



A methodology using Gaussian-based density map approximation to assess sets of cryo-electron microscopy density maps

Slavica Jonić

► To cite this version:

Slavica Jonić. A methodology using Gaussian-based density map approximation to assess sets of cryo-electron microscopy density maps. *Journal of Structural Biology*, In press, 10.1016/j.jsb.2018.07.014 . hal-01853731

HAL Id: hal-01853731

<https://hal.science/hal-01853731>

Submitted on 3 Aug 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A methodology using Gaussian-based density map approximation to assess sets of cryo-electron microscopy density maps

Slavica Jonić

Sorbonne Université, UMR CNRS 7590, Muséum National d'Histoire Naturelle, IRD, Institut de
Minéralogie, de Physique des Matériaux et de Cosmochimie, IMPMC, 75005 Paris, France.

All correspondence in relation to this article should be sent to:

Slavica Jonic, IMPMC-UMR CNRS 7590, Sorbonne Université, CC 115, 4 Place Jussieu, 75005 Paris,
France; Phone: +33 1 44 27 72 05; Fax: +33 1 44 27 37 85; E-mail: Slavica.Jonic@upmc.fr

Running title: Gaussian-based density map assessment

Keywords: cryo-electron microscopy; density map; assessment; Gaussian functions; macromolecular
complex; structure

Abstract:

This article presents a methodology to assess a set of density maps, as used in the Blind Assessment Phase of the 2015/2016 Map Challenge (EMDataBank Validation Challenges). The synthetic and experimental cryo-electron microscopy (cryo-EM) density maps obtained by different single particle analysis protocols and by different participants, submitted in the Challenge Phase for assessment, were analyzed with this methodology and the obtained results are presented and discussed here. The goal of using such a methodology was to blindly identify the density maps with globally similar structural information, meaning the maps with the structural information mostly “reproduced” by different protocols. To this end, the density maps are “coarsened” using Gaussian-based approximations, with the same input approximation parameters for all maps of the target biological complex. The approximated maps are then represented in a common reduced-dimension (here, 3D) space of their correlation-coefficient-based distances, in which close maps mean similar maps. The distance analysis allows identifying maps with the most “reproduced” structural information by different protocols. The obtained results are also discussed taking into account the detailed information about the protocols that has been released after the end of the Blind Assessment Phase.

INTRODUCTION

Single particle cryo-electron microscopy (cryo-EM) is a powerful structural biology technique that has proved to be useful in determining near-atomic-resolution structures of various macromolecular complexes (Bai et al., 2015; Banerjee et al., 2016; Bartesaghi et al., 2015; Hospenthal et al., 2016; Khatter et al., 2015; Liao et al., 2013; Yu et al., 2008; Zhang et al., 2008). The near-atomic resolutions of the density maps obtained by cryo-EM should facilitate deriving of *de novo* structural models from the density maps (Baker et al., 2010; Wang et al.,

2015). However, *de novo* modeling from density maps is not an easy task and the major question is how to assess the actual quality of the obtained cryo-EM maps.

The standard method to assess the quality of cryo-EM density maps is Fourier Shell Correlation (FSC) (Harauz and van Heel, 1986). When a structure of the same or similar complex could have been obtained at atomic resolution by other experimental techniques such as X-ray crystallography or NMR, this structure is first aligned (via rigid-body or elastic alignment) with the cryo-EM density map and the density map simulated from this atomic-resolution structure is then compared with the cryo-EM density map over different spatial frequencies using the FSC (i.e., the normalized cross-correlation coefficient between the two density maps over corresponding shells in Fourier space). When no atomic-resolution structure is available, the FSC is used to compare two density maps computed independently from two subsets of the same set of images that yielded the resulting density map. The FSC (between 0 and 1) informs about consistency of the density maps that are being compared over spatial frequencies. When the FSC is computed between the density maps obtained from subsets of images, it informs about reproducibility of structural reconstruction over different image subsets. Yet, the inverse spatial frequency corresponding to a FSC threshold value is commonly reported as the resolution of the density map (as the frequency to which the structural information can be considered reliable). Regarding the FSC threshold, the value of 0.143 is used when the alignment of each half-subset of particles is conducted independently (currently, the standard procedure), whereas the value of 0.5 has traditionally been used when all particles were aligned together.

When a density map is obtained with an image analysis protocol, the question is also whether some other image analysis protocol could result in a density map of better quality. Indeed,

different image processing strategies (different regarding image processing steps, algorithms, or parameters used) may result in more or less different density maps from the same set of images. For instance, the image analysis protocols usually involve removal of some particle images, in order to keep for the final reconstruction only images of the best quality and those that are the most consistent among each other regarding their content (structural information). However, different image analysis protocols use different criteria or combinations of them to decide which images should be kept or removed. Depending on how many and which images are kept for the final reconstruction, the quality of the reconstructed density map may be better or worse.

Taking into account all of this, the question is what strategies one could use to assess quality of density maps reconstructed using different image analysis protocols and software. In this article, we present one such strategy. This strategy has been used in the 2015/2016 Map Challenge to blindly assess sets of density maps reconstructed by different researchers using different single particle analysis protocols and software (the maps were made available by the participants of the Challenge Phase for assessment by the participants of the Blind Assessment Phase). This strategy allows identifying maps with most “reproduced” structural information by different protocols and software. To this end, the density maps are “coarsened” using Gaussian-based approximations, with the same input approximation parameters for all maps of the target biological complex. The approximated maps are then represented in a common reduced-dimension (here, 3D) space of their correlation-coefficient-based distances. In this space, the closest maps correspond to the most similar maps and can be interpreted as containing the structural information that has been “reproduced” by the protocols/software. If several maps gather close to each other in this space, the structural information in these maps can be considered as “consensus”.

This article also presents the results of using this methodology to analyze 66 density maps submitted by 27 participants. A set of 7-13 density maps was available for each one of 7 target biological systems. Six sets of density maps were obtained from experimental images and one from simulated (*in silico*) images. The density maps from experimental images were obtained for β -Galactosidase (BetaGal) (Bartesaghi et al., 2014), Brome Mosaic Virus (BMV) (Wang et al., 2014), Apo-ferritin (Russo and Passmore, 2014), T20S proteasome (Campbell et al., 2015), 80S ribosome (Wong et al., 2014), and TRPV1 channel (Liao et al., 2013). The density maps from simulated images were obtained for GroEL (Vulovic et al., 2013). The maps and associated metadata submitted to the Challenge are available at <https://doi.org/10.5281/zenodo.1185426>.

This article also includes results of map distance analysis using original maps (maps without Gaussian-based approximations). These results are shown to be globally consistent with the results of using Gaussian-based map approximations, with a few exceptions that indicate that Gaussian-based map approximation facilitates the analysis. Indeed, the use of Gaussian functions for map coarsening facilitates the comparison of maps not only because it reduces the noise in the original maps but also because it “uniformizes” the density representation of different maps of the same target system (the Gaussian functions are used as the basis functions of the map approximation).

The obtained results are discussed taking into account the detailed information on the protocols and software used to obtain the maps that has been released after the end of the Challenge.

METHODS

This section describes the methodology to assess similarities and dissimilarities among a set of density maps of the same target biological macromolecular complex to identify maps with most reproduced structural information by different image analysis protocols and software for density map reconstruction from cryo-EM single particle images.

This methodology is recommended to be used with “unfiltered” density maps. It is inspired by the one proposed in (Sanchez Sorzano et al., 2016). This previously published methodology has been proposed in a different context, which is to study continuous conformational changes of complexes (Sanchez Sorzano et al., 2016). Contrary to this previously published methodology that is based on flexible alignment among pairs of density maps (via elastic deformation of one map to fit the other map in the pair), the methodology proposed here analyzes similarities among rigid-body aligned maps approximated using 3D Gaussian functions. The Gaussian-based map approximations, obtained using the method proposed in (Jonic and Sorzano, 2016a), can be seen as denoised (Jonic et al., 2016) or, generally, simplified or approximate (“coarse”) versions of the given density maps. Given a standard deviation of Gaussian functions and a target approximation error, the Gaussian-based map approximation method aims at determining the number of Gaussian functions, their positions, and weights such that the map approximation error is below the target approximation error. The standard deviation of Gaussian functions and the target approximation error are the main input parameters to the Gaussian-based map approximation method as they usually affect the maximum achievable accuracy of the map approximation and therefore require adjustments. The other input approximation parameters are the initial number of Gaussian functions, the speed of adding Gaussian functions, and the minimum distance between the Gaussian functions. These additional parameters can be kept to their default values that are small enough to avoid affecting the maximum achievable accuracy

of the map approximation (they only affect the speed of convergence of the algorithm (Jonic and Sorzano, 2016a)). More details about the Gaussian-based map approximation method are provided in (Jonic and Sorzano, 2016a) and in **Supporting Material** here.

Given a set of density maps, one of them is used as a common reference to rigid-body align and resize the other maps so that they can be compared (the maps were resized to have the same number of voxels and the same voxel size as the reference map). For the majority of target complexes in this article, the reference map was one of the smallest-size maps. This means that the size of other maps was typically reduced to the size of the reference map. To speed up computations in two cases of very large maps, the density-map size was additionally reduced by a factor of 2 in each dimension (from 420^3 to 210^3 voxels for BMV and from 380^3 to 190^3 voxels for 80S ribosome, **Table 1**).

In the next step, all density maps of the same target complex are approximated with Gaussian functions, using the same input approximation parameters. This Gaussian-based map approximation is done within a mask created from the reference map using a selected density level threshold and morphological operators. This means that a great amount of background noise that may be present and different in different density maps is not incorporated in the Gaussian-based approximated map. The parameters used for the Gaussian-based approximation are summarized in **Table 1**. All other parameters of the Gaussian-based approximation method were set to their default values (the initial number of Gaussian functions to 300, the speed of adding Gaussian functions to 30 %, and the minimum distance between the Gaussian functions to 0.001 voxel) (Jonic and Sorzano, 2016a).

Then, Pearson correlation coefficient (CC) is calculated among the approximated maps within the area determined by the mask that was used for the map approximation. For n given density maps, the n -by- n matrix of distances (dissimilarities) among the maps ($1-CC$) is then projected onto a low-dimensional space (1D, 2D, or 3D space is usually used for visualization) using non-metric multidimensional scaling method (Cox and Cox, 2001) that creates a configuration of points whose Euclidean distances approximate a monotonic transformation of the original dissimilarities (the method approximates the ranks of the dissimilarities instead of the dissimilarity values, meaning that it creates a configuration of points with large interpoint distances corresponding to large dissimilarities and small interpoint distances corresponding to small dissimilarities). The maps are then visualized in this low-dimensional distance space. As in (Sanchez Sorzano et al., 2016), the non-metric multidimensional scaling was performed with the MATLAB *mdscale* method using the default values of its parameters, which results in a unique low-dimensional projection of the original (higher-dimensional) distance space. However, it should be noted that, with some other parameters (e.g., a random initial configuration of points or a different maximum number of iterations), a different projection of the same original distance space would be obtained for each run of the method.

We here use 3D distance spaces to approximate the original distance spaces because a 3D approximation is generally more accurate than a 2D or 1D approximation (the errors of 2D and 3D approximations for the data used in this article are compared in **Table S9 in Supporting Material**).

Finally, the maps gathering closest to each other in a localized area of the low-dimensional distance space are identified as the maps with the most reproduced structural information by different image analysis protocols and software.

RESULTS

This section presents the blind analysis results obtained using the methodology described in **Methods** section and 7 available sets of density maps deposited as “unfiltered” in the Challenge Phase of the 2015/2016 Map Challenge. For each of the 7 target biological systems, we provide a figure that shows a 3D space of distances among the corresponding Gaussian-based approximated maps (**Figs. 1-7**). The figures of corresponding 3D representations of distances among the original density maps (without Gaussian-based approximations) are provided in **Supporting Material (Figs. S1, S3, S5, S7, S9, S11, S13)**. The mentioned figures show the distance spaces with automatically determined axes limits that allow to better visualize smaller distances among maps. The distance spaces (among the original maps and among their Gaussian-based approximations) visualized using the same axes limits for all 7 target systems are also provided and allow to better visualize larger distances among maps (**Figs. S15 and S16 in Supporting Material**). The distances among original maps and the distances among their Gaussian-based approximations, before and after projecting the maps to 3D distance space, are also part of **Supporting Material (Tables S10-S37)**. **Supporting Material** also includes Chimera visualization of the original density maps (**Figs. S2, S4, S6, S8, S10, S12, S14**), their Gaussian-based approximations (**Figs. S17-S23**), the contours of the masks used for the map approximation and for computing *CC* for the approximated and original maps (**Fig. S24**), as well as the number of 3D Gaussian functions and the density-map approximation error obtained using these Gaussian functions (**Tables S1-S7**) for the parameters provided in **Table 1**.

It should be noted that the same target approximation error (10 %) was first tried for all 7 biological systems but with no success to approximate the majority of the density maps of some systems up to this error, even when reducing the standard deviation of Gaussian functions.

Therefore, a smaller target approximation error (5 %) was used for these systems (**Table 1**), which allowed the density map approximations to at least up to around 10% in the majority of cases (**Tables S1-S7**). Even if the approximation error of some density maps is much higher than this “virtual” target approximation error of 10%, this does not prevent the method to distinguish similar from dissimilar maps. On the contrary, very different approximation errors among the maps of the same system would mean that the densities in these maps are really different and accentuating these differences using such “poor” Gaussian-based approximation makes the method distinguish more easily among different and similar maps. An example of this is the case of GroEL density maps whose approximation errors are higher for maps 158, 168, and 169 than for other maps (~20-45% for maps 158, 168, 169 and ~10% for other maps, for the Gaussian standard deviation of 1.5 voxel and the target approximation error of 10%, **Table S8**) and the method detected these maps as being different from other maps (**Fig. 4 and Fig. S15D**). Though a smaller standard deviation of Gaussian functions (1 voxel) and a smaller target approximation error (5 %) reduce the approximation error of these 3 density maps (**Table S8**) and make that some density maps look more different or more similar than for larger values of the two parameters (e.g., maps 104 and 132 or maps 132 and 158, **Fig. S25**), these differences are not significant (the distance between maps 104 and 132 changes only by 0.01 and between maps 158 and 132 by 0.02, **Tables S23 and S39**). In general, the maps whose distances are below 0.05 can be considered to be highly similar. On the contrary, the differences between the maps whose distances are above 0.1 can be considered to be more important.

In one case (TRPV1 channel), the approximation errors of the density maps are much higher than 10 % and very different among each other (**Table S7**) and the method detects that the most different maps are 146, 156, 161, and 163 (**Fig. 7, Fig. S16C, and Table S35**). These differences and higher final approximation errors for these maps are probably due to a less well

resolved density and/or high level of noise in the original maps (note that the mask does not totally remove the noise outside the useful density and some noise is also present inside the useful part of the map) (**Fig. S14, Fig. S23, and Fig. S24G**).

Detailed analysis of similarities and dissimilarities among maps of 7 target systems

β-Galactosidase (BetaGal):

This data set contains 12 density maps (map indexes: 106, 113, 116, 134, 138, 139, 154, 157, 159, 160, 164, and 167). Maps 138 and 139 can be considered to contain the same information as they were projected onto the same point in the Gaussian-map distance space (**Fig. 1**). This has been verified by overlapping the two maps in Chimera (**Fig. S2A**). The same is valid for maps 159 and 164 (**Fig. S2B**). As being very close to each other in the Gaussian-map distance space, maps 106, 134, and 154 can be considered to be very similar (**Fig. S2C**). Maps 138 and 139 are close to maps 106, 134, and 154 (**Fig. 1**), which means that these two groups of maps are relatively similar. Their display in Chimera shows that maps 138 and 139 are slightly smoother versions of maps 106, 134, 154 (**Fig. S2D**). Following the same reasoning based on the map distance space, map 113 can be considered to be more similar to maps 138 and 139 than to maps 106, 134, 154 (**Fig. 1**) and their display in Chimera indicates that map 113 is a smoother version of maps 138 and 139 (**Fig. S2D**). Furthermore, based on the distance space and visualization in Chimera, maps 116, 157, 159, 160, 164, and 167 can be considered to be more different from 106, 134, 138, 139, and 154 than map 113 is (**Fig. 1 and Fig. S2E**). For instance, one can notice that maps 159 and 167 are locally much smoother than other maps (**Fig. S2E**). Based on these blind analysis results, the most reproduced structural information by different protocols and users is in maps 106, 134, 138, 139, and 154. Similar results have been obtained in non-Gaussian-based distance space, except that map 113 is much closer to maps

138 and 139 in the non-Gaussian-based distance space than in the Gaussian-based distance space (**Fig. S1, Fig. S15A**). The distance between maps 113 and 138 (or 139) is 0.03 in the non-Gaussian-based distance space (**Tables S13**), which means that they can be considered as almost identical based on the results in this distance space. However, the information in these maps is different (as already mentioned, map 113 is a relatively smooth version of maps 138 and 139, **Fig. S2D**), which was better detected in the Gaussian-based distance space (the distance between maps 113 and 138 is larger (0.09) in the Gaussian-based distance space) (**Fig. 1, Tables S11**).

Brome Mosaic Virus (BMV):

This data set contains 7 density maps (map indexes: 102, 110, 136, 137, 140, 142, and 152). Maps 137 and 142 can be considered to be almost the same as they were projected onto almost the same point in the Gaussian-map distance space (**Fig. 2**), which can be verified by overlapping these maps in Chimera (**Fig. S4A**). Maps 136 and 102 were projected in a close neighborhood of maps 137 and 142 as well as map 140 was projected close to map 102 (**Fig. 2**). This indicates that these five maps (102, 136, 137, 140, and 142) are similar and their slight differences can largely be explained by different local smoothing effects (**Figs. S4C and S4D**). Maps 110 and 152 have different density distributions than other maps (**Fig. S4B**) and were projected farther away (**Fig. 2**). Actually, the radius of the capsid seems to be larger in these two maps and one can observe the differences in the densities using the isosurface density levels for which similar capsid sizes are obtained in different maps (**Fig. S4B**). Based on these blind analysis results, the most reproduced structural information by different protocols and users is in maps 102, 136, 137, 140, and 142. These results are similar to those obtained in non-Gaussian-based distance space (**Fig. S3, Fig. S15B**), except that map 102 was in non-Gaussian-based distance space projected closer (found to be more similar) to map 110. A visual

inspection of the maps indicates that map 102 is actually quite different from map 110 (**Fig. S4**), which is indeed better described in the Gaussian-based distance space than in the non-Gaussian-based distance space (the distance between maps 102 and 110 is 0.52 in the Gaussian-based distance space while it is 0.20 in the non-Gaussian-based distance space, **Tables S15 and S17, Fig. S15B**).

Apo-ferritin:

This data set contains 8 density maps (map indexes: 112, 118, 121, 122, 124, 147, 155, and 166). In the Gaussian-map distance space, maps 112, 118, and 166 were projected relatively close to each other while the other maps were projected farther away from these three maps (**Fig. 3**). This means that, in terms of Gaussian-based map approximations, maps 112, 118, and 166 are similar among each other and different from the other maps. Indeed, maps 118 and 166 look like smoothed versions of map 112 (**Fig. S6A and S6B**). Also, though map 121 looks similar to map 112, slight differences can be observed in the conformation of the complex in the two maps using Chimera by overlapping the two maps (**Fig. S6C**) and morph mapping. Furthermore, map 124 seems to contain a slightly different conformation than maps 118 and 166, as observed by overlapping the maps (**Fig. S6D**) and by morph mapping in Chimera. Finally, the remaining maps (122, 147, and 155) are quite noisy and less well resolved (**Fig. S6E**). Based on these blind analysis results, the most reproduced structural information by different protocols and users is in maps 112, 118, 166. These results are similar to those obtained in non-Gaussian-based distance space. More precisely, the differences that exist between the configuration of points in the Gaussian and non-Gaussian-based distance spaces can be considered of little or no importance (on average, the interpoint distances differ between the two spaces by less than 0.05) (**Fig. S5, Fig. S15C, Tables S19 and S21**).

GroEL *in silico*:

This data set contains 9 density maps (map indexes: 104, 120, 132, 143, 153, 158, 165, 168, and 169). Maps 104 and 143 were projected almost onto the same point in the Gaussian-map distance space (**Fig. 4**), which means that they are almost identical (**Fig. S8A**). Also, the projections of maps 132 and 153 in the Gaussian-map distance space are almost the same (**Fig. 4**), meaning that these two maps are also almost the same (**Fig. S8B**). Interestingly enough, maps 132 and 153 were projected close to maps 104 and 143, meaning that they are similar and this can be checked with Chimera (**Figs. S8A, S8B, and S8C**). Maps 120 and 165 were also projected in a close neighborhood of maps 104, 132, 143, and 153 and the distances among these 6 maps are smaller than 0.05, which indicates a high similarity among them (**Fig. 4, Table S23, Figs. S8A-S8E**). They look like containing slightly different conformations (**Figs. S8D and S8E**), but this finding could not have been checked as the information on how the input data has been simulated and whether it contains conformational variability or not was unavailable here. The remaining maps (158, 168, and 169) are different among each other while being less well resolved or noisier than other maps, which explains the result that the projections of these maps are far away from each other and from the projections of other maps (**Fig. 4, Fig. S8F, Table S23**). Based on these blind analysis results, the most reproduced structural information by different protocols and users is in maps 104, 132, 143, and 153, followed by maps 120 and 165. These results are consistent with those obtained in non-Gaussian-based distance space (**Fig. S7, Fig. S15D, Table S25**).

