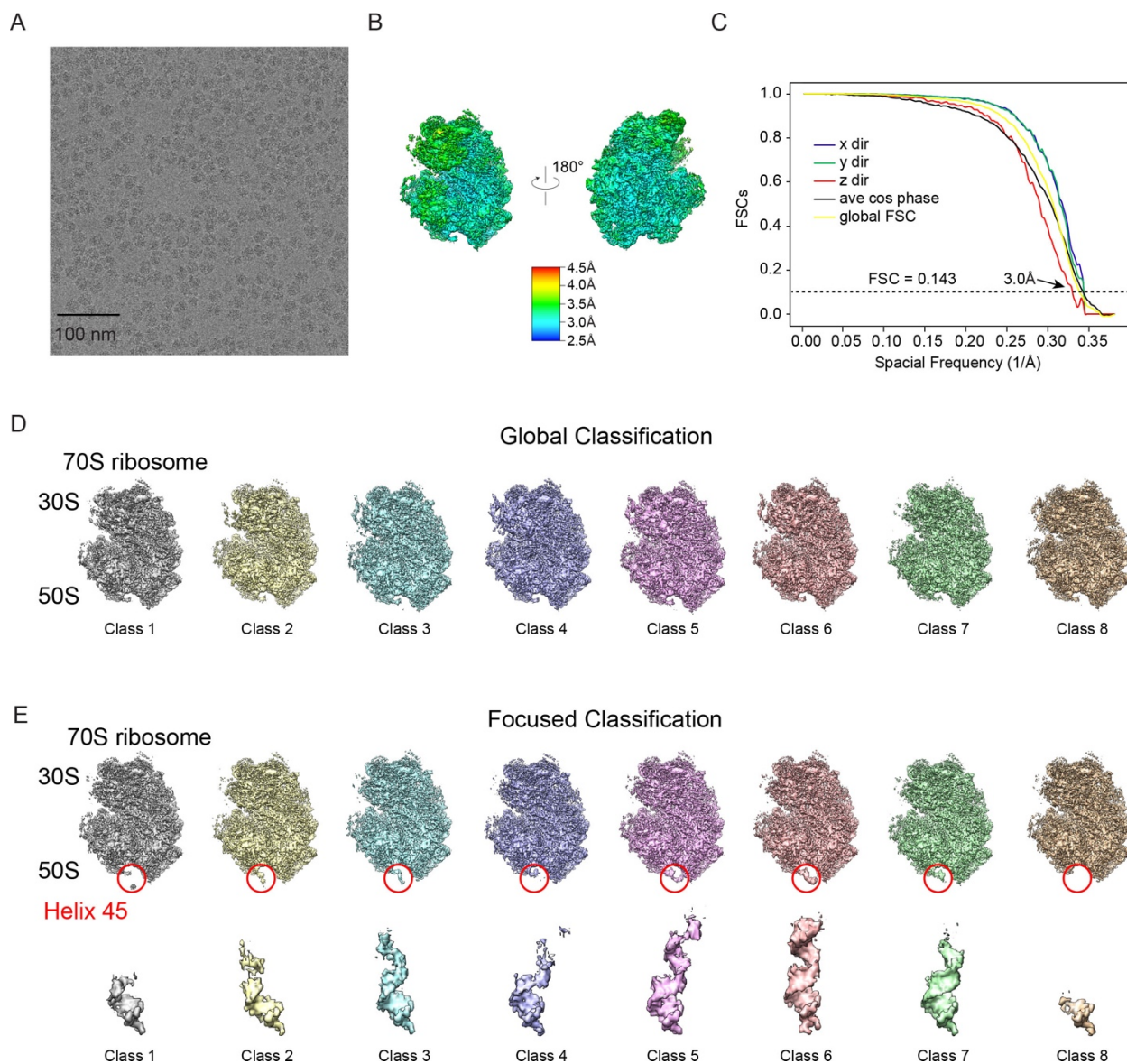
707 varied, and the results were followed by monitoring κ. (A) Two different mask sizes are displayed, applied

708 to an asymmetric unit around the arms/hands. (B) The results of focused classification with different mask

709 radii. Here, a 60 Å mask performs optimally, which effectively represents a tight mask that completely

710 encompasses only the mobile area.