T20S proteasome:

This data set contains 9 density maps (map indexes: 103, 107, 108, 130, 131, 141, 144, 145, and 162). In the Gaussian-map distance space, maps 144 and 145 were projected onto almost the same point and maps 107, 141, and 162 are closest to them (**Fig. 5, Table S27**), meaning that

maps 107, 141, and 162 are slightly different from maps 144 and 145. The differences among these maps are mostly due to different amounts of noise and possible slight differences in conformation in these maps (**Figs. S10A and S10C**). Maps 103, 130, 131, and 108 were projected farther away (found to be different) from other maps (**Fig. 5, Table S27**), which can be explained by the fact that maps 103, 130, 131, and 108 are noisier and locally slightly less well resolved than other maps (**Figs. S10D and S10E**). Finally, maps 130 and 131 are almost the same and were projected onto the same point (**Fig. S10B**). Based on these blind analysis results, the most reproduced structural information by different protocols and users is in maps 107, 141, 144, 145, and 162. These results are similar to those obtained in non-Gaussian-based distance space (**Fig. S9, Fig. S16A, Table S29**). The most important difference between the two spaces is that map 103 (an outlier in the Gaussian-based distance space) was projected even farther from other maps in non-Gaussian-based distance space (e.g., when going from Gaussian-based distance space to non-Gaussian-based distance space, its distance from map 107 increases from 0.14 to 0.20 and its distance from map 141 increases from 0.16 to 0.23). The other interpoint distances differ between the two spaces by less than 0.05 on average, which can be considered of little or no importance.

80S ribosome:

This data set contains 13 density maps (map indexes: 111, 114, 119, 123, 125, 126, 127, 128, 129, 148, 149, 150, and 151). In this case, the majority of maps were projected in the Gaussian-map distance space close to each other (**Fig. 6, Table S31**). This suggests that they are similar, which can be verified with Chimera (**Fig. S12A**). The remaining maps (111, 119, and 129) were projected farther (**Fig. 6, Table S31**). This suggests that maps 111, 119, and 129 are different from other maps, which can be verified with Chimera (**Fig. S12B**). Indeed, the densities are locally less well resolved in these maps than in other maps, as shown using two different

density isosurface levels in Chimera (**Figs. S12C and S12D**). These results are consistent with those obtained in non-Gaussian-based distance space (**Fig. S11, Fig. S16B, Table S33**). Based on these blind analysis results, the same structural information was reproduced by different protocols and users in the majority of maps (more precisely, in maps 114, 123, 125, 126, 127, 128, 148, 149, 150, and 151).

TRPV1 channel:

This data set contains 8 density maps (map indexes: 101, 115, 133, 135, 146, 156, 161, and 163). In the Gaussian-map distance space, maps 133 and 135 were projected onto almost the same point (**Fig. 7**). This means that they are very similar, which was verified with Chimera (**Fig. S14A**). The remaining maps were projected more or less far from these two maps and closest to them are maps 101 and 115 (**Fig. 7, Table S35**). This means that maps 133 and 135 are more or less different from the remaining maps and most similar to them are maps 101 and 115, which can be verified with Chimera (**Fig. S14**). Indeed, Chimera indicates more noise and less well resolved densities in the remaining maps (**Figs. S14A and S14B**). Based on these blind analysis results, the most reproduced structural information by different protocols and users is in maps 101, 115, 133, and 135. These results are similar to those obtained in non-Gaussian-based distance space (**Fig. S13, Fig. S16C, Table S37**). The most important difference between the two spaces is that maps 146 and 156 (two outliers in the Gaussian-based distance space, **Table S35**) were projected even farther from other maps in the non-Gaussian-based distance space (**Table S37**).

DISCUSSION AND CONCLUSION

We presented a methodology that allows identifying similar (or dissimilar) density maps from a set of density maps of the same macromolecular complex. This methodology is based on

projecting Gaussian-based approximations of the density maps onto a low-dimensional space of their distances in which the closest density maps correspond to the most similar maps and the most distant maps correspond to the most dissimilar ones. We showed that the methodology based on calculating distances (correlation coefficients) among Gaussian-based approximated maps is more appropriate for the task of identifying similar and dissimilar density maps than the methodology based on distances among original maps. The Gaussian-based map representations were obtained with the method proposed in (Jonic and Sorzano, 2016a). Several different applications of this method have already been shown, including normal-mode-based deformation modeling for continuous conformational variability analysis and EM map denoising (Jin et al., 2014; Jonic and Sorzano, 2016b; Jonic et al., 2016; Sanchez Sorzano et al., 2016). Though this method can be used to fully denoise the maps (Jonic et al., 2016), it should be noted that this method was here used with a different goal. Its goal was to represent the different maps using the same “building blocks” (Gaussian functions that are the basis functions in the function approximation framework used in this method) i.e. to “uniformize” the representation of different maps of the same target system in order to get more comparable maps. Therefore, the use of Gaussian functions facilitated the comparison of maps not only because it reduced noise in the maps but also because it “uniformized” the representation of the map density.

We also presented the results obtained with this methodology in analyzing density maps of 7 macromolecular complexes reconstructed from experimental and synthetic cryo-EM images by the participants of the 2015/2016 Map Challenge. This methodology was used in the Blind Assessment Phase of the Map Challenge to identify most reproduced structural information by different image analysis protocols and software and the obtained results are supported by the results of analyzing the density maps in Chimera. We identified the maps that are either

globally similar or globally dissimilar. A potential local analysis (focusing on regions of the maps that look different to illustrate the performance of the analysis on these regions) was out of scope of this article. Furthermore, though an atomic model is available for each of the seven data sets, this model was not used here because the idea was to devise a method that would be complementary to methods that use atomic models for map evaluation. Also, a further map interpretation to address the question whether the different maps would lead to different atomic models was out of scope of this work.

In the Blind Assessment Phase, beside density maps, additional information was available (provided by the participants of the Challenge Phase), such as on the movie frames used to determine the orientation of images and the frames used for the final map reconstruction; whether the particle coordinates provided with images were used or new particle picking was performed; the initial number of particle images; the number of particles at the start of refinement and the number of particles included in the final density map; whether 2D and 3D classifications, local drift correction, and exposure weighting were performed; the resolution of the final map and the method used for its determination; whether other methods were used to validate the map and which ones; the voxel size; the temperature factor and other filtering applied on the final map; and whether the mask was submitted. Also, other additional information was provided by the Map Challenge organizers, such as the FSC curves calculated using the FSC server (<http://pdbe.org/fsc>) between the submitted half-maps (without using masks and, if a mask was provided by the participant, using this mask) and the inverse spatial frequencies at which these FSC curves crossed 0.5 and 0.143. From all this information, in the experiments presented here, we have only used the information on the voxel size.

After the end of the Blind Assessment Phase, the organizers have provided more details about the image analysis protocols used in the Challenge Phase, such as details on movie frame alignment; CTF estimation and correction; particles selection; initial model (from a data base or from images); 2D and 3D classification protocols; local drift correction; exposure weighting; map refinement software; and map validation methods other than FSC-based. For this article, we have analyzed this additional information to see whether the maps identified as similar using the proposed methodology were obtained using similar image analysis protocols. This analysis has showed that, for the same target complex, different protocols often resulted in density maps that were identified, using our methodology, as globally containing similar structural information. Over different target complexes, the same researcher has usually used the same image analysis protocol within the same software package (or a combination of software packages) but with a few different parameters (e.g., the start and stop frames and the number of particles at different processing steps). Such basically identical protocols have resulted in density maps that were, depending on the target complex, identified as similar to or different from the maps from other protocols.

The map differences identified by our methodology are sometimes difficult to interpret in terms of differences in image analysis protocols. A reason is that some information about how the density maps were reconstructed is still missing. For instance, the participants have not always made clear whether symmetries had been imposed or not during or at the end of refinement in the cases of symmetric complexes (6 out of 7 target complexes), whereas our methodology identifies non-symmetrized maps as different from symmetrized maps. Also, as some participants have pointed out in the provided information, the CTF amplitude correction is often an integral part of the refinement/reconstruction protocol (which is not always clearly indicated) and it commonly involves Wiener filtering. This means that the density maps deposited as

“unfiltered” (used by our methodology) may still be considered as filtered, with the amount of filtering that depends on the protocol used.

With the provided density maps, a full and fair comparison of different protocols over different target complexes is unfortunately impossible. The main reason for this is that the same image analysis protocol was not systematically used with all density maps. Regarding the software packages, only Relion was used with each target complex. More precisely, different researches have used Relion at different steps of their protocols and sometimes in combination with other software packages. When Relion was used for refinement, the resulting density maps of the same target complex are similar among each other, even when different frame alignment and particle sorting methods were used, which can be explained by the fact that the initial models used in all these cases were low-pass-filtered published maps. According to our methodology, depending on the target complex, the density maps resulting from other protocols are more or less similar to (or different from) the Relion-based maps. However, other software packages and other protocols were less often used, among which those refining *de novo* initial models (calculated by the participant from selected particle images). Therefore, it is difficult to make fair statistics and comparisons of different protocols using the provided data.

However, our methodology describes well similarities and differences among density maps of the same target complex obtained using different software packages and protocols, as showed by analyzing these maps in Chimera. This methodology could also be used to analyze density maps of the same complex obtained from different sets of images using the same image analysis protocol. Averaging of density maps identified, using this methodology, as most similar among each other could be expected to yield a density map with a better signal-to-noise ratio than the one of individual density maps. The methodology proposed here could also be

used to obtain a “consensus” density map using several different image analysis protocols, within the same software package or different packages. The “consensus” map could be the average of maps identified as most similar among each other or one of these maps, whichever of these two options results in the highest signal-to-noise-ratio density map.

SUPPORTING MATERIAL

The Supporting Material contains a brief description of the method for Gaussian-based density map approximation, 25 supplementary figures (**Figs. S1-S25**), and 39 supplementary tables (**Tables S1-S39**).

ACKNOWLEDGMENTS

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

REFERENCES

- Bai, X.C., Yan, C., Yang, G., Lu, P., Ma, D., Sun, L., Zhou, R., Scheres, S.H., Shi, Y., 2015. An atomic structure of human gamma-secretase. *Nature* 525, 212-217.
- Baker, M.L., Zhang, J., Ludtke, S.J., Chiu, W., 2010. Cryo-EM of macromolecular assemblies at near-atomic resolution. *Nat Protoc* 5, 1697-1708.
- Banerjee, S., Bartesaghi, A., Merk, A., Rao, P., Bulfer, S.L., Yan, Y., Green, N., Mroczkowski, B., Neitz, R.J., Wipf, P., Falconieri, V., Deshaies, R.J., Milne, J.L., Huryn, D., Arkin, M., Subramaniam, S., 2016. 2.3 Å resolution cryo-EM structure of human p97 and mechanism of allosteric inhibition. *Science* 351, 871-875.
- Bartesaghi, A., Matthies, D., Banerjee, S., Merk, A., Subramaniam, S., 2014. Structure of beta-galactosidase at 3.2-Å resolution obtained by cryo-electron microscopy. *Proc Natl Acad Sci U S A* 111, 11709-11714.

- Bartesaghi, A., Merk, A., Banerjee, S., Matthies, D., Wu, X., Milne, J.L., Subramaniam, S., 2015. 2.2 Å resolution cryo-EM structure of beta-galactosidase in complex with a cell-permeant inhibitor. *Science* 348, 1147-1151.
- Campbell, M.G., Veessler, D., Cheng, A., Potter, C.S., Carragher, B., 2015. 2.8 Å resolution reconstruction of the *Thermoplasma acidophilum* 20S proteasome using cryo-electron microscopy. *Elife* 4.
- Cox, T.F., Cox, M.A.A., 2001. *Multidimensional Scaling*. 2nd ed. Chapman & Hall/CRC, Boca Raton, FL.
- Harauz, G., van Heel, M., 1986. Exact filters for general geometry three dimensional reconstruction. *Optik* 73, 146-156.
- Hospenthal, M.K., Redzej, A., Dodson, K., Ukleja, M., Frenz, B., Rodrigues, C., Hultgren, S.J., DiMaio, F., Egelman, E.H., Waksman, G., 2016. Structure of a Chaperone-Usher Pilus Reveals the Molecular Basis of Rod Uncoiling. *Cell* 164, 269-278.
- Jin, Q., Sorzano, C.O., de la Rosa-Trevin, J.M., Bilbao-Castro, J.R., Nunez-Ramirez, R., Llorca, O., Tama, F., Jonic, S., 2014. Iterative elastic 3D-to-2D alignment method using normal modes for studying structural dynamics of large macromolecular complexes. *Structure* 22, 496-506.
- Jonic, S., Sorzano, C.O.S., 2016a. Coarse-graining of volumes for modeling of structure and dynamics in electron microscopy: Algorithm to automatically control accuracy of approximation. *IEEE Journal of Selected Topics in Signal Processing* 10, 161 - 173.
- Jonic, S., Sorzano, C.O., 2016b. Versatility of Approximating Single-Particle Electron Microscopy Density Maps Using Pseudoatoms and Approximation-Accuracy Control. *Biomed Res Int* 2016, 7060348.
- Jonic, S., Vargas, J., Melero, R., Gomez-Blanco, J., Carazo, J.M., Sorzano, C.O., 2016. Denoising of high-resolution single-particle electron-microscopy density maps by their approximation using three-dimensional Gaussian functions. *J Struct Biol* 194, 423-433.
- Khatter, H., Myasnikov, A.G., Natchiar, S.K., Klaholz, B.P., 2015. Structure of the human 80S ribosome. *Nature* 520, 640-645.

- Liao, M., Cao, E., Julius, D., Cheng, Y., 2013. Structure of the TRPV1 ion channel determined by electron cryo-microscopy. *Nature* 504, 107-112.
- Russo, C.J., Passmore, L.A., 2014. Electron microscopy: Ultrastable gold substrates for electron cryomicroscopy. *Science* 346, 1377-1380.
- Sanchez Sorzano, C.O., Alvarez-Cabrera, A.L., Kazemi, M., Carazo, J.M., Jonic, S., 2016. StructMap: Elastic Distance Analysis of Electron Microscopy Maps for Studying Conformational Changes. *Biophys J* 110, 1753-1765.
- Vulovic, M., Ravelli, R.B., van Vliet, L.J., Koster, A.J., Lazic, I., Lucken, U., Rullgard, H., Oktem, O., Rieger, B., 2013. Image formation modeling in cryo-electron microscopy. *J Struct Biol* 183, 19-32.
- Wang, R.Y., Kudryashev, M., Li, X., Egelman, E.H., Basler, M., Cheng, Y., Baker, D., DiMaio, F., 2015. De novo protein structure determination from near-atomic-resolution cryo-EM maps. *Nat Methods* 12, 335-338.
- Wang, Z., Hryc, C.F., Bammes, B., Afonine, P.V., Jakana, J., Chen, D.H., Liu, X., Baker, M.L., Kao, C., Ludtke, S.J., Schmid, M.F., Adams, P.D., Chiu, W., 2014. An atomic model of brome mosaic virus using direct electron detection and real-space optimization. *Nat Commun* 5, 4808.
- Wong, W., Bai, X.C., Brown, A., Fernandez, I.S., Hanssen, E., Condrón, M., Tan, Y.H., Baum, J., Scheres, S.H., 2014. Cryo-EM structure of the Plasmodium falciparum 80S ribosome bound to the anti-protozoan drug emetine. *Elife* 3.
- Yu, X., Jin, L., Zhou, Z.H., 2008. 3.88 Å structure of cytoplasmic polyhedrosis virus by cryo-electron microscopy. *Nature* 453, 415-419.
- Zhang, X., Settembre, E., Xu, C., Dormitzer, P.R., Bellamy, R., Harrison, S.C., Grigorieff, N., 2008. Near-atomic resolution using electron cryomicroscopy and single-particle reconstruction. *Proc Natl Acad Sci U S A* 105, 1867-1872.

FIGURE LEGENDS

Figure 1: Projection of Gaussian-based approximations of β -Galactosidase maps onto a 3D space of their distances. In the distance space, each map is represented by a point and the map index.

Figure 2: Projection of Gaussian-based approximations of Brome Mosaic Virus maps onto a 3D space of their distances. In the distance space, each map is represented by a point and the map index.

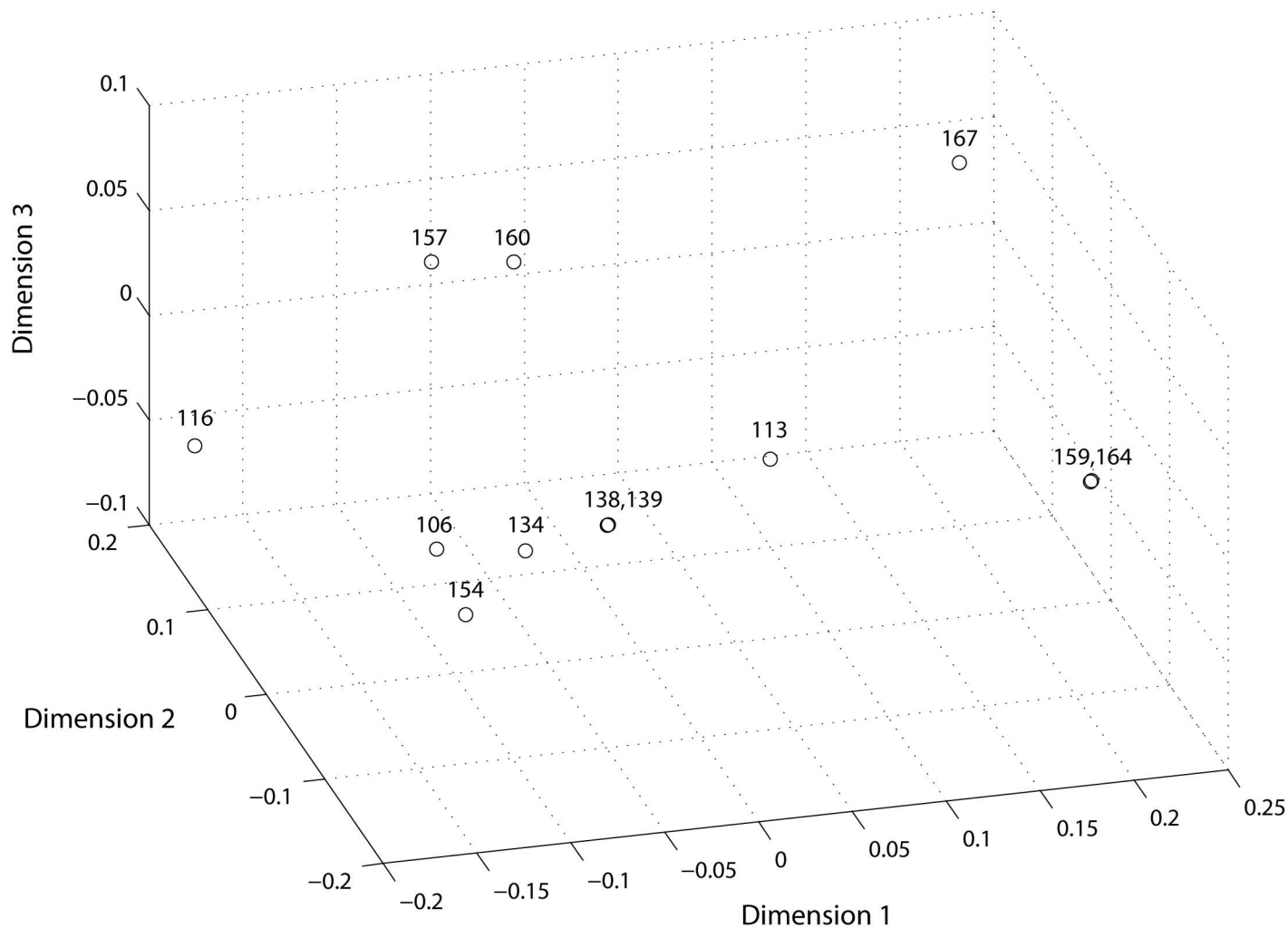
Figure 3: Projection of Gaussian-based approximations of Apo-ferritin maps onto a 3D space of their distances. In the distance space, each map is represented by a point and the map index.

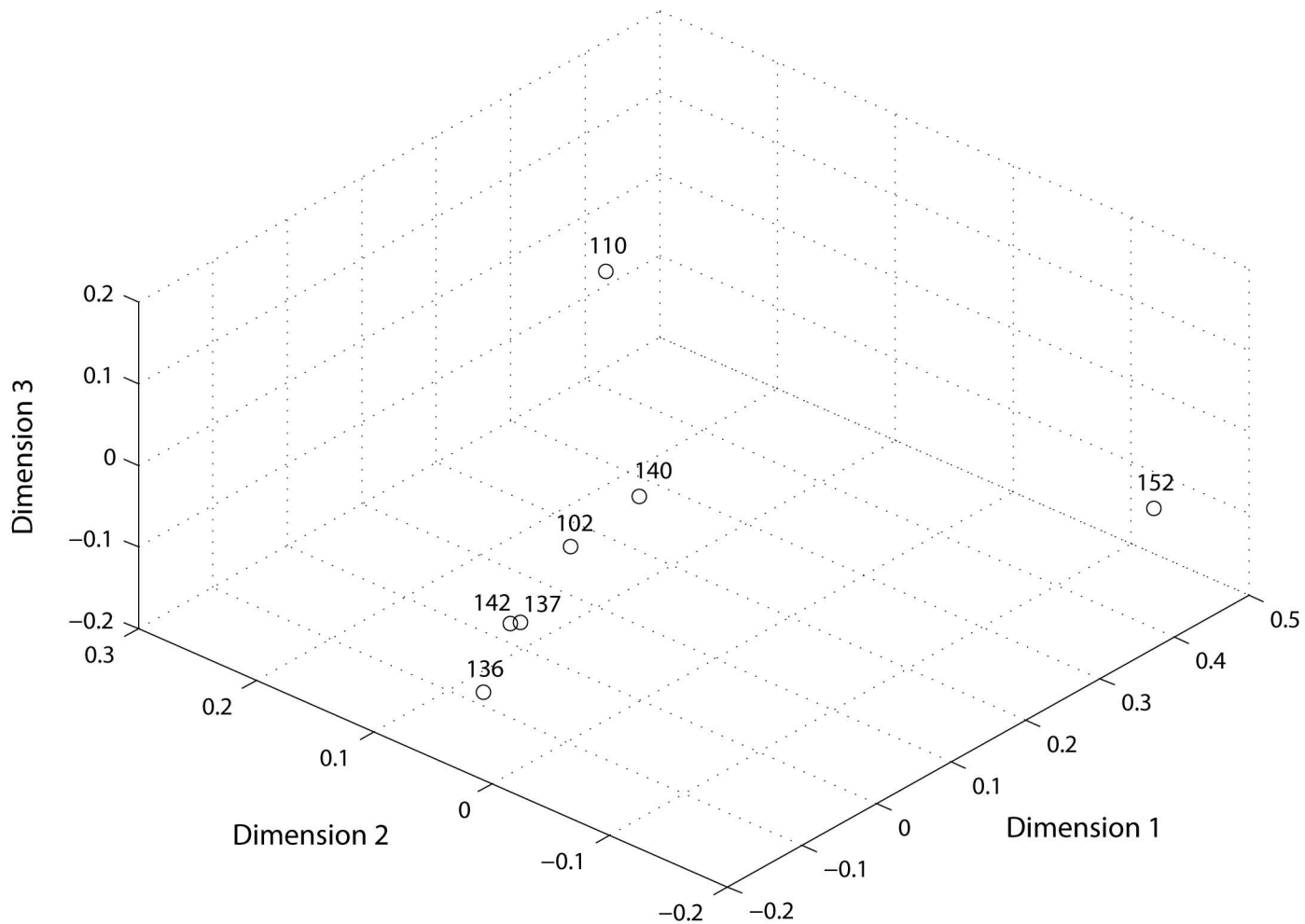
Figure 4: Projection of Gaussian-based approximations of GroEL *in silico* maps onto a 3D space of their distances. In the distance space, each map is represented by a point and the map index.

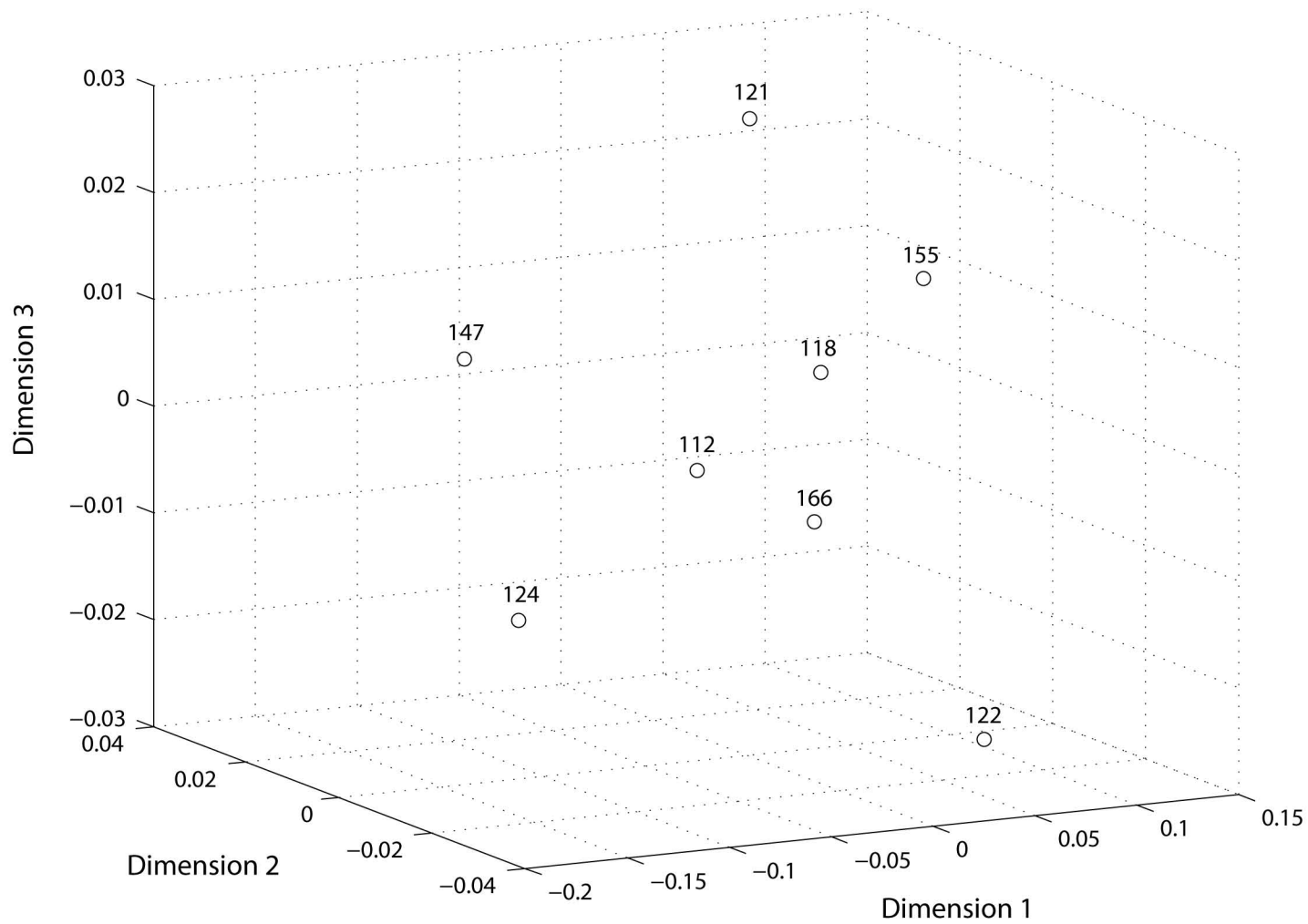
Figure 5: Projection of Gaussian-based approximations of T20S proteasome maps onto a 3D space of their distances. In the distance space, each map is represented by a point and the map index.

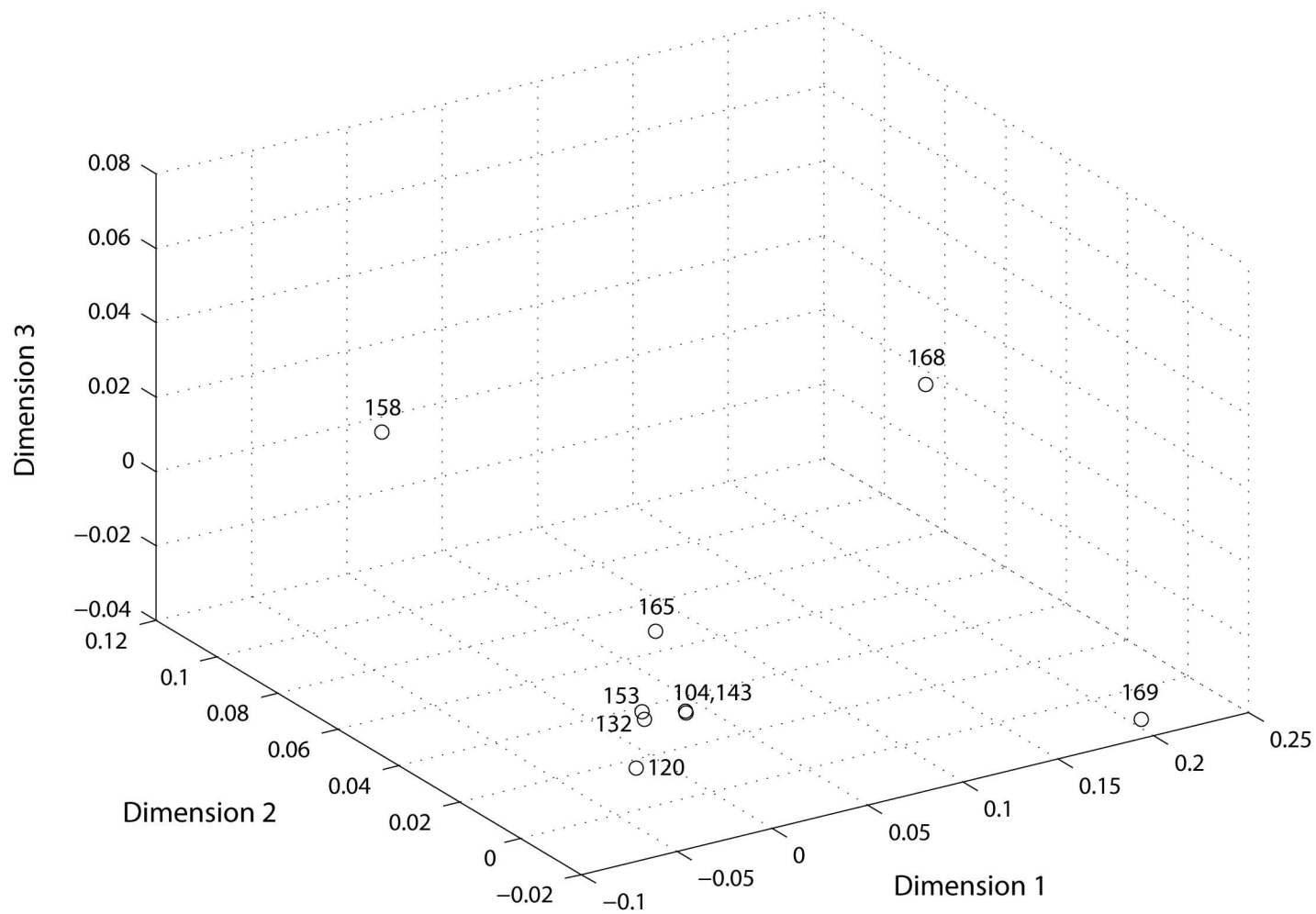
Figure 6: Projection of Gaussian-based approximations of 80S ribosome maps onto a 3D space of their distances. In the distance space, each map is represented by a point and the map index.

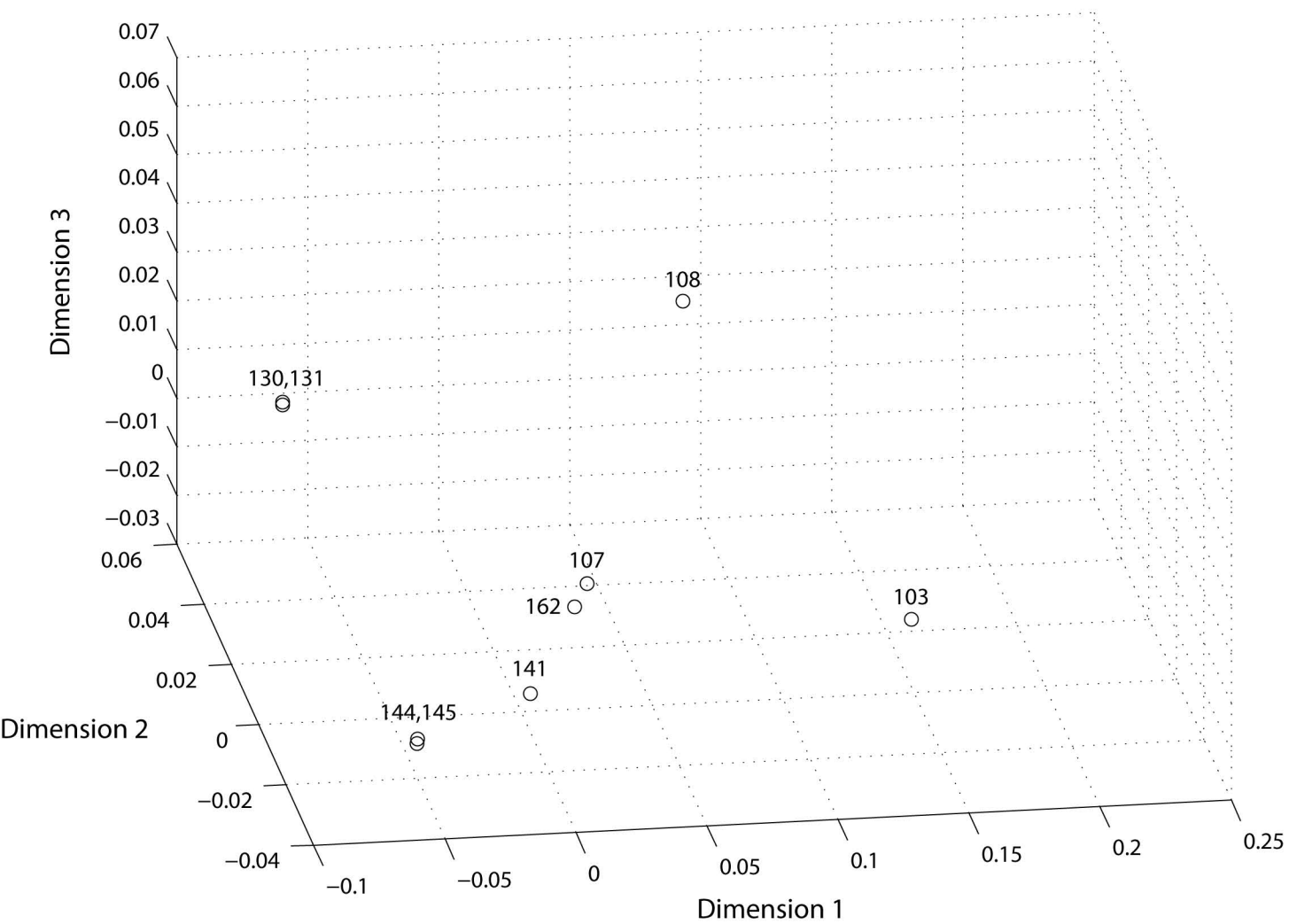
Figure 7: Projection of Gaussian-based approximations of TRPV1 channel maps onto a 3D space of their distances. In the distance space, each map is represented by a point and the map index.

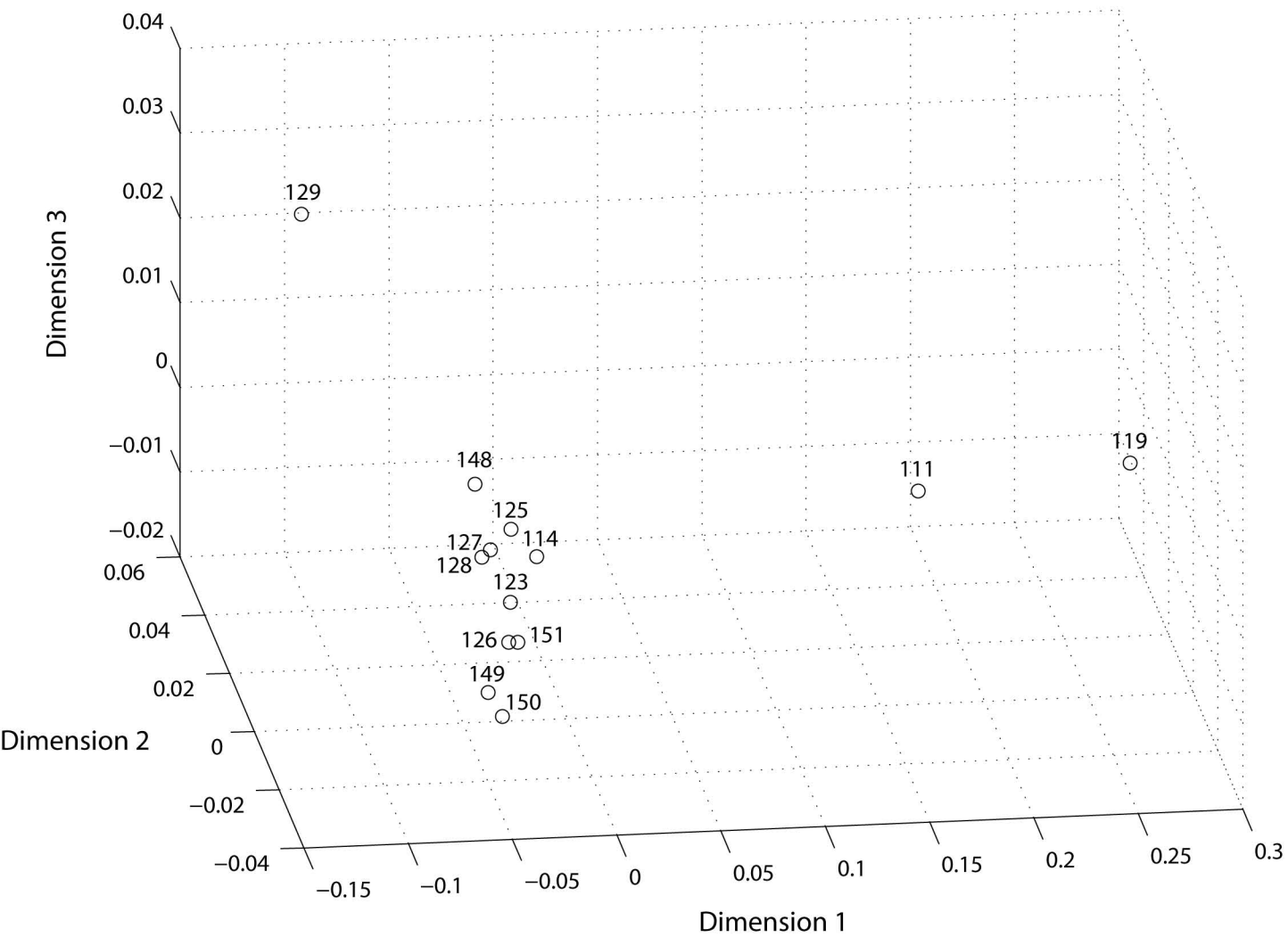


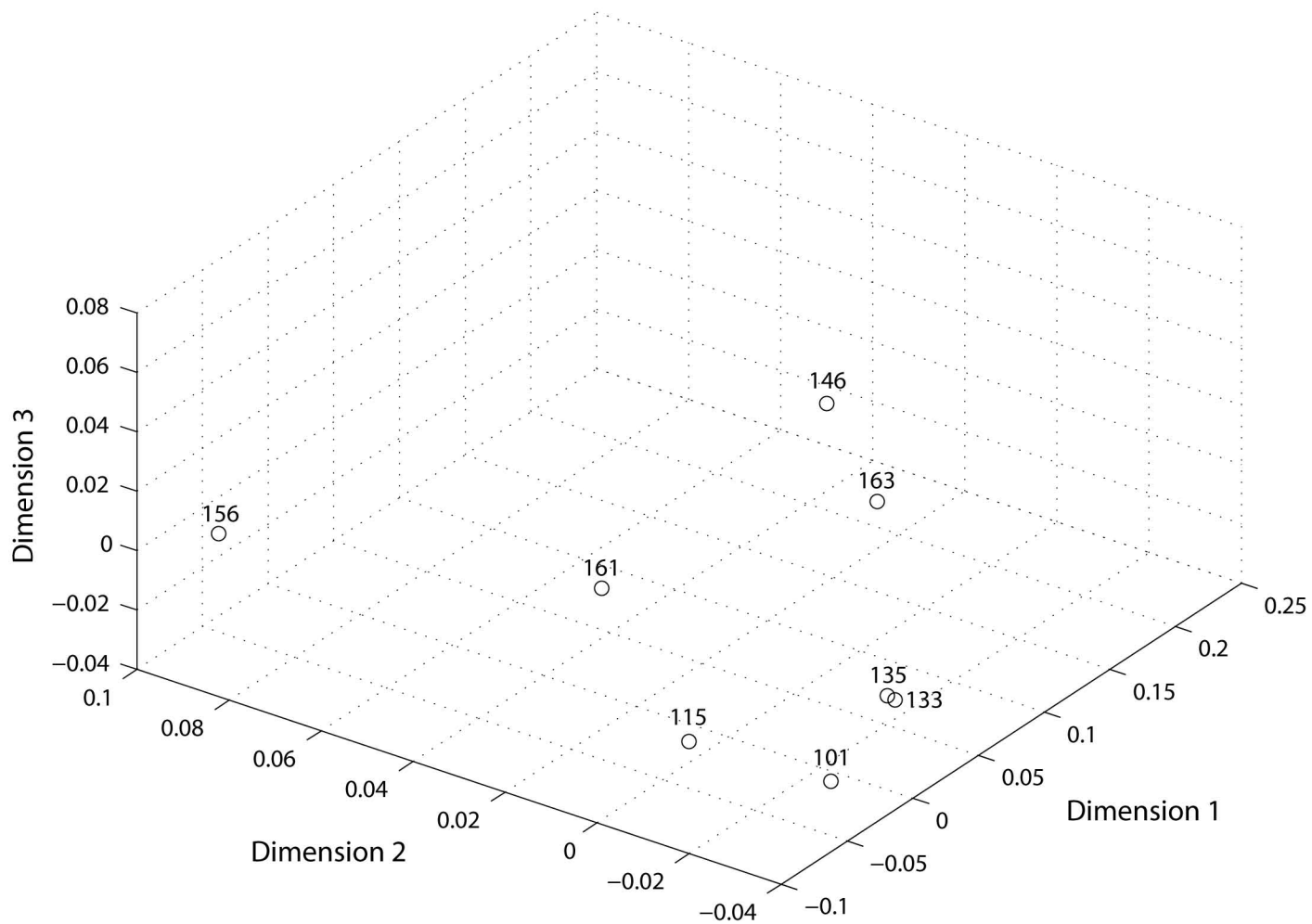












| | Reference map to align and resample maps | Size of reference map [voxel] | Size of each map after final resampling [voxel] | Voxel size after final resampling [Å] | Reference-map density threshold to create mask | Standard deviation of Gaussian functions [voxel] | Target approximation error [%] |
|---------------------------|--|-------------------------------|---|---------------------------------------|--|--|--------------------------------|
| β-Galactosidase | 106 | 196 ³ | 196 ³ | 1.275 | 0.0161 | 0.6 | 5 |
| Brome Mosaic Virus | 102 | 420 ³ | 210 ³ | 1.98 | 0.00682 | 0.75 | 10 |
| Apo-ferritin | 122 | 132 ³ | 132 ³ | 1.346 | 0.0313 | 1.5 | 10 |
| GroEL <i>in silico</i> | 104 | 200 ³ | 200 ³ | 1.42 | 0.0132 | 1.5 | 10 |
| T20S proteasome | 107 | 240 ³ | 240 ³ | 1.315 | 0.00675 | 0.6 | 10 |
| 80S ribosome | 114 | 380 ³ | 190 ³ | 2.68 | 0.0684 | 1.0 | 5 |
| TRPV1 channel | 101 | 256 ³ | 256 ³ | 1.2 | 0.00959 | 1.0 | 5 |

Table 1: Parameters used to compute Gaussian-based approximations of cryo-EM density maps of 7 target biological systems.