711

712

713

714 **Supplementary Figure 4 – Cryo-EM data for HIV-1 TAR—ribosome fusions**. (A) Example raw image
715 collected for TAR-labeled ribosomes. (B) Initial single-model refinement, colored by local resolution and
716 (C) the corresponding FSC curves. (D) Classes generated from global 3-D classification showing a lack of
717 density in the region of helix 45. (E) Classes from focused 3-D classification, with the mask applied to the
718 region of TAR fusion, denoted by a red circle with the corresponding densities of the TAR hairpin in the
719 absence of the ribosome scaffold below.

720

721

References

722

723

724 [1] Y. Cheng, N. Grigorieff, P.A. Penczek, T. Walz, A primer to single-particle cryo-electron
725   microscopy, Cell. 161 (2015) 438–449. doi:10.1016/j.cell.2015.03.050.
726 [2] E. Nogales, The development of cryo-EM into a mainstream structural biology technique, Nature
727   Methods. 13 (2016) 24–27. doi:10.1038/nmeth.3694.
728 [3] A. Bartesaghi, C. Aguerrebere, V. Falconieri, S. Banerjee, L.A. Earl, X. Zhu, et al., Atomic
729   Resolution Cryo-EM Structure of β-Galactosidase, Structure. (2018).
730   doi:10.1016/j.str.2018.04.004.
731 [4] Y.Z. Tan, S. Aiyer, M. Mietzsch, J.A. Hull, R. McKenna, J. Grieger, et al., Sub-2 Å Ewald
732   Curvature Corrected Single-Particle Cryo-EM, bioRxiv. (2018) 305599. doi:10.1101/305599.
733 [5] K. Murata, M. Wolf, Cryo-electron microscopy for structural analysis of dynamic biological
734   macromolecules, Biochim. Biophys. Acta. 1862 (2018) 324–334.
735   doi:10.1016/j.bbagen.2017.07.020.
736 [6] S.H.W. Scheres, Processing of Structurally Heterogeneous Cryo-EM Data in RELION, in: The
737   Resolution Revolution: Recent Advances in cryoEM, Elsevier, 2016: pp. 125–157.
738   doi:10.1016/bs.mie.2016.04.012.
739 [7] H. Gao, M. Valle, M. Ehrenberg, J. Frank, Dynamics of EF-G interaction with the ribosome
740   explored by classification of a heterogeneous cryo-EM dataset, J. Struct. Biol. 147 (2004) 283–
741   290. doi:10.1016/j.jsb.2004.02.008.
742 [8] D. Lyumkis, A.F. Brilot, D.L. Theobald, N. Grigorieff, Likelihood-based classification of cryo-
743   EM images using FREALIGN, J. Struct. Biol. 183 (2013) 377–388.
744   doi:10.1016/j.jsb.2013.07.005.
745 [9] S.H.W. Scheres, H. Gao, M. Valle, G.T. Herman, P.P.B. Eggermont, J. Frank, et al.,
746   Disentangling conformational states of macromolecules in 3D-EM through likelihood
747   optimization, Nature Methods. 4 (2007) 27–29. doi:10.1038/nmeth992.
748 [10] C.M. Spahn, P.A. Penczek, Exploring conformational modes of macromolecular assemblies by
749   multiparticle cryo-EM, Curr. Opin. Struct. Biol. 19 (2009) 623–631.
750   doi:10.1016/j.sbi.2009.08.001.
751 [11] M. Valle, J. Sengupta, N.K. Swami, R.A. Grassucci, N. Burkhardt, K.H. Nierhaus, et al., Cryo-
752   EM reveals an active role for aminoacyl-tRNA in the accommodation process, Embo J. 21
753   (2002) 3557–3567. doi:10.1093/emboj/cdf326.
754 [12] Q. Zhou, X. Huang, S. Sun, X. Li, H.-W. Wang, S.-F. Sui, Cryo-EM structure of SNAP-SNARE
755   assembly in 20S particle, Cell Res. 25 (2015) 551. doi:10.1038/cr.2015.47.
756 [13] X.-C. Bai, E. Rajendra, G. Yang, Y. Shi, S.H.W. Scheres, Sampling the conformational space of
757   the catalytic subunit of human γ-secretase, Elife. 4 (2015). doi:10.7554/eLife.11182.
758 [14] P.A. Penczek, J. Frank, C.M. Spahn, A method of focused classification, based on the bootstrap
759   3D variance analysis, and its application to EF-G-dependent translocation, J. Struct. Biol. 154
760   (2006) 184–194. doi:10.1016/j.jsb.2005.12.013.
761 [15] S.L. Ilca, A. Kotecha, X. Sun, M.M. Poranen, D.I. Stuart, J.T. Huiskonen, Localized
762   reconstruction of subunits from electron cryomicroscopy images of macromolecular complexes,
763   Nat Commun. 6 (2015) 8843. doi:10.1038/ncomms9843.
764 [16] P.D. Abeyrathne, C.S. Koh, T. Grant, N. Grigorieff, A.A. Korostelev, S. Subramaniam,
765   Ensemble cryo-EM uncovers inchworm-like translocation of a viral IRES through the ribosome,
766   Elife. 5 (2016) e14874. doi:10.7554/eLife.14874.
767 [17] O. von Loeffelholz, S.K. Natchiar, N. Djabeur, A.G. Myasnikov, H. Kratzat, J.-F. Ménétret, et
768   al., Focused classification and refinement in high-resolution cryo-EM structural analysis of
769   ribosome complexes, Curr. Opin. Struct. Biol. 46 (2017) 140–148.
770   doi:10.1016/j.sbi.2017.07.007.
771 [18] A.B. Loveland, G. Demo, N. Grigorieff, A.A. Korostelev, Ensemble cryo-EM elucidates the