SUPPORTING MATERIAL

A methodology using Gaussian-based density map approximation to assess sets of cryo-electron microscopy density maps

Slavica Jonić

Sorbonne Université, UMR CNRS 7590, Muséum National d'Histoire Naturelle, IRD, Institut de Minéralogie, de Physique des Matériaux et de Cosmochimie, IMPMC, 75005 Paris, France.

All correspondence in relation to this article should be sent to:

Slavica Jonic, IMPMC-UMR CNRS 7590, Sorbonne Université, CC 115, 4 Place Jussieu, 75005 Paris, France; Phone: +33 1 44 27 72 05; Fax: +33 1 44 27 37 85; E-mail: Slavica.Jonic@upmc.fr

CONTENTS:

This Supporting Material contains a brief description on the method for Gaussian-based density map approximation (Jin et al., 2014; Jonic and Sorzano, 2016a; Jonic et al., 2016; Sanchez Sorzano et al., 2016), 25 supplementary figures (**Figs. S1-S25**), and 39 supplementary tables (**Tables S1-S39**).

BACKGROUND ON THE METHOD FOR GAUSSIAN-BASED MAP APPROXIMATION

A function $f(\mathbf{r})$ ($\mathbf{r} \in \mathbb{R}^3$) can be approximated using Gaussian radial basis functions (RBFs) by $\hat{f}_N(\mathbf{r}) = \sum_{i=1}^N \omega_i K_\sigma(\|\mathbf{r} - \mathbf{r}_i\|)$, where $K_\sigma(r)$ is the RBF kernel that is a Gaussian function with the standard deviation σ and the amplitude of 1, N is the number of Gaussian functions, \mathbf{r}_i is the vector of the center coordinates of the i th Gaussian function, $\|\mathbf{r} - \mathbf{r}_i\|$ is the Euclidean distance between the vectors \mathbf{r} and \mathbf{r}_i , and $\omega_i > 0$ is the weight (contribution) of the i th Gaussian function. Given an EM density map $f(\mathbf{r})$ ($\mathbf{r} \in \mathbb{R}^3$), a Gaussian-function standard deviation σ , and a target approximation error ε , our approach determines the number of Gaussian functions N , their positions \mathbf{r}_i , and weights ω_i such that the approximation error e_N satisfies $e_N = \frac{1}{V} \sum_{j=1}^V \frac{|f(\mathbf{r}_j) - \hat{f}_N(\mathbf{r}_j)|}{\Delta f} < \varepsilon$. Here, Δf is the effective range of values in the EM map, \mathbf{r}_j is the voxel location at which the given EM map is compared with its approximation, and V is the total number of evaluated voxels (the evaluation can be done in a region of interest defined by a mask). To avoid getting trapped into local minima of the error e_N , new Gaussian functions are added progressively in regions with large errors, from a given initial number of Gaussian functions (referred to as the initial seeds parameter) using a given speed of adding

the Gaussian functions (referred to as the grow seeds parameter), and weights and positions of the current number of Gaussian functions are determined by a gradient descent minimization of e_N while respecting a given minimum distance between the Gaussian functions d_{min} (the Gaussian functions will not be placed closer than the distance d_{min}). It should be noted that Gaussian function positions \mathbf{r}_i do not necessarily coincide with voxel positions \mathbf{r}_j because the Gaussian function positions vary continuously within the EM density map. Also, it should be noted that σ can be expressed in angstroms, but we usually express it in voxels.

Different coarse-grain EM map representations can be obtained with this approach by varying σ and ε . Smaller values of σ and ε result in larger numbers of Gaussian functions and vice versa. Their values (usual range: $\sigma = 0.5 - 2.5$ voxel and $\varepsilon = 1 - 15\%$) should be chosen to suit the target application of the method, as explained in (Jonic and Sorzano, 2016a; Jonic and Sorzano, 2016b). Small enough values of the minimum distance between Gaussian functions d_{min} , the initial seeds parameter, and the grow seeds parameter will only affect the speed of convergence of the algorithm but not the maximum achievable accuracy of the density map approximation (Jonic and Sorzano, 2016a). These parameters are usually kept to their default values, which are 0.001 voxel, 300, and 30 %, for d_{min} , the initial seeds parameter, and the grow seeds parameter, respectively (Jonic and Sorzano, 2016a).

SUPPLEMENTARY FIGURE LEGENDS:

Figure S1: Projection of β -Galactosidase maps with no Gaussian-based approximation onto a 3D space of distances among maps. In the distance space, each map is represented by a point and the map index. See also **Figure 1** in the main text and **Figure S2**.

Figure S2: Density maps of β -Galactosidase with no Gaussian-based approximation displayed in Chimera using comparable isosurface density thresholds. The maps are indicated by their indexes. In panels, G stands for grey, Y for yellow, C for cyan, B for blue, M for magenta, T for tan, and P for pink. See also **Figure 1** in the main text and **Figure S1**.

Figure S3: Projection of Brome Mosaic Virus maps with no Gaussian-based approximation onto a 3D space of distances among maps. In the distance space, each map is represented by a point and the map index. See also **Figure 2** in the main text and **Figure S4**.

Figure S4: Density maps of Brome Mosaic Virus with no Gaussian-based approximation displayed in Chimera using comparable isosurface density thresholds. The maps are indicated by their indexes. In panels A and D, G stands for grey, Y for yellow, C for cyan, M for magenta, and B for blue. See also **Figure 2** in the main text and **Figure S3**.

Figure S5: Projection of Apo-ferritin maps with no Gaussian-based approximation onto a 3D space of distances among maps. In the distance space, each map is represented by a point and the map index. See also **Figure 3** in the main text and **Figure S6**.

Figure S6: Density maps of Apo-ferritin with no Gaussian-based approximation displayed in Chimera using comparable isosurface density thresholds. The maps are indicated by their indexes. In panels B-D, G stands for grey (panels B-C) or green (panel D), C for cyan, Y for yellow, and M for magenta. See also **Figure 3** in the main text and **Figure S5**.

Figure S7: Projection of GroEL maps with no Gaussian-based approximation onto a 3D space of distances among maps. In the distance space, each map is represented by a point and the map index. See also **Figure 4** in the main text and **Figure S8**.

Figure S8: Density maps of GroEL with no Gaussian-based approximation displayed in Chimera using comparable isosurface density thresholds. The maps are indicated by their indexes. In panels A-F, C stands for cyan, B for blue, Y for yellow, G for grey (panels B-C and E) or green (panel F), M for magenta, P for pink, and T for tan. See also **Figure 4** in the main text and **Figure S7**.

Figure S9: Projection of T20S proteasome maps with no Gaussian-based approximation onto a 3D space of distances among maps. In the distance space, each map is represented by a point and the map index. See also **Figure 5** in the main text and **Figure S10**.

Figure S10: Density maps of T20S proteasome with no Gaussian-based approximation displayed in Chimera using comparable isosurface density thresholds. The maps are indicated by their indexes. In panels A-E, G stands for grey (panel A left and panels C-D) or green (panel A right and panel E), Y for yellow, M for magenta, P for pink, C for cyan, B for blue, and T for tan. See also **Figure 5** in the main text and **Figure S9**.

Figure S11: Projection of 80S ribosome maps with no Gaussian-based approximation onto a 3D space of distances among maps. In the distance space, each map is represented by a point and the map index. See also **Figure 6** in the main text and **Figure S12**.

Figure S12: Density maps of 80S ribosome with no Gaussian-based approximation displayed in Chimera, using comparable isosurface density thresholds in panels A-C and a different isosurface density threshold in panel D. The maps are indicated by their indexes. In panels A-B, Y stands for

yellow, M for magenta, C for cyan, T for tan, B for blue, and G for green. See also **Figure 6** in the main text and **Figure S11**.

Figure S13: Projection of TRPV1 channel maps with no Gaussian-based approximation onto a 3D space of distances among maps. In the distance space, each map is represented by a point and the map index. See also **Figure 7** in the main text and **Figure S14**.

Figure S14: Density maps of TRPV1 channel with no Gaussian-based approximation displayed in Chimera using comparable isosurface density thresholds. The maps are indicated by their indexes. In panel A, G stands for grey, Y for yellow, C for cyan, B for blue, M for magenta, P for pink, and R for red. See also **Figure 7** in the main text and **Figure S13**.

Figure S15: Projection of β -Galactosidase (A), Brome Mosaic Virus (B), Apo-ferritin (C), and GroEL (D) density maps with (left) and without (right) Gaussian-based approximation onto a 3D space of distances among maps shown using the same axes limits for all 7 target systems. In the distance space, each map is represented by a point and the map index.

Figure S16: Projection of T20S proteasome (A), 80S ribosome (B), and TRPV1 channel (C) density maps with (left) and without (right) Gaussian-based approximation onto a 3D space of distances among maps shown using the same axes limits for all 7 target systems. In the distance space, each map is represented by a point and the map index.

Figure S17: Density maps of β -Galactosidase (gray, transparent) superposed onto their respective Gaussian-based approximations (yellow, solid). The map indexes are provided near the maps.

Figure S18: Density maps of Brome Mosaic Virus (gray, solid) superposed onto their respective Gaussian-based approximations (yellow, solid). The map indexes are provided near the maps.

Figure S19: Density maps of Apo-ferritin (gray, transparent) superposed onto their respective Gaussian-based approximations (yellow, solid). The map indexes are provided near the maps.

Figure S20: Density maps of GroEL (gray, transparent) superposed onto their respective Gaussian-based approximations (yellow, solid). The map indexes are provided near the maps.

Figure S21: Density maps of T20S proteasome (gray, transparent) superposed onto their respective Gaussian-based approximations (yellow, solid). The map indexes are provided near the maps.

Figure S22: Density maps of 80S ribosome (gray, transparent) superposed onto their respective Gaussian-based approximations (yellow, solid). The map indexes are provided near the maps.

Figure S23: Density maps of TRPV1 channel (gray, transparent) superposed onto their respective Gaussian-based approximations (yellow, solid). The map indexes are provided near the maps.

Figure S24: Masked β -Galactosidase (A), Brome Mosaic Virus (B), Apo-ferritin (C), GroEL (D), T20S proteasome (E), 80S ribosome (F), and TRPV1 channel (G) density maps visualized using isosurfaces corresponding to very low density values to show the contours of the masks. The same mask was used for all maps of the same target system. In this figure, only one masked map is shown for each of the 7 target systems and its index (selected arbitrarily) is shown below the map.

Figure S25: Projection of GroEL density maps with Gaussian-based approximation onto a 3D space of distances among maps (shown using the same axes limits for all 7 target systems) for the following two sets of values of the Gaussian-function standard deviation and the target approximation error, respectively: 1) 1.5 voxel and 10 % (A); and 2) 1 voxel and 5 % (B). In the distance space, each map is represented by a point and the map index. Panel A in this figure is the same as panel D in **Figure S15**. See also **Table S8**.

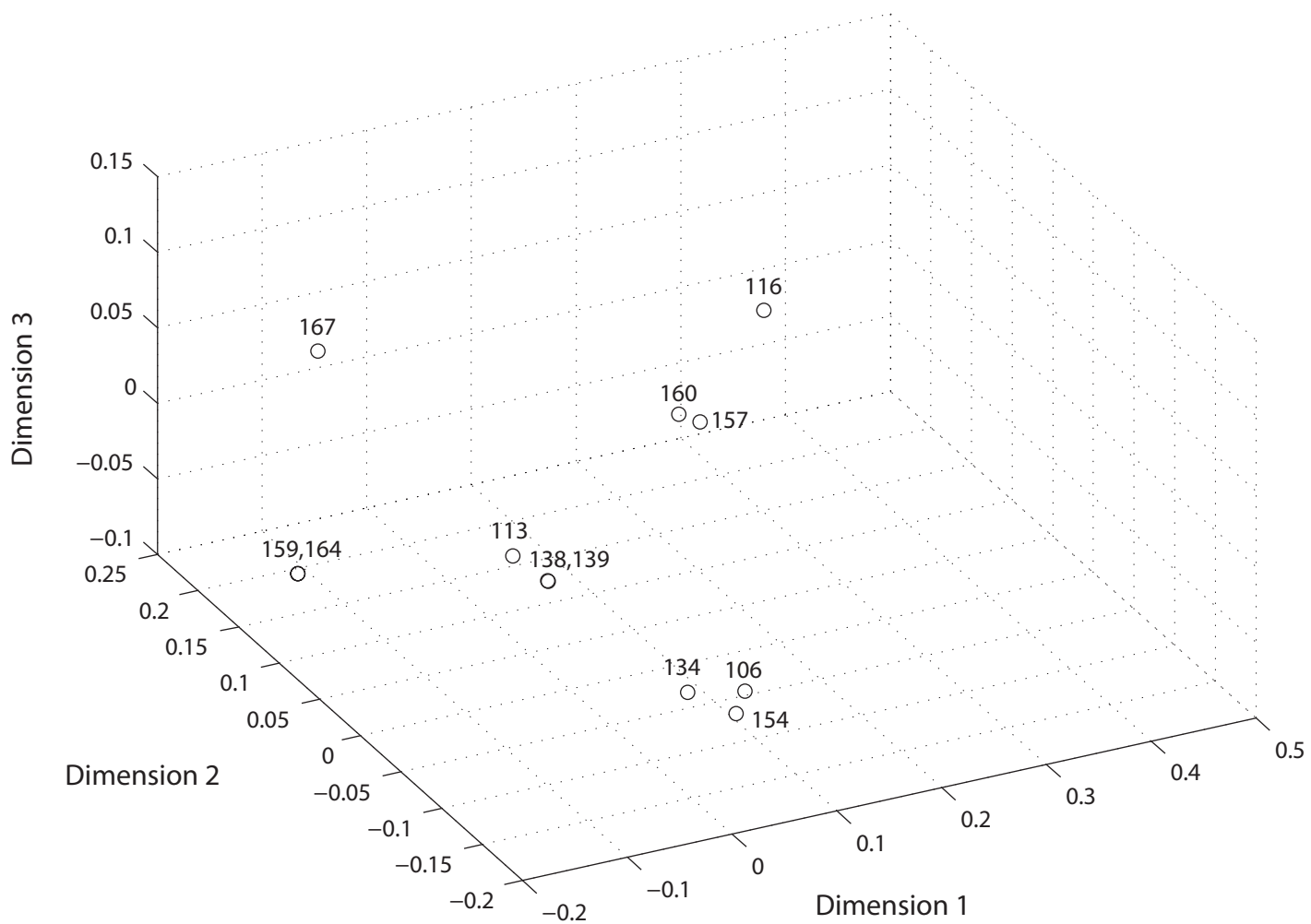


Figure S1

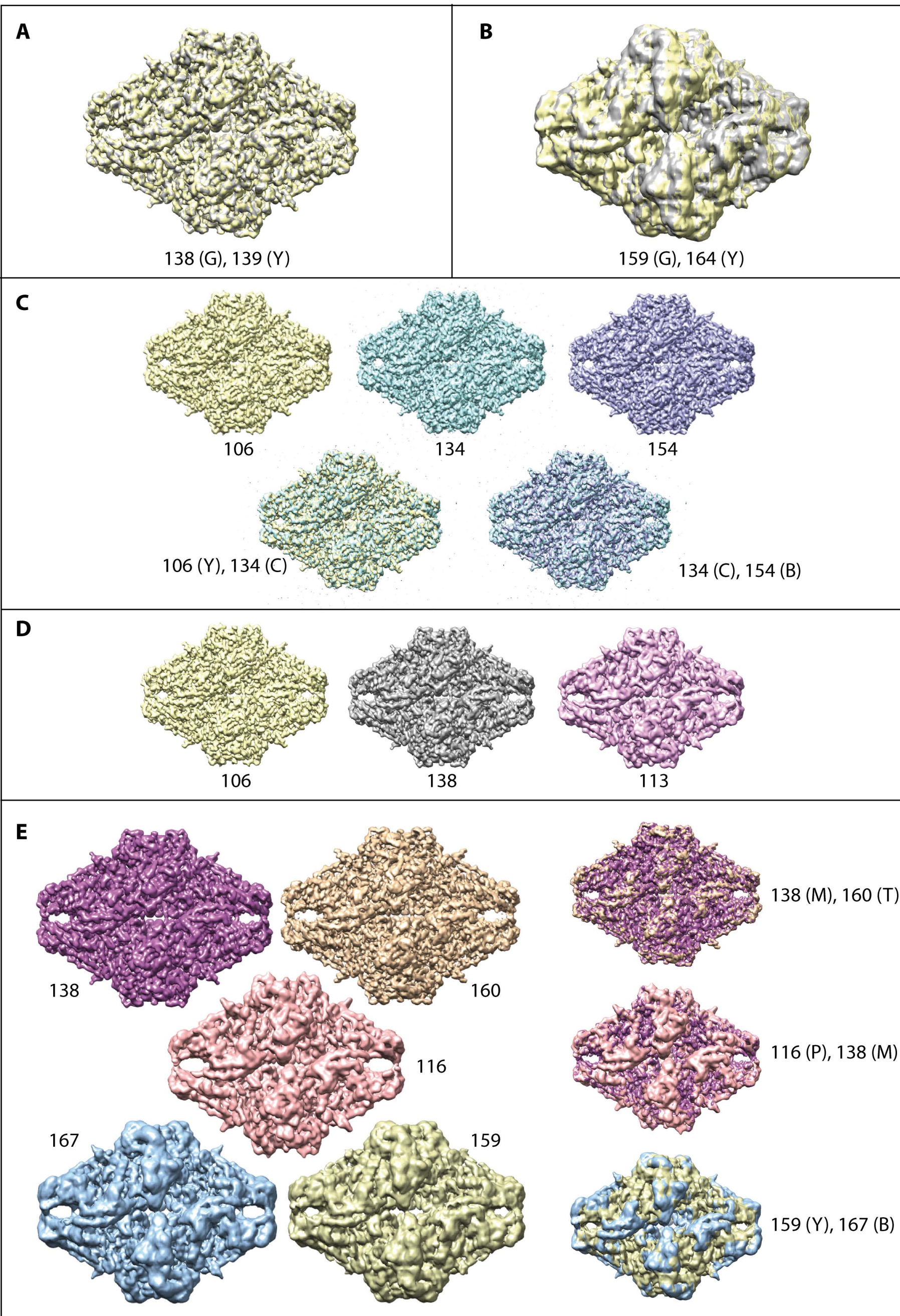


Figure S2

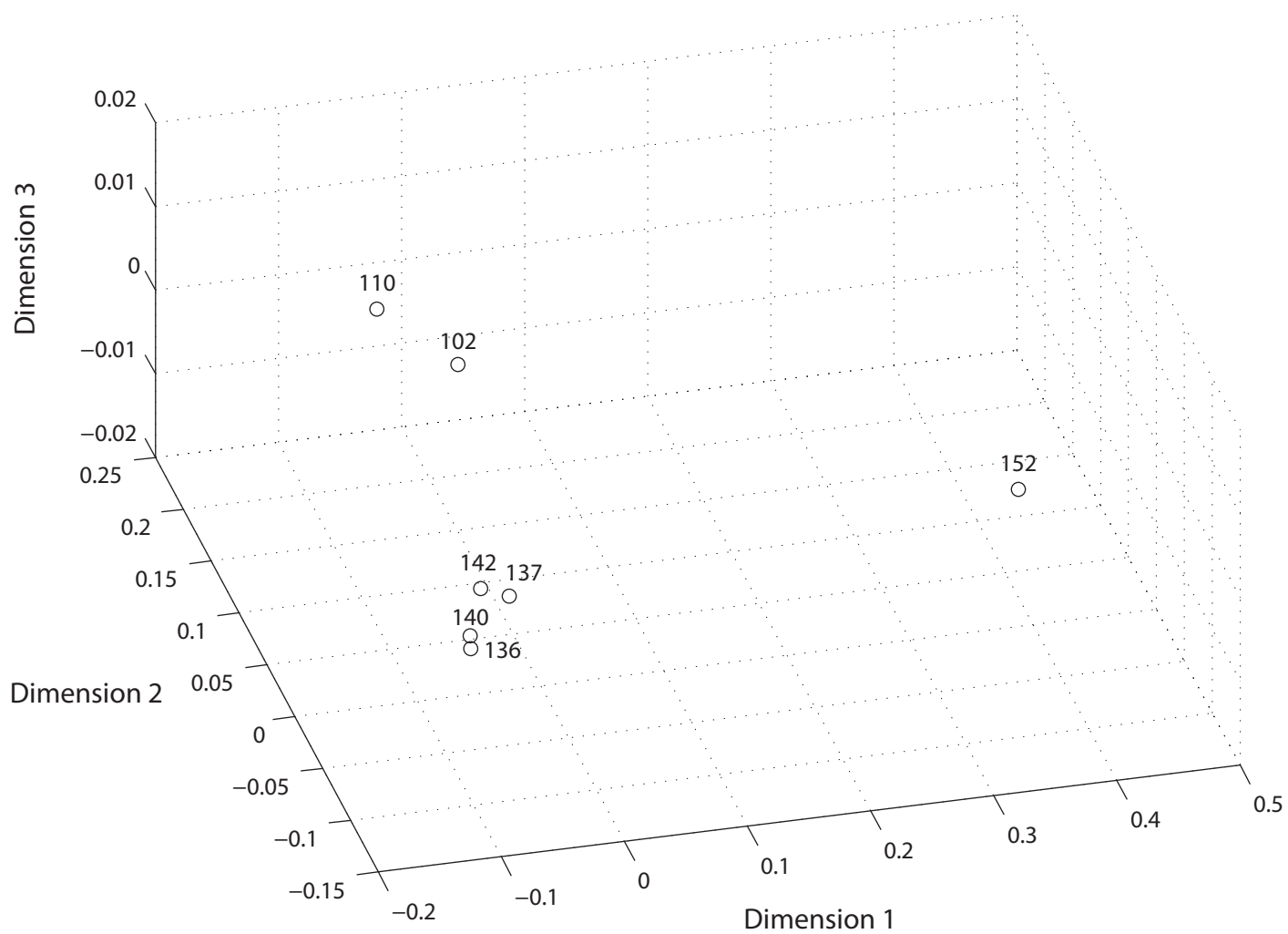


Figure S3

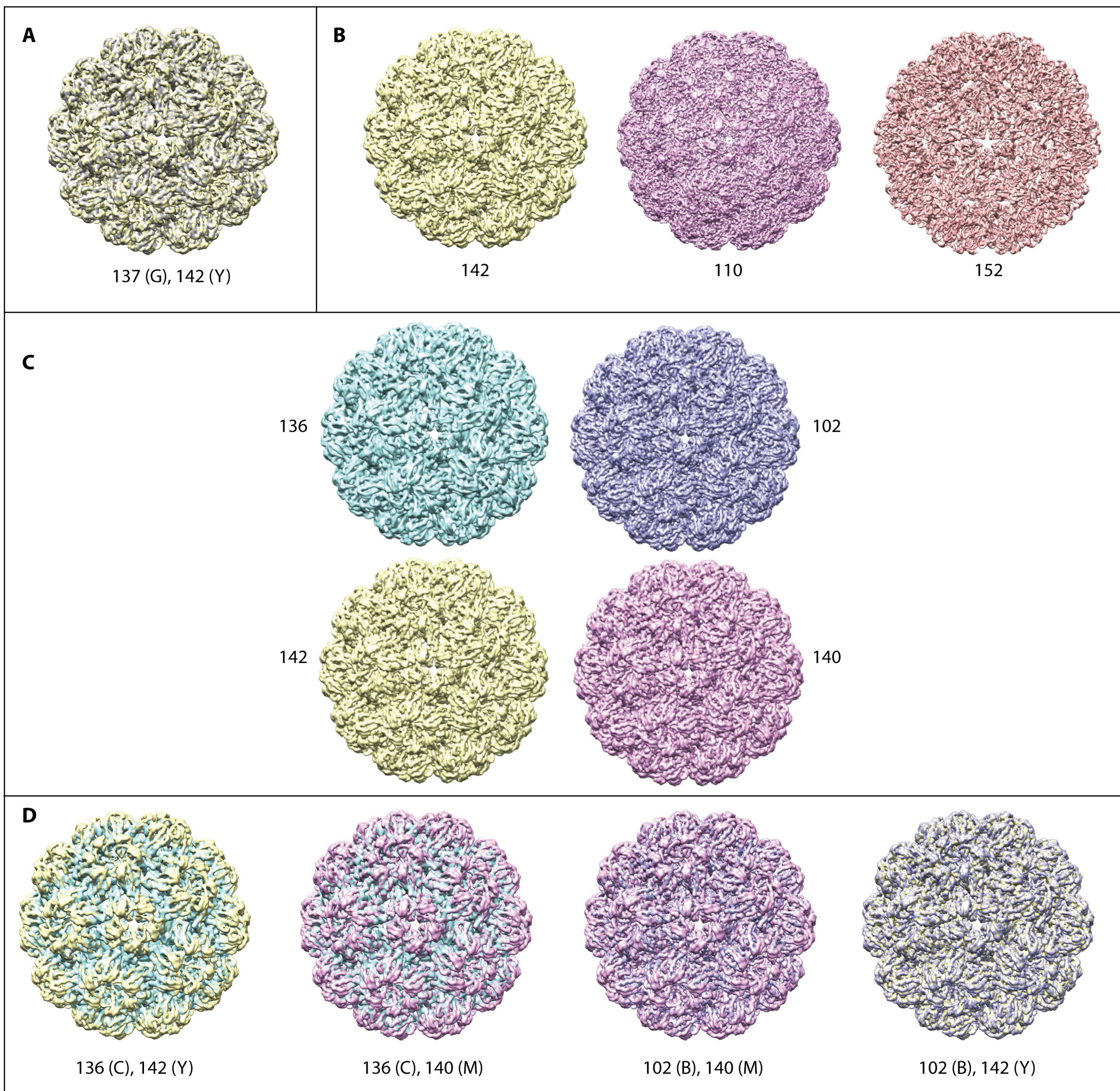


Figure S4

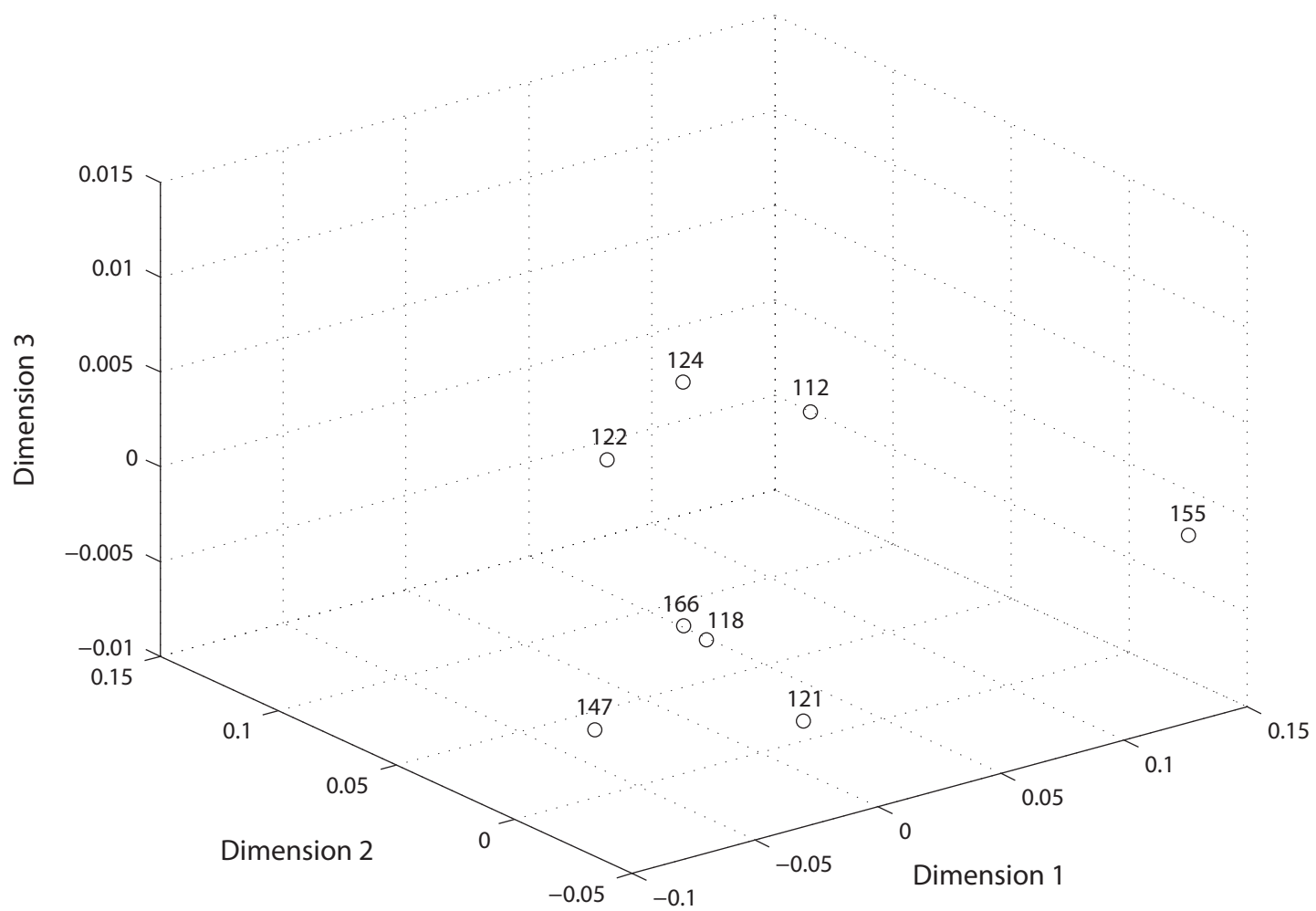


Figure S5

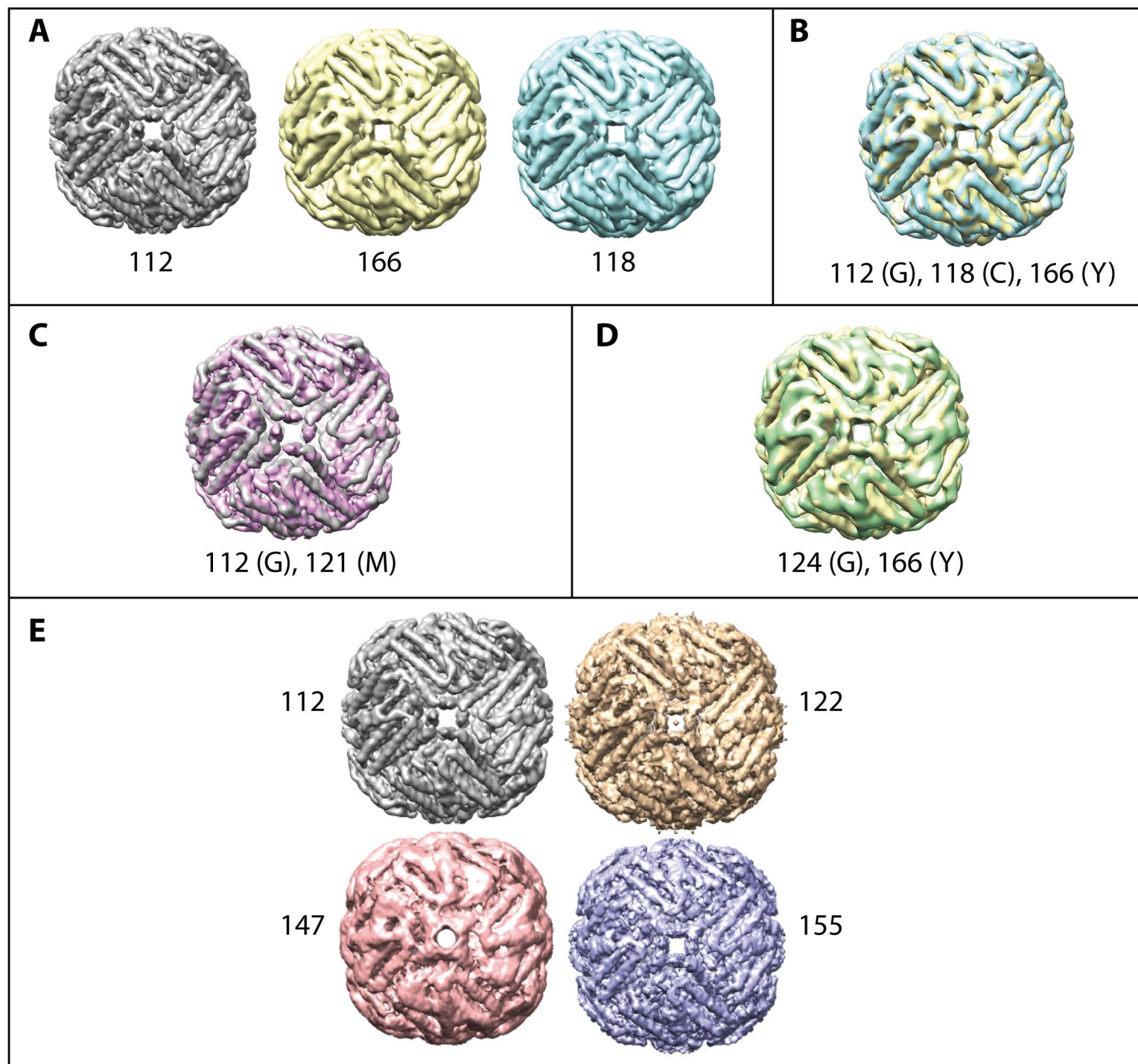


Figure S6

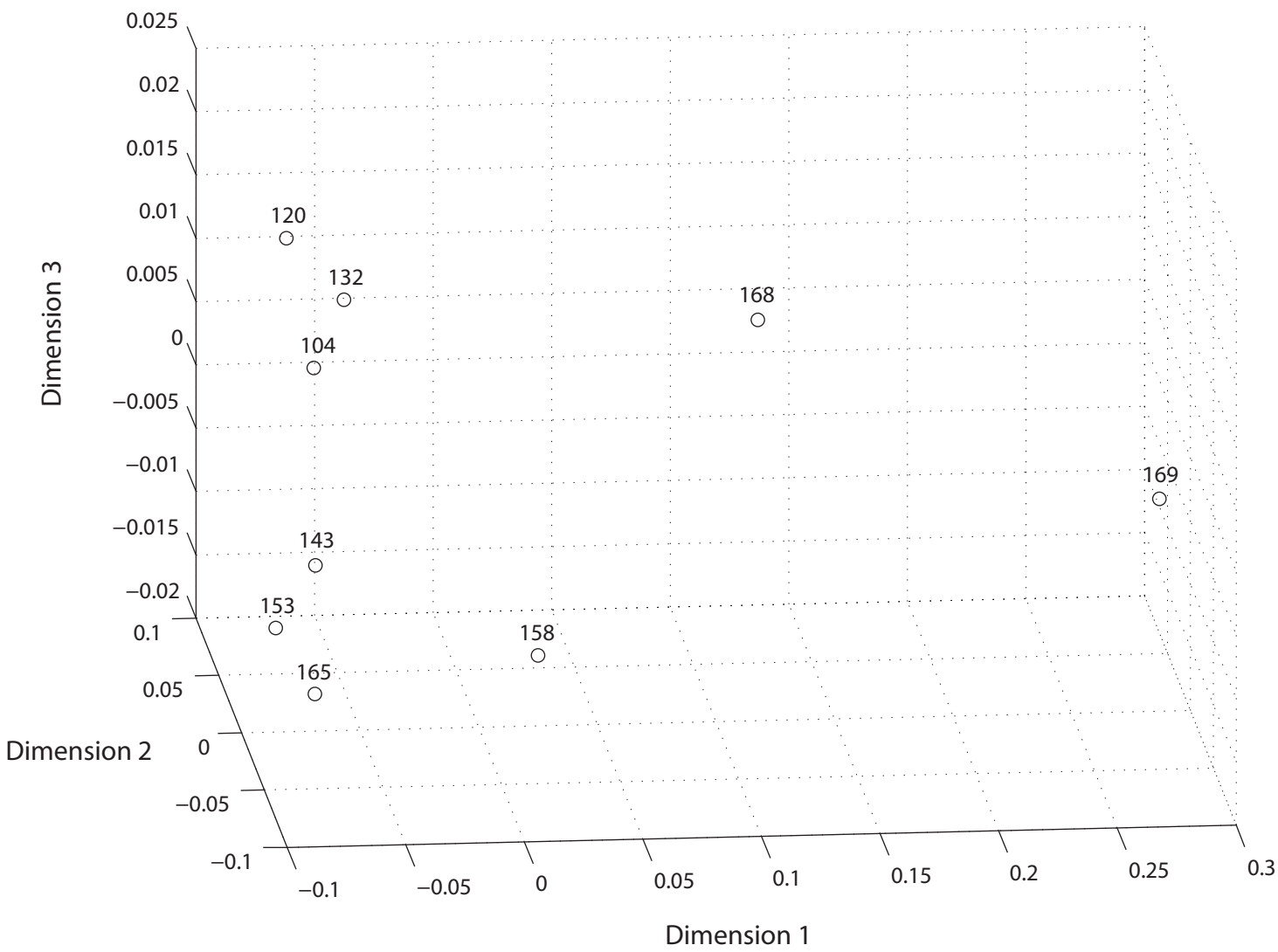
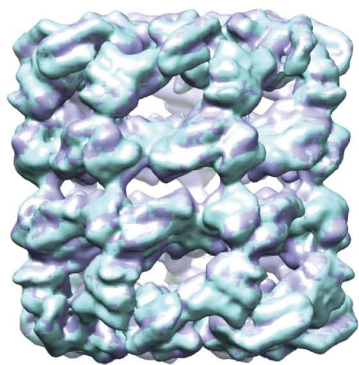
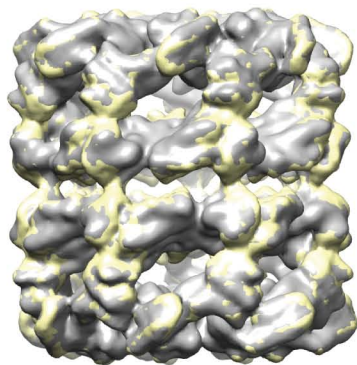


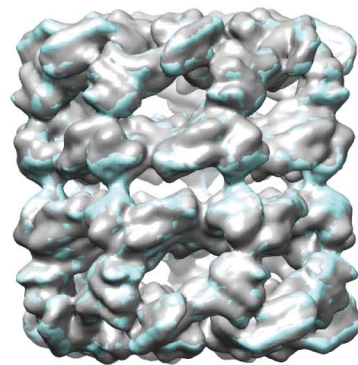
Figure S7

A

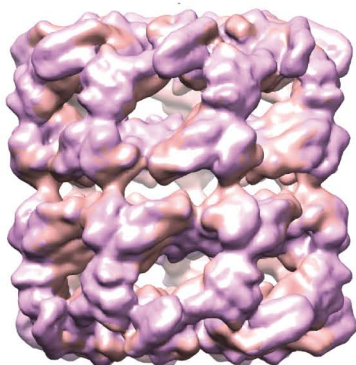
104 (C), 143 (B)

B

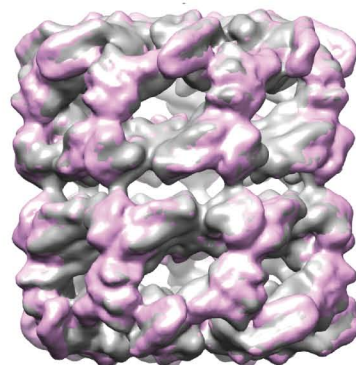
132 (Y), 153 (G)

C

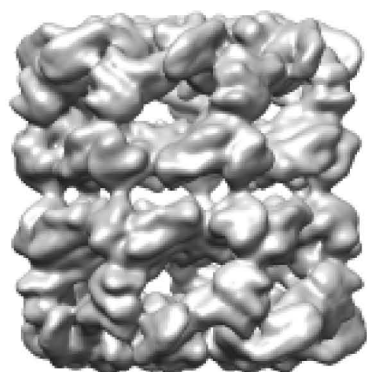
104 (C), 153 (G)