31

772   mechanism of translation fidelity, Nature. 546 (2017) 113–117. doi:10.1038/nature22397.

773   [19]   T.H.D. Nguyen, W.P. Galej, X.-C. Bai, C.G. Savva, A.J. Newman, S.H.W. Scheres, et al., The
774   architecture of the spliceosomal U4/U6.U5 tri-snRNP, Nature. 523 (2015) 47–52.
775   doi:10.1038/nature14548.

776   [20]   H.E. Autzen, A.G. Myasnikov, M.G. Campbell, D. Asarnow, D. Julius, Y. Cheng, Structure of
777   the human TRPM4 ion channel in a lipid nanodisc, Science. 359 (2018) 228–232.
778   doi:10.1126/science.aar4510.

779   [21]   M.L. Oldham, N. Grigorieff, J. Chen, Structure of the transporter associated with antigen
780   processing trapped by herpes simplex virus, Elife. 5 (2016) 213. doi:10.7554/eLife.21829.

781   [22]   N. Grigorieff, Frealign: An Exploratory Tool for Single-Particle Cryo-EM, Meth. Enzymol. 579
782   (2016) 191–226. doi:10.1016/bs.mie.2016.04.013.

783   [23]   A. Ballandras-Colas, M. Brown, N.J. Cook, T.G. Dewdney, B. Demeler, P. Cherepanov, et al.,
784   Cryo-EM reveals a novel octameric integrase structure for betaretroviral intasome function,
785   Nature. 530 (2016) 358–361. doi:10.1038/nature16955.

786   [24]   T. Grant, A. Rohou, N. Grigorieff, cisTEM, user-friendly software for single-particle image
787   processing, Elife. 7 (2018) e14874. doi:10.7554/eLife.35383.

788   [25]   J.T. Huiskonen, H.T. Jäälinoja, J.A.G. Briggs, S.D. Fuller, S.J. Butcher, Structure of a hexameric
789   RNA packaging motor in a viral polymerase complex, J. Struct. Biol. 158 (2007) 156–164.
790   doi:10.1016/j.jsb.2006.08.021.

791   [26]   D.O. Passos, M. Li, R. Yang, S.V. Rebensburg, R. Ghirlando, Y. Jeon, et al., Cryo-EM
792   structures and atomic model of the HIV-1 strand transfer complex intasome, Science. 355 (2017)
793   89–92. doi:10.1126/science.aah5163.

794   [27]   J.T. Huiskonen, Image processing for cryogenic transmission electron microscopy of symmetry-
795   mismatched complexes, Bioscience Reports. (2018) BSR20170203. doi:10.1042/BSR20170203.