D

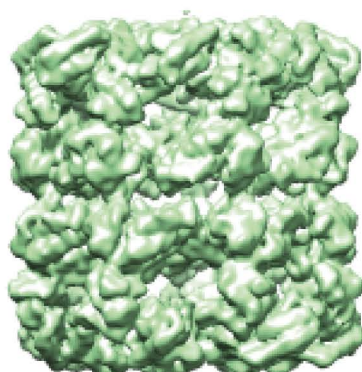
120 (M), 165 (P)

E

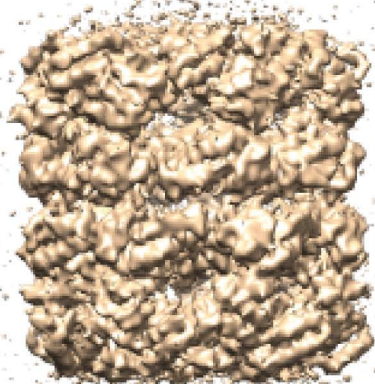
120 (M), 153 (G)

F

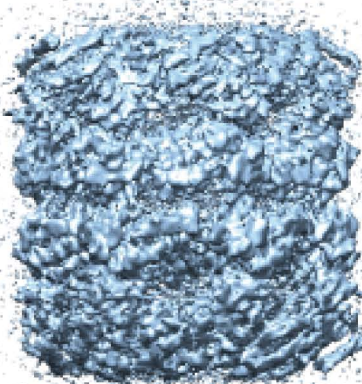
153



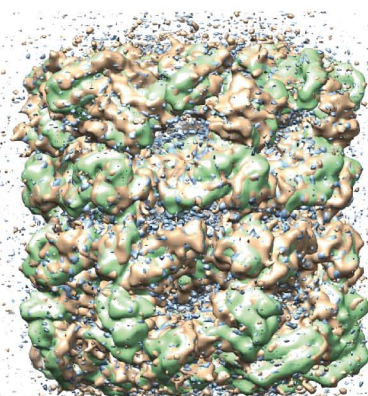
158



168



169



158 (G), 168 (T), 169 (B)

Figure S8

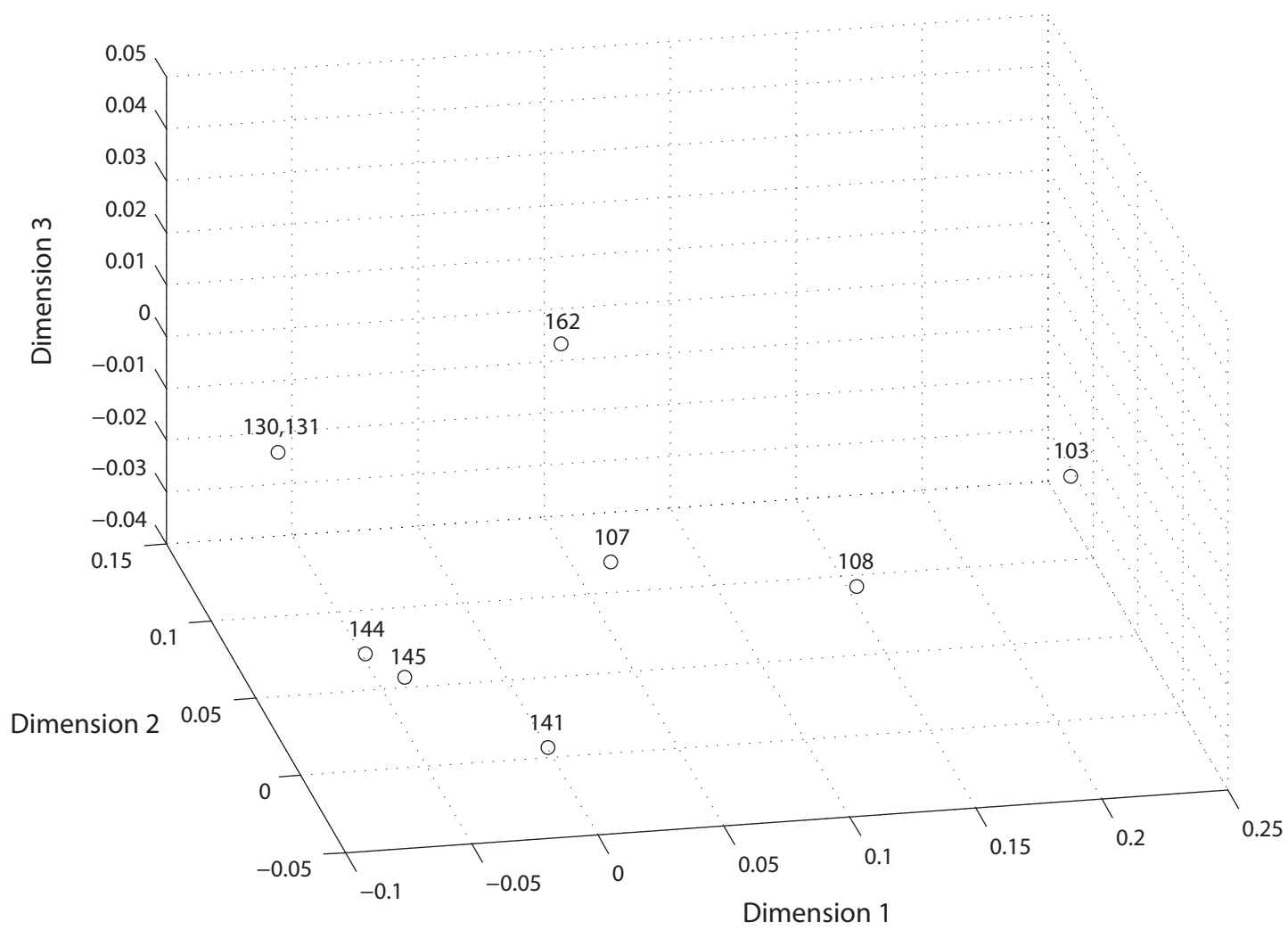
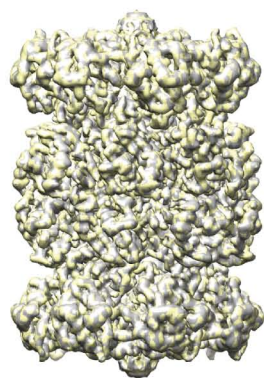
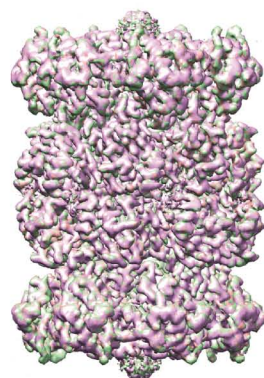


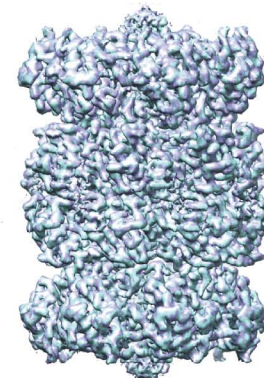
Figure S9

A

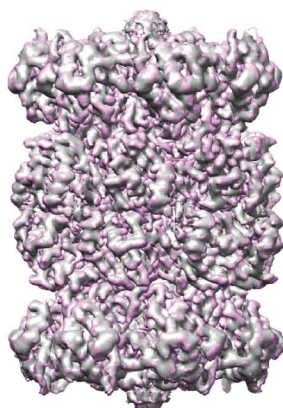
144 (G), 145 (Y)



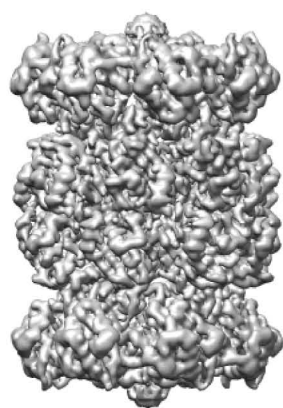
107 (G), 141 (M), 162 (P)

B

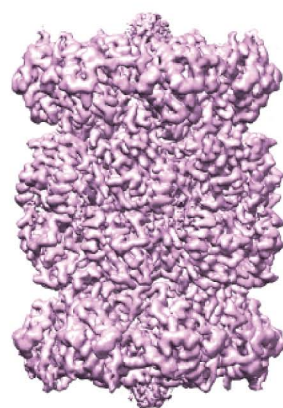
130 (C), 131 (B)

C

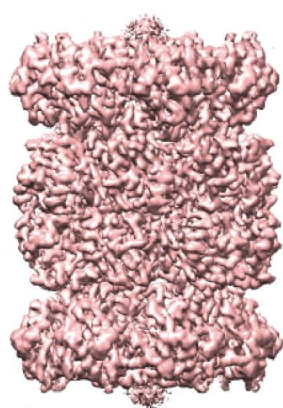
141 (M), 144 (G)



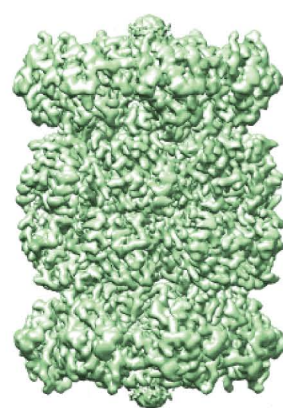
144



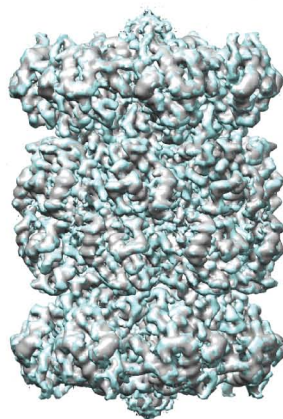
141



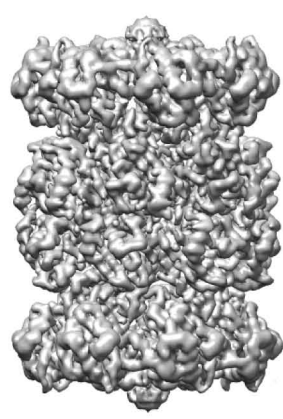
162



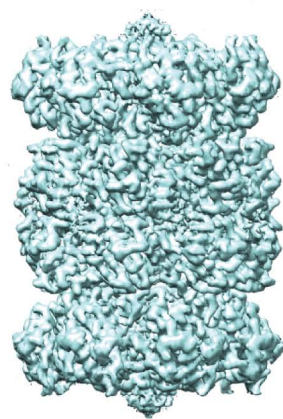
107

D

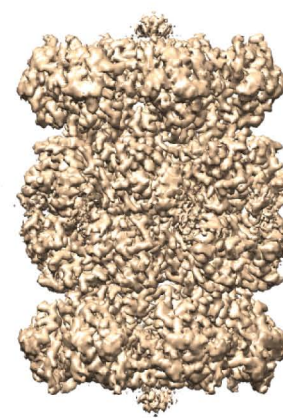
130 (C), 144 (G)



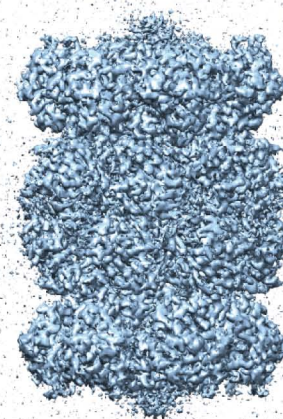
144



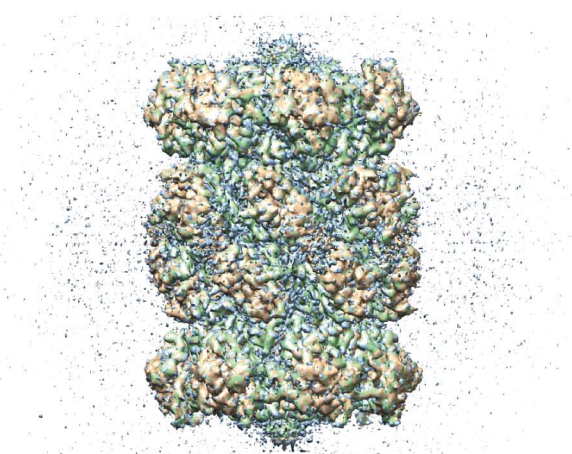
130



108



103

E

103 (B), 107 (G), 108 (T)

Figure S10

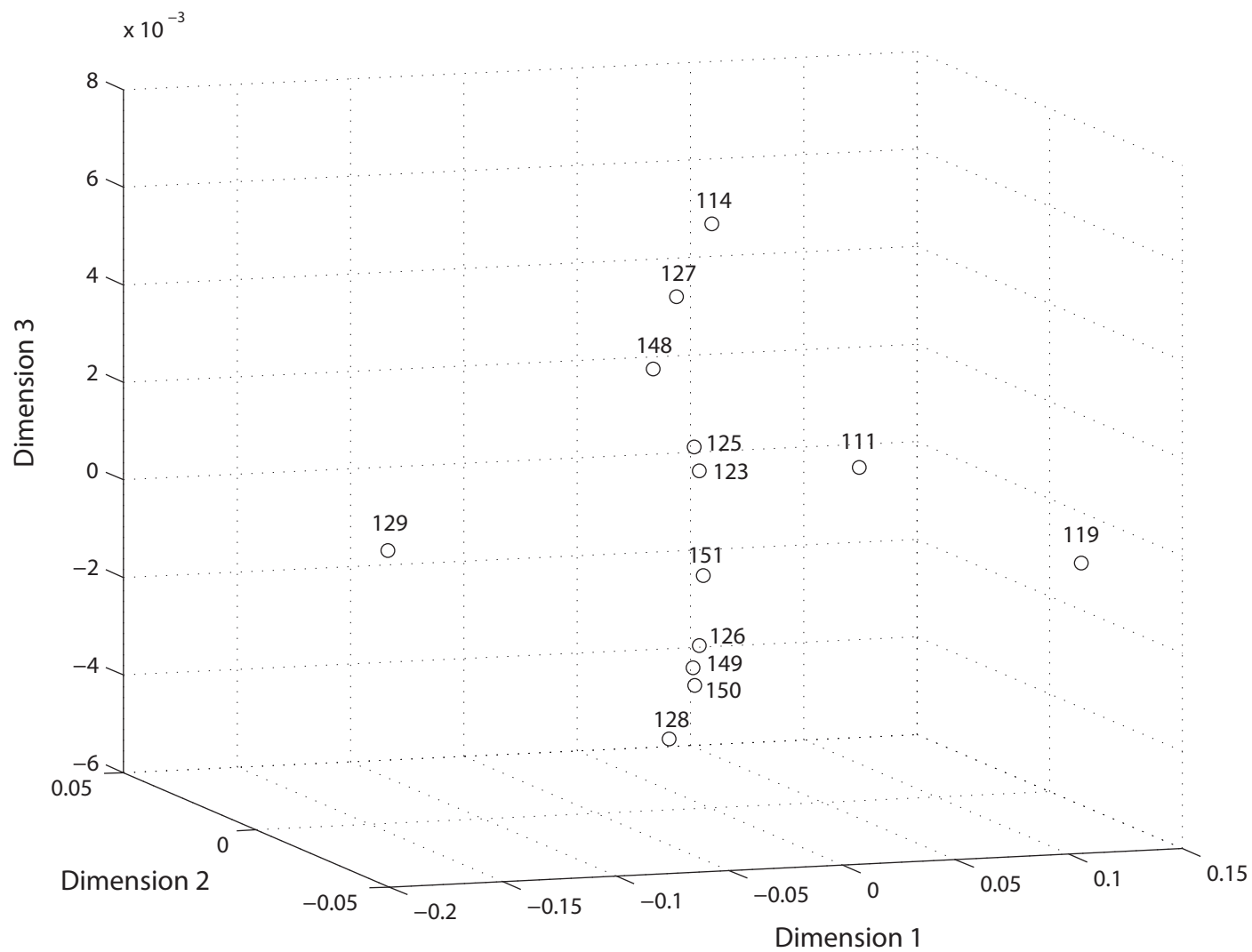
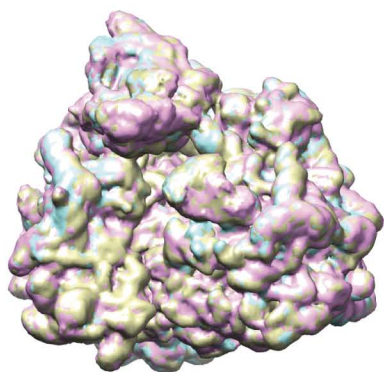
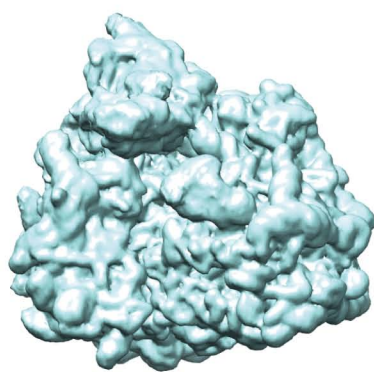


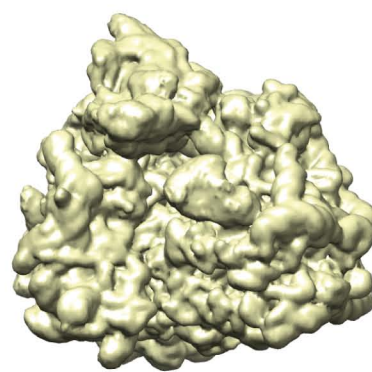
Figure S11

A

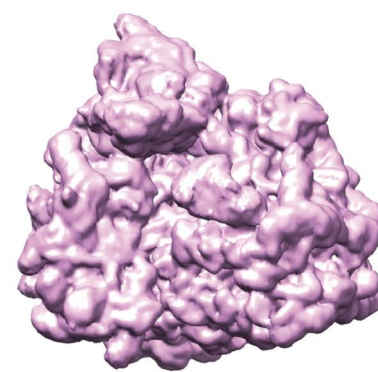
123 (Y), 148 (M), 150 (C)



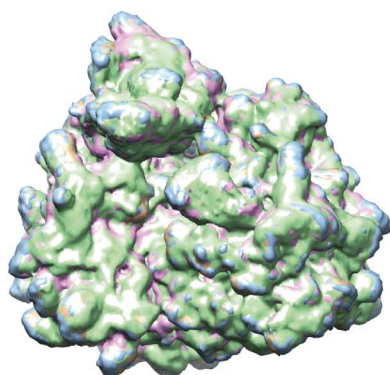
150



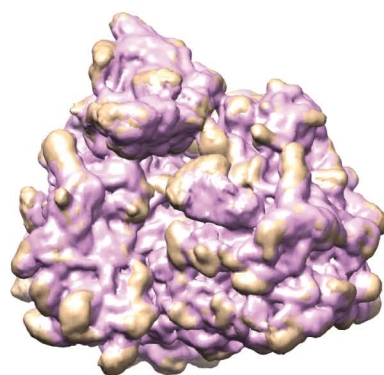
123



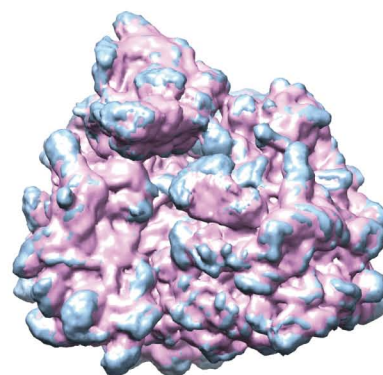
148

B

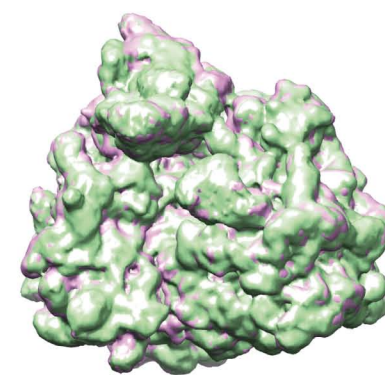
111 (T), 119 (B), 129 (G), 148 (M)



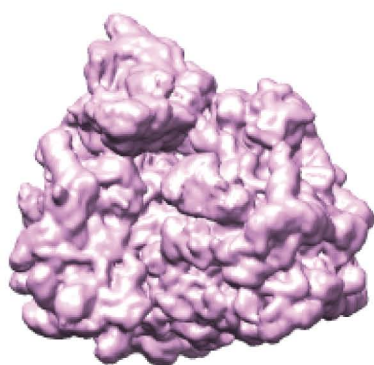
111 (T), 148 (M)



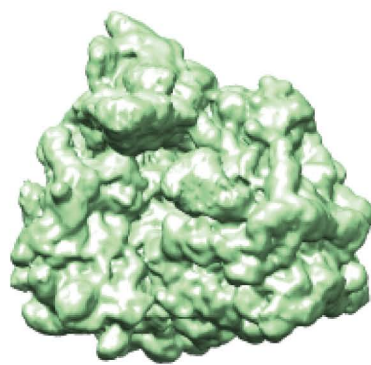
119 (B), 148 (M)



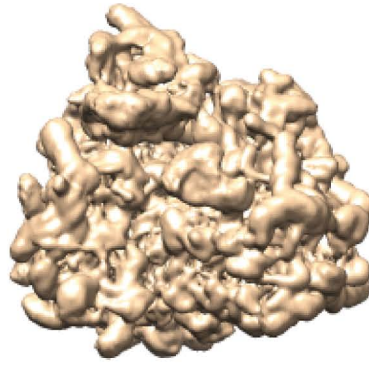
129 (G), 148 (M)

C

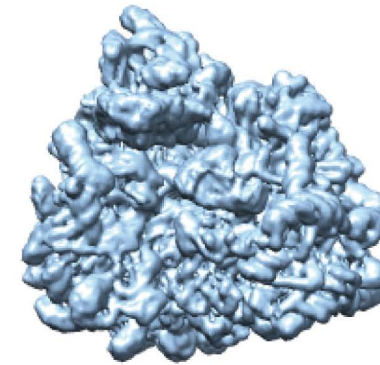
148



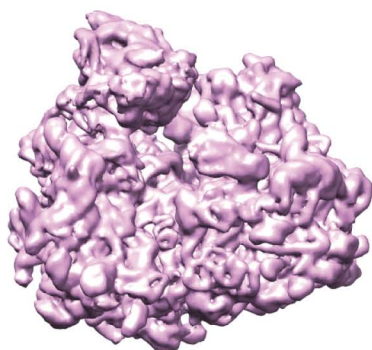
129



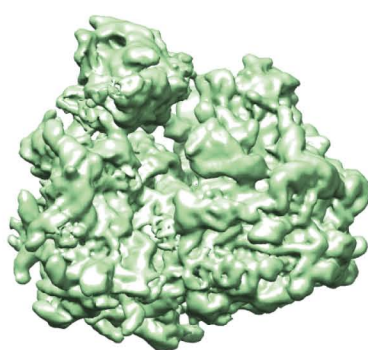
111



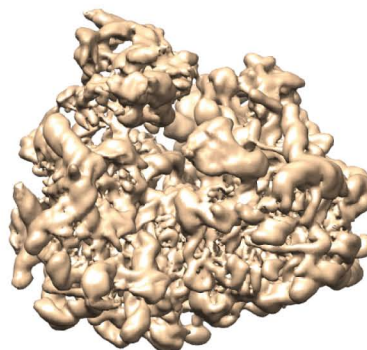
119

D

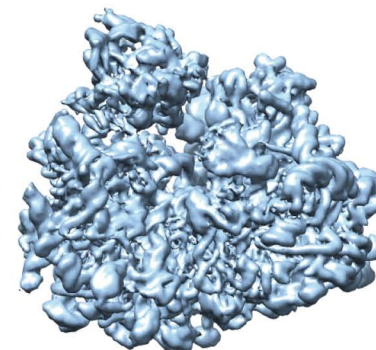
148



129



111



119

Figure S12

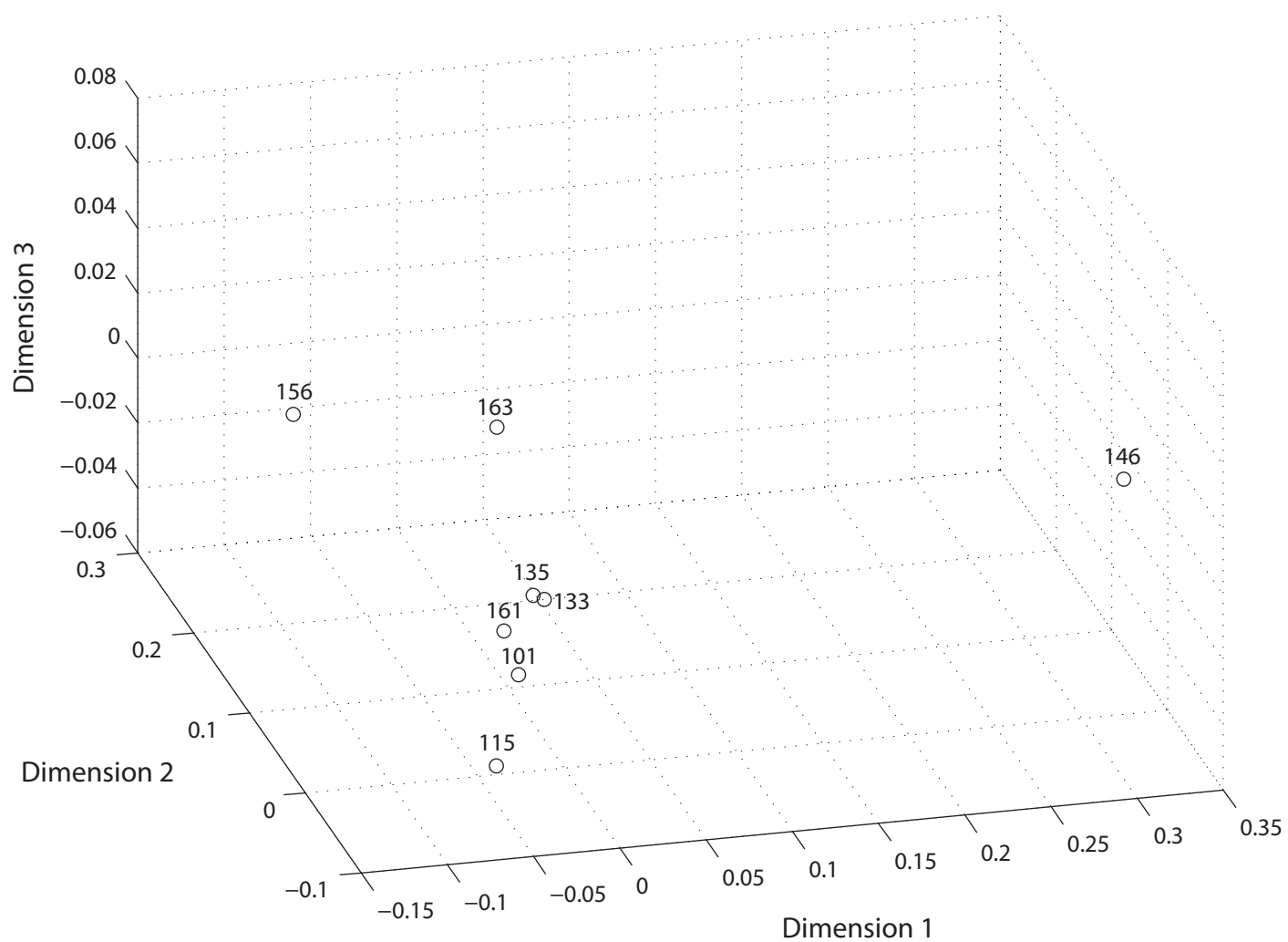
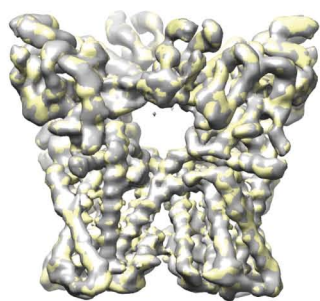
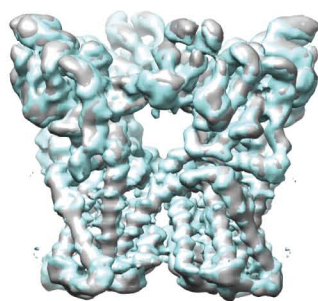


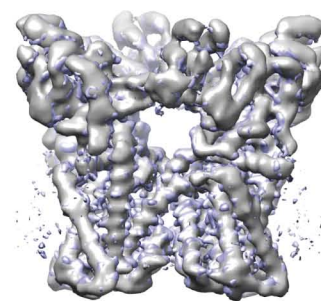
Figure S13

A

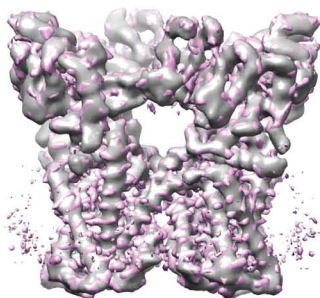
133 (G), 135 (Y)



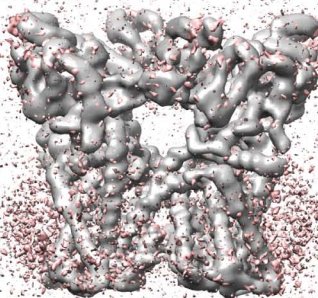
133 (G), 101 (C)



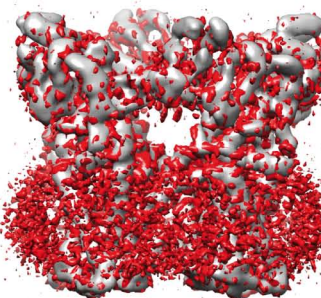
133 (G), 115 (B)



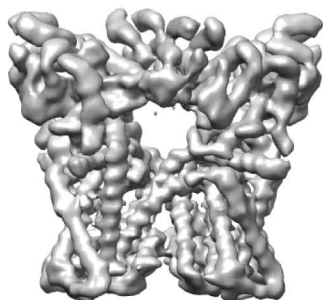
133 (G), 161 (M)



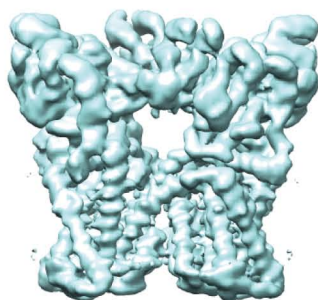
133 (G), 146 (P)



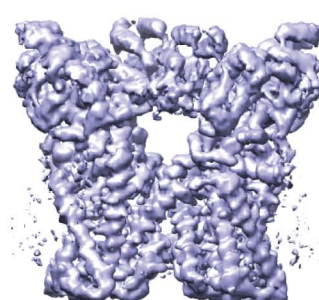
133 (G), 156 (R)

B

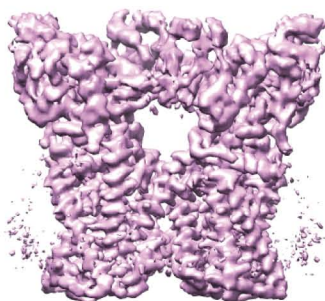
133



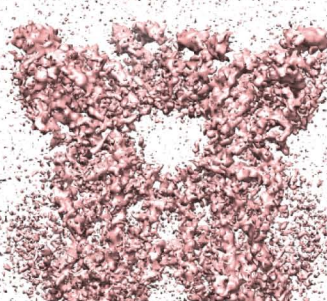
101



115



161



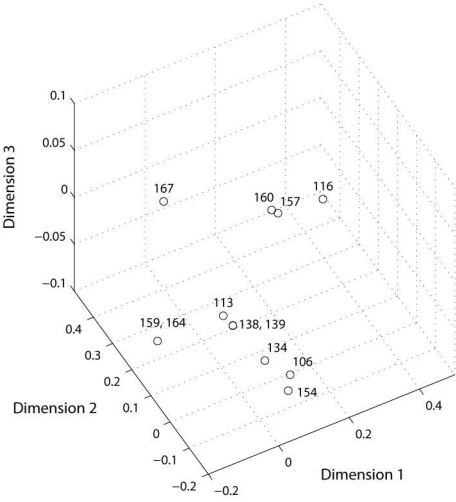
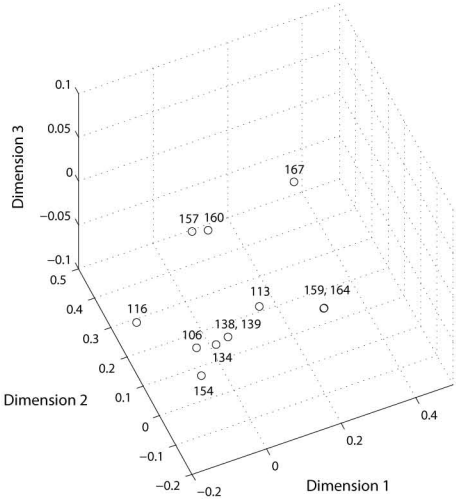
146



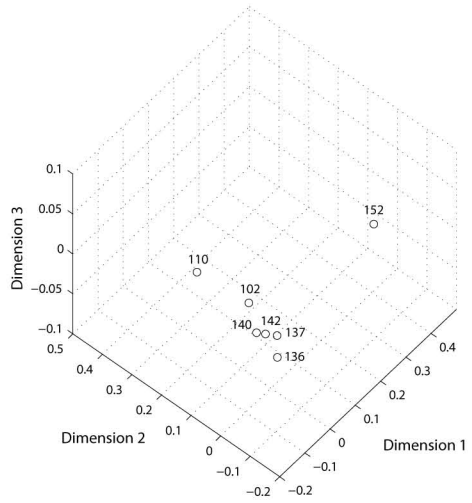
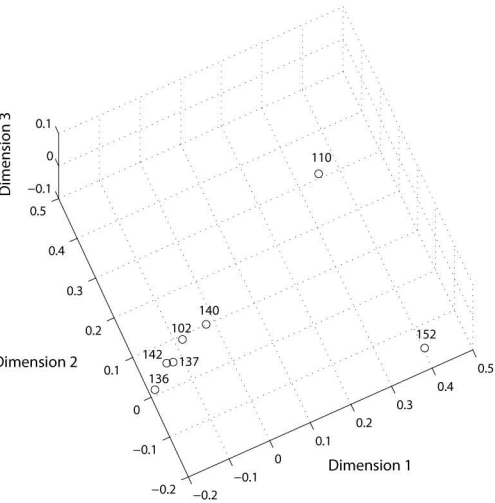
156

Figure S14

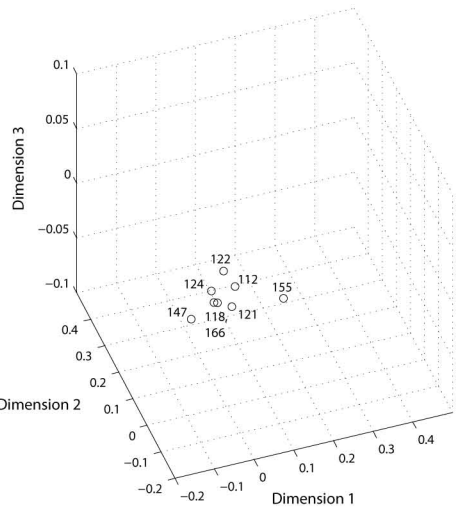
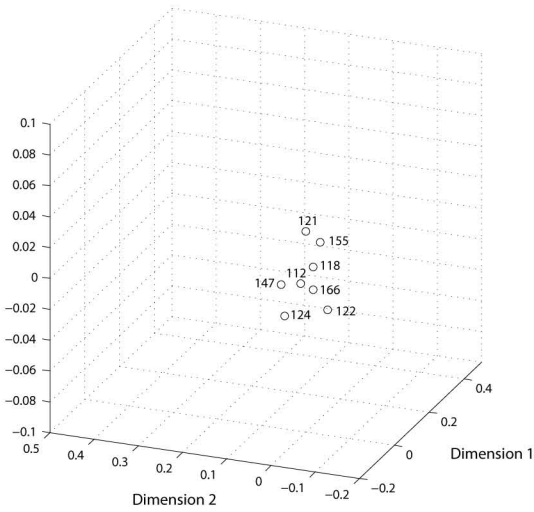
A



B



C



D

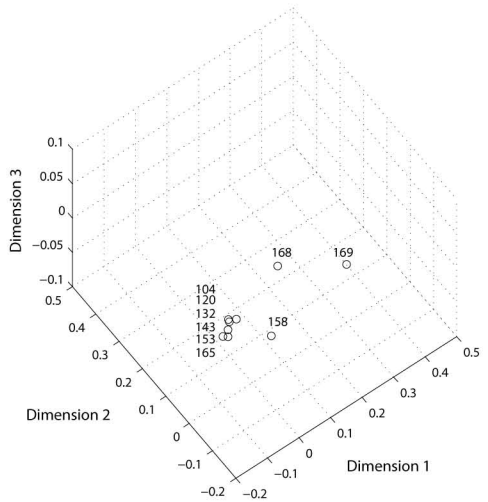
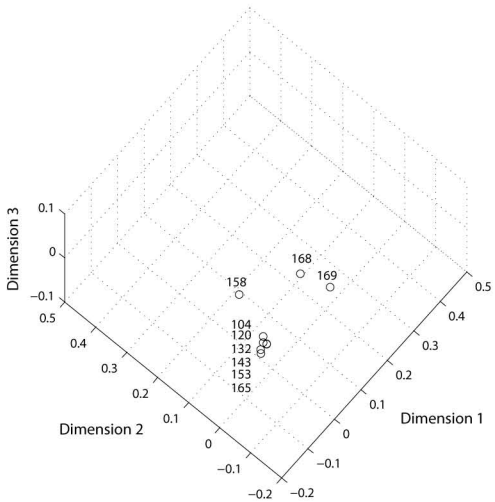
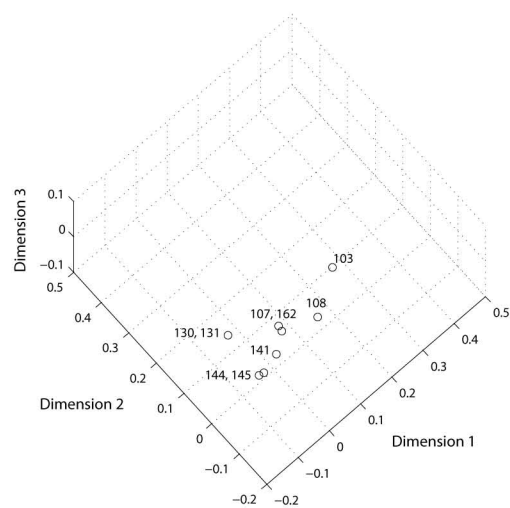
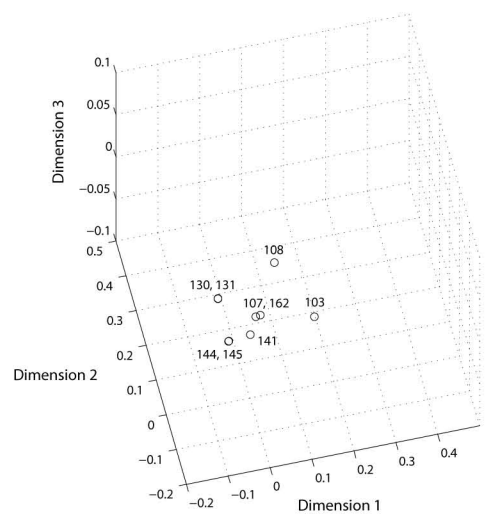
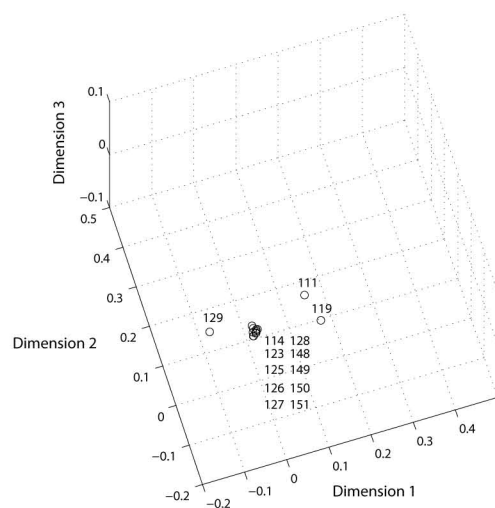
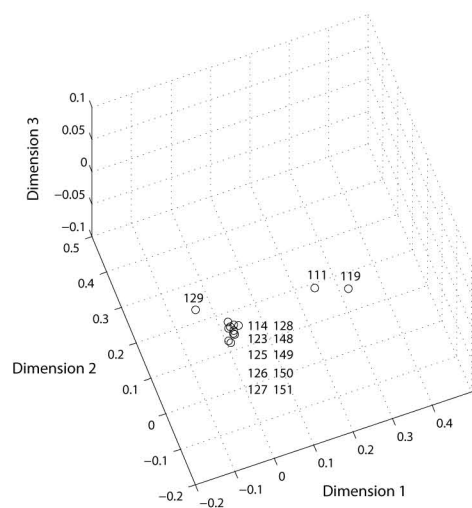
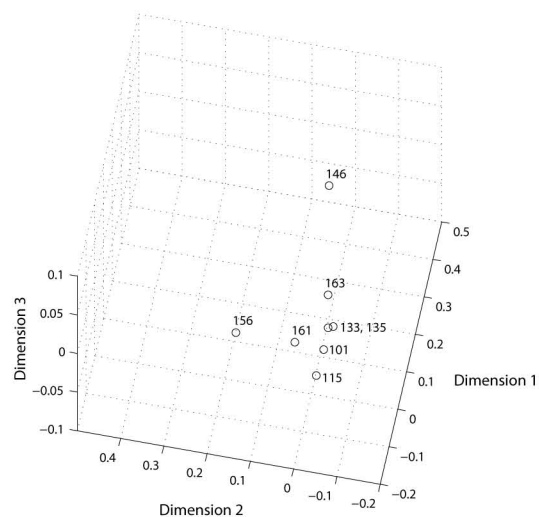
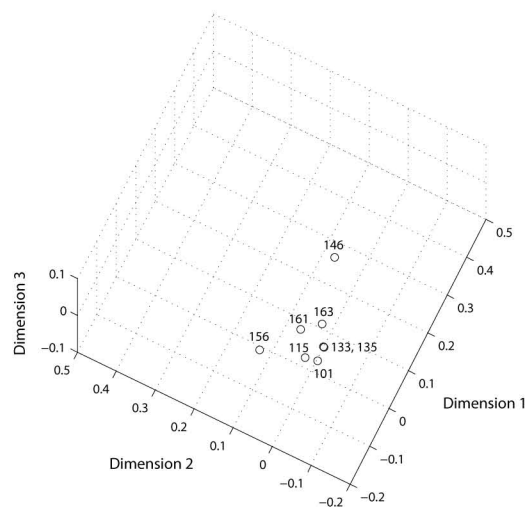
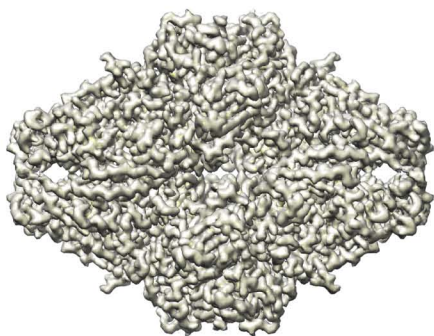