796   [28]   R.I. Koning, J. Gomez-Blanco, I. Akopjana, J. Vargas, A. Kazaks, K. Tars, et al., Asymmetric
797   cryo-EM reconstruction of phage MS2 reveals genome structure *in situ*, Nat Commun. 7
798   (2016) 12524. doi:10.1038/ncomms12524.

799   [29]   S.J. Ludtke, P.R. Baldwin, W. Chiu, EMAN: semiautomated software for high-resolution single-
800   particle reconstructions, J. Struct. Biol. 128 (1999) 82–97. doi:10.1006/jsbi.1999.4174.

801   [30]   Q. Liu, K. Fredrick, Contribution of intersubunit bridges to the energy barrier of ribosomal
802   translocation, Nucl. Acids Res. 41 (2013) 565–574. doi:10.1093/nar/gks1074.

803   [31]   D. Qin, N.M. Abdi, K. Fredrick, Characterization of 16S rRNA mutations that decrease the
804   fidelity of translation initiation, Rna. 13 (2007) 2348–2355. doi:10.1261/rna.715307.

805   [32]   T. Asai, D. Zaporojets, C. Squires, C.L. Squires, An Escherichia coli strain with all
806   chromosomal rRNA operons inactivated: Complete exchange of rRNA genes between bacteria,
807   Proceedings of the National Academy of Sciences. 96 (1999) 1971–1976.
808   doi:10.1073/pnas.96.5.1971.

809   [33]   P. Gay, D. Le Coq, M. Steinmetz, T. Berkelman, C.I. Kado, Positive selection procedure for
810   entrapment of insertion sequence elements in gram-negative bacteria, J. Bacteriol. 164 (1985)
811   918–921.

812   [34]   J. Chiu, P.E. March, R. Lee, D. Tillett, Site-directed, Ligase-Independent Mutagenesis (SLIM): a
813   single-tube methodology approaching 100% efficiency in 4 h, Nucl. Acids Res. 32 (2004) e174–
814   e174. doi:10.1093/nar/gnh172.

815   [35]   C. Suloway, J. Pulokas, D. Fellmann, A. Cheng, F. Guerra, J. Quispe, et al., Automated
816   molecular microscopy: the new Leginon system, J. Struct. Biol. 151 (2005) 41–60.
817   doi:10.1016/j.jsb.2005.03.010.

818   [36]   G.C. Lander, S.M. Stagg, N.R. Voss, A. Cheng, D. Fellmann, J. Pulokas, et al., Appion: an
819   integrated, database-driven pipeline to facilitate EM image processing, J. Struct. Biol. 166
820   (2009) 95–102.

821   [37]   S.Q. Zheng, E. Palovcak, J.-P. Armache, K.A. Verba, Y. Cheng, D.A. Agard, MotionCor2:
822   anisotropic correction of beam-induced motion for improved cryo-electron microscopy, Nature

823   Methods. 14 (2017) 331–332. doi:10.1038/nmeth.4193.

824   [38]   T. Grant, N. Grigorieff, Measuring the optimal exposure for single particle cryo-EM using a 2.6
825   Å reconstruction of rotavirus VP6, Elife. 4 (2015) e06980. doi:10.7554/eLife.06980.

826   [39]   A. Rohou, N. Grigorieff, CTFFIND4: Fast and accurate defocus estimation from electron
827   micrographs, J. Struct. Biol. 192 (2015) 216–221. doi:10.1016/j.jsb.2015.08.008.

828   [40]   A.M. Roseman, FindEM--a fast, efficient program for automatic selection of particles from
829   electron micrographs, J. Struct. Biol. 145 (2004) 91–99.

830   [41]   K. Zhang, Gctf: Real-time CTF determination and correction, J. Struct. Biol. 193 (2016) 1–12.
831   doi:10.1016/j.jsb.2015.11.003.

832   [42]   D. Kimanius, B.O. Forsberg, S.H. Scheres, E. Lindahl, S. Subramaniam, Accelerated cryo-EM
833   structure determination with parallelisation using GPUs in RELION-2, Elife. 5 (2016) e18722.
834   doi:10.7554/eLife.18722.

835   [43]   S.H.W. Scheres, RELION: Implementation of a Bayesian approach to cryo-EM structure
836   determination, J. Struct. Biol. 180 (2012) 519–530. doi:10.1016/j.jsb.2012.09.006.

837   [44]   E.F. Pettersen, T.D. Goddard, C.C. Huang, G.S. Couch, D.M. Greenblatt, E.C. Meng, et al.,
838   UCSF Chimera--a visualization system for exploratory research and analysis, J Comput Chem.
839   25 (2004) 1605–1612. doi:10.1002/jcc.20084.

840   [45]   G. Harauz, M. van Heel, Exact filters for general geometry 3-dimensional reconstruction, Optik.
841   73 (1986) 146–156.

842   [46]   Y.Z. Tan, P.R. Baldwin, J.H. Davis, J.R. Williamson, C.S. Potter, B. Carragher, et al.,
843   Addressing preferred specimen orientation in single-particle cryo-EM through tilting, Nature
844   Methods. 14 (2017) 793–796. doi:10.1038/nmeth.4347.

845   [47]   M. Hohn, G. Tang, G. Goodyear, P.R. Baldwin, Z. Huang, P.A. Penczek, et al., SPARX, a new
846   environment for Cryo-EM image processing, J. Struct. Biol. 157 (2007) 47–55.
847   doi:10.1016/j.jsb.2006.07.003.

848   [48]   N. Fischer, P. Neumann, A.L. Konevega, L.V. Bock, R. Ficner, M.V. Rodnina, et al., Structure
849   of the E. coli ribosome-EF-Tu complex at <3 Å resolution by Cs-corrected cryo-EM, Nature.
850   520 (2015) 567–570. doi:10.1038/nature14275.

851   [49]   P. Emsley, B. Lohkamp, W.G. Scott, K. Cowtan, Features and development of Coot, Acta
852   Crystallogr. D Biol. Crystallogr. 66 (2010) 486–501. doi:10.1107/S0907444910007493.

853   [50]   N.R. Voss, D. Lyumkis, A. Cheng, P.-W. Lau, A. Mulder, G.C. Lander, et al., A toolbox for ab
854   initio 3-D reconstructions in single-particle electron microscopy, J. Struct. Biol. 169 (2010) 389–
855   398. doi:10.1016/j.jsb.2009.12.005.

856   [51]   J. Cohen, A coefficient of agreement for nominal scales, Educational and Psychologicla
857   Measurement. XX (1960) 37–46.

858   [52]   W.J. Youden, Index for rating diagnostic tests, Cancer. 3 (1950) 32–35. doi:10.1002/1097-
859   0142(1950)3:1<32::AID-CNCR2820030106>3.0.CO;2-3.

860   [53]   A.B. Loveland, A.A. Korostelev, Structural dynamics of protein S1 on the 70S ribosome
861   visualized by ensemble cryo-EM, Methods. 137 (2018) 55–66.
862   doi:10.1016/j.ymeth.2017.12.004.

863   [54]   Y. Liu, S. Gonen, T. Gonen, T.O. Yeates, Near-atomic cryo-EM imaging of a small protein
864   displayed on a designed scaffolding system, Proceedings of the National Academy of Sciences.
865   115 (2018) 3362–3367. doi:10.1073/pnas.1718825115.

866   [55]   T.G. Martin, T.A.M. Bharat, A.C. Joerger, X.-C. Bai, F. Praetorius, A.R. Fersht, et al., Design of
867   a molecular support for cryo-EM structure determination, Proc. Natl. Acad. Sci. U.S.a. 113
868   (2016) E7456–E7463. doi:10.1073/pnas.1612720113.

869   [56]   X. Agirrezabala, J. Lei, J.L. Brunelle, R.F. Ortiz-Meoz, R. Green, J. Frank, Visualization of the
870   hybrid state of tRNA binding promoted by spontaneous ratcheting of the ribosome, Mol. Cell. 32
871   (2008) 190–197. doi:10.1016/j.molcel.2008.10.001.

872   [57]   D. Lyumkis, D. Oliveira Dos Passos, E.B. Tahara, K. Webb, E.J. Bennett, S. Vinterbo, et al.,
873   Structural basis for translational surveillance by the large ribosomal subunit-associated protein

874   quality control complex, Proceedings of the National Academy of Sciences. 111 (2014) 15981–
875   15986. doi:10.1073/pnas.1413882111.
876