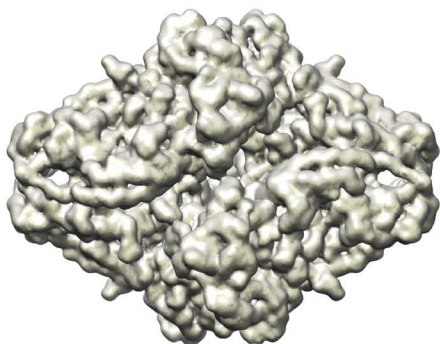
Figure S15

A**B****C****Figure S16**

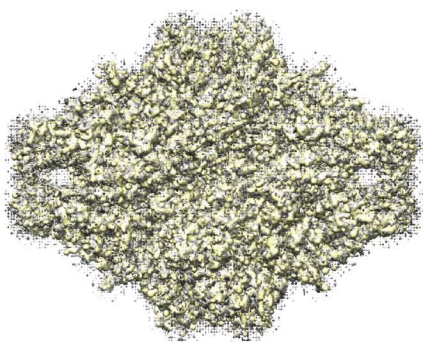
106



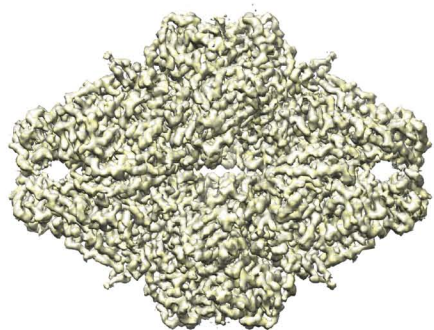
113



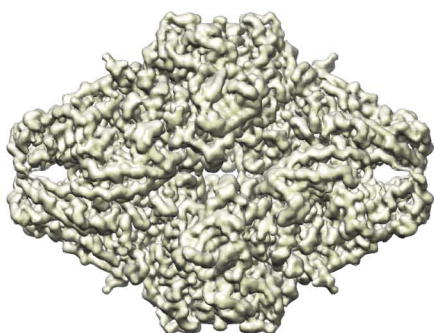
116



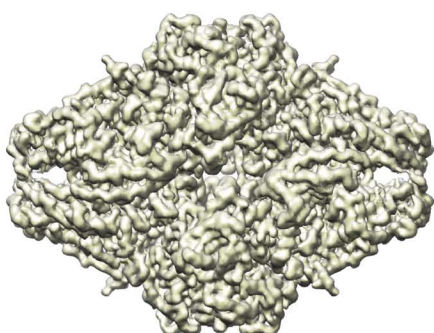
134



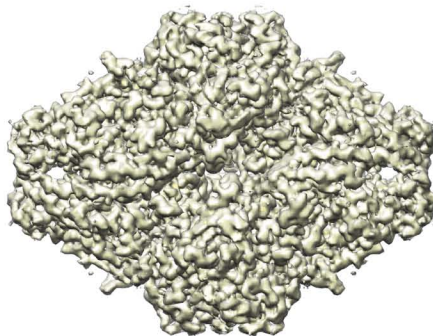
138



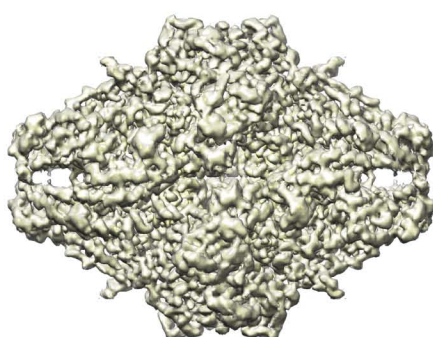
139



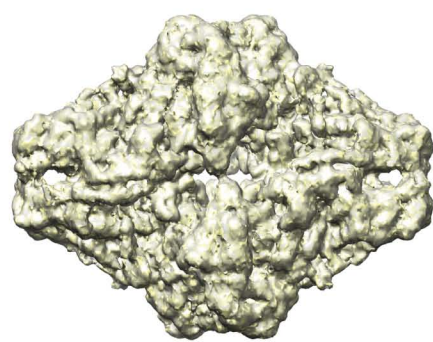
154



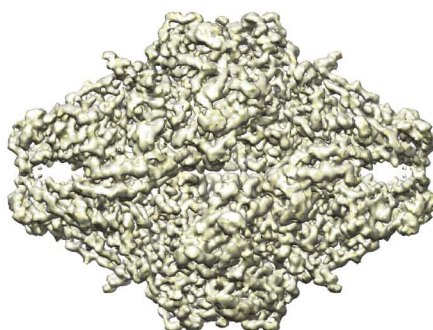
157



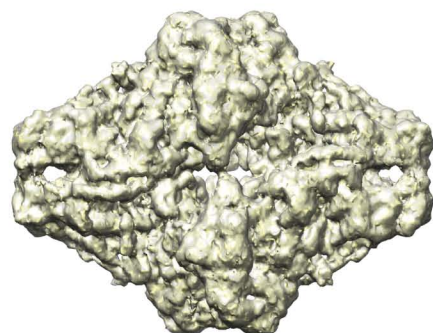
159



160



164



167

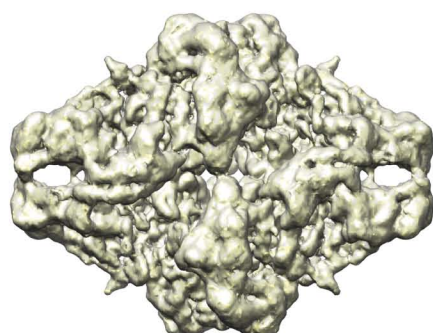
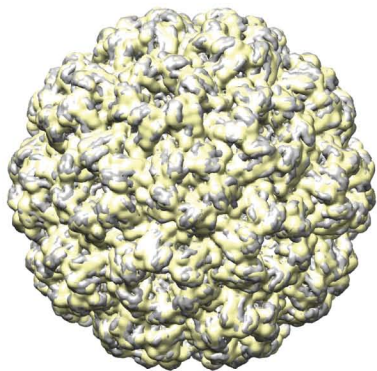
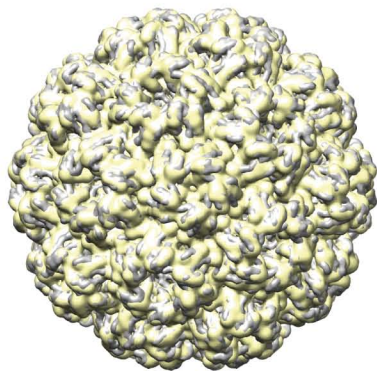


Figure S17

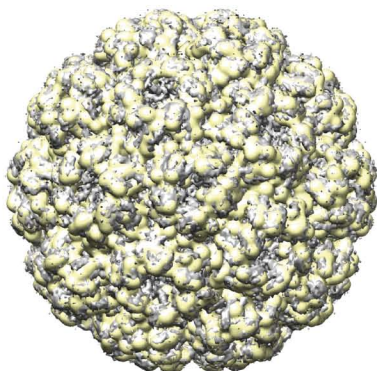
102



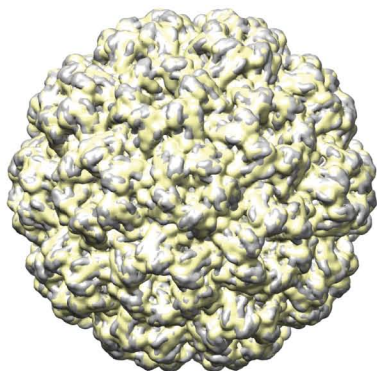
140



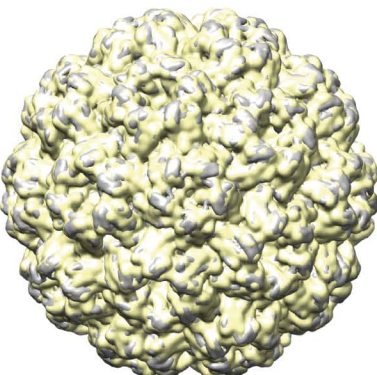
110



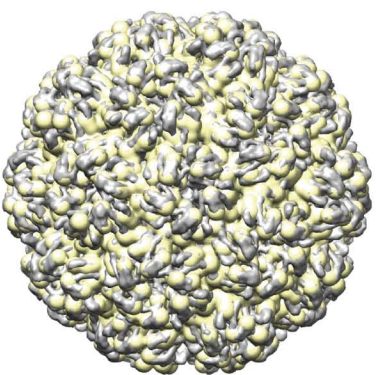
142



136



152



137

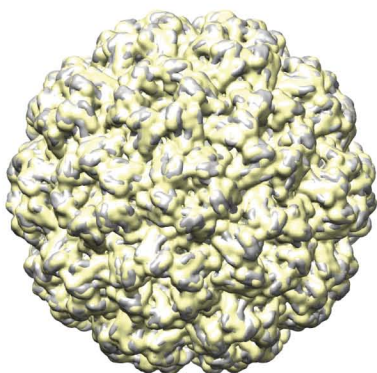
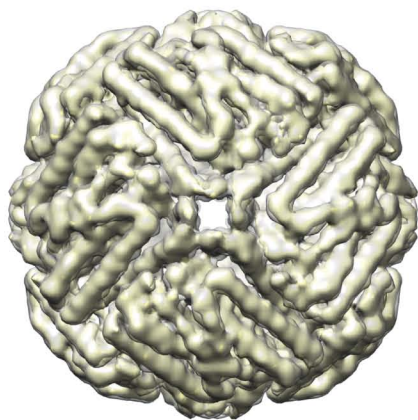
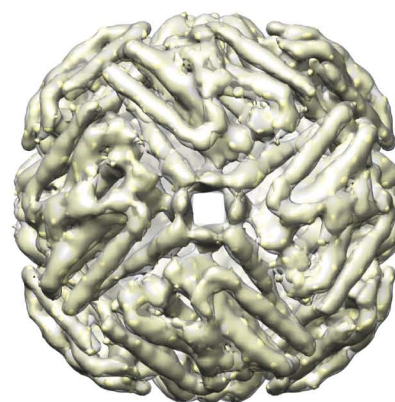


Figure S18

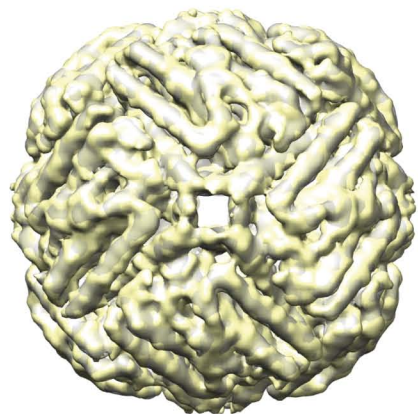
112



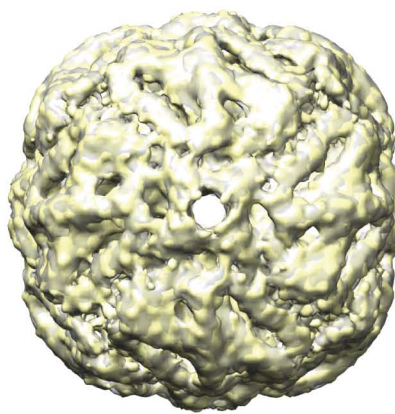
124



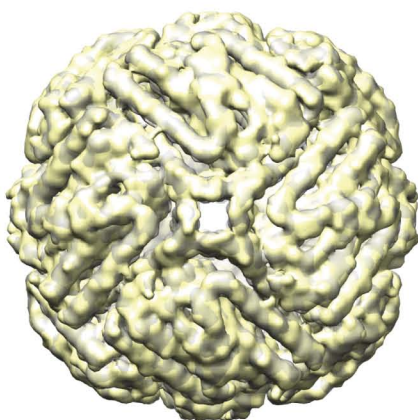
118



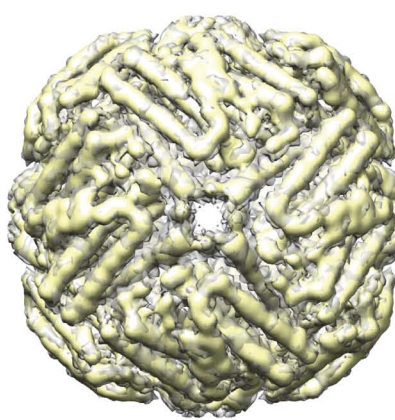
147



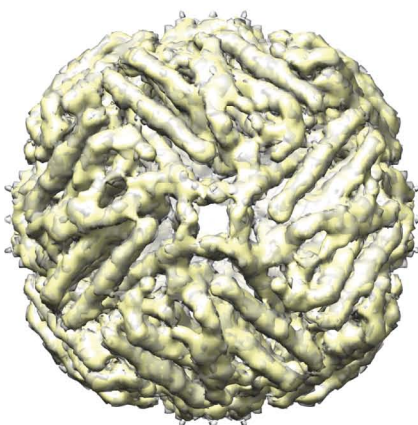
121



155



122



166

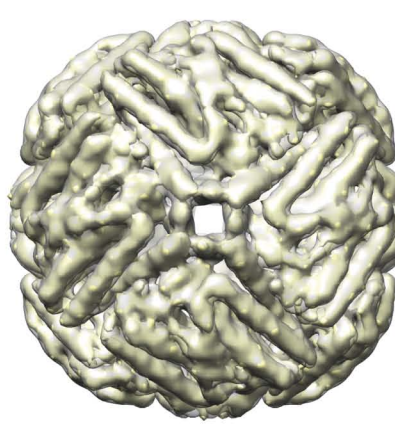
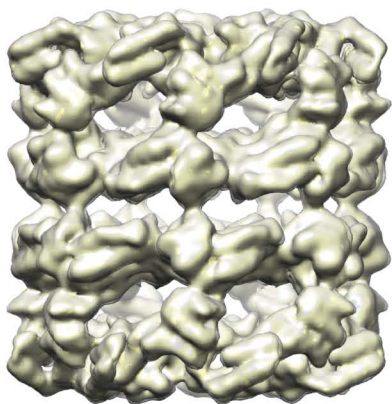
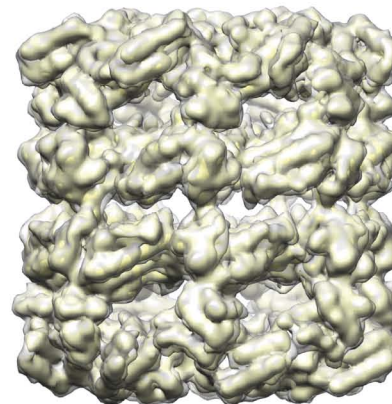


Figure S19

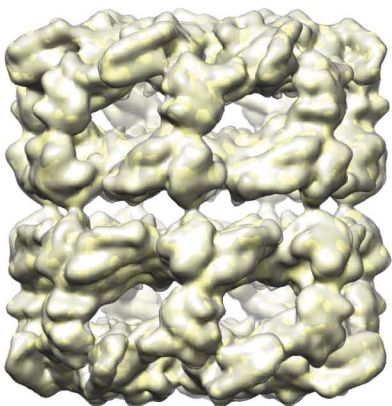
104



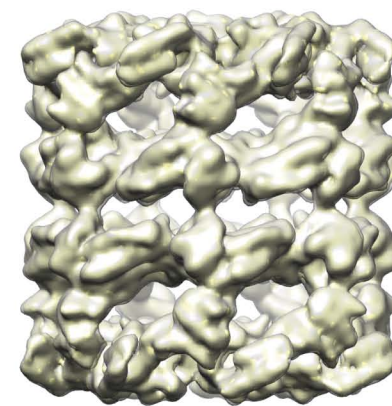
158



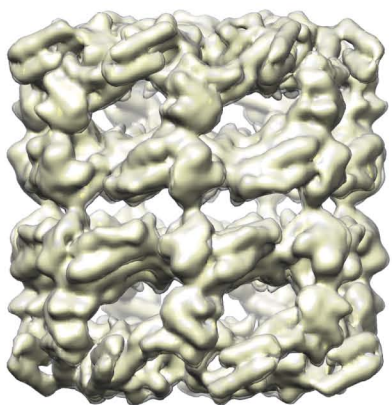
120



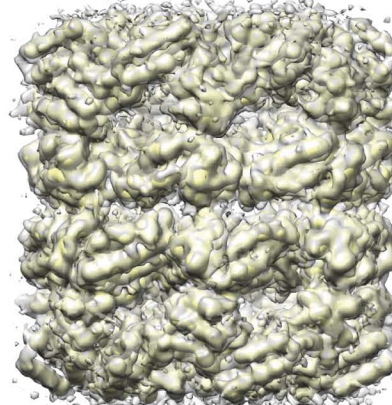
165



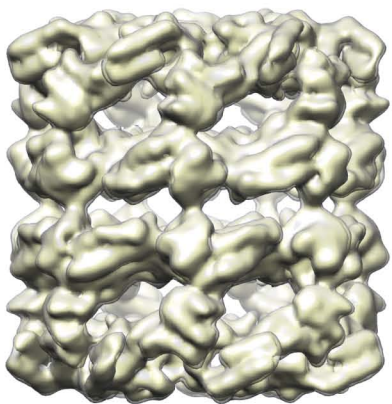
132



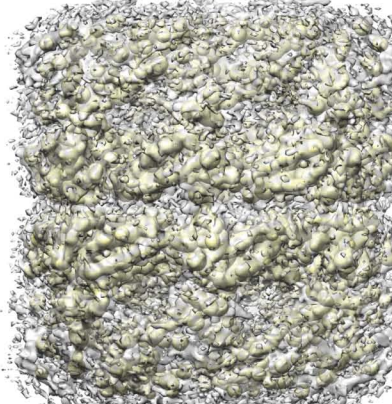
168



143



169



153

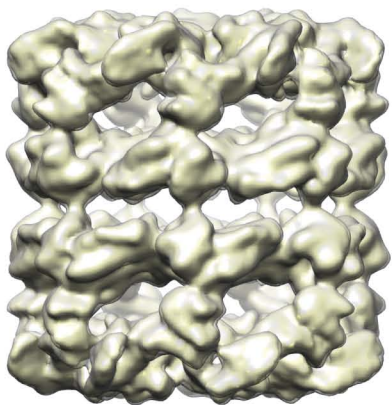
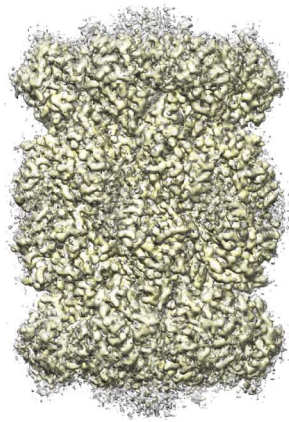
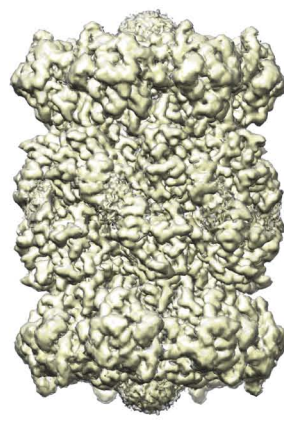


Figure S20

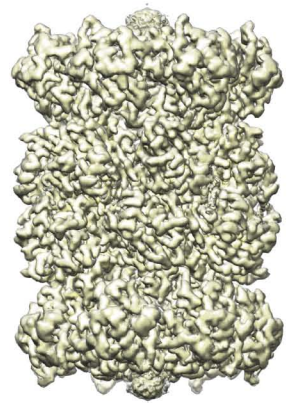
103



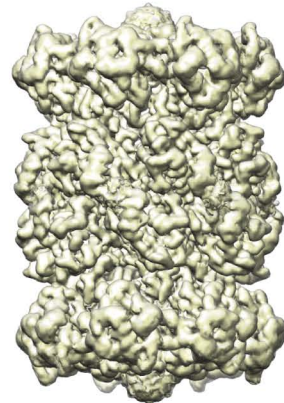
141



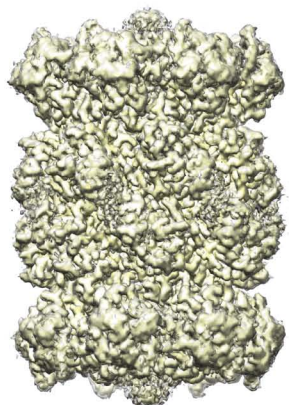
107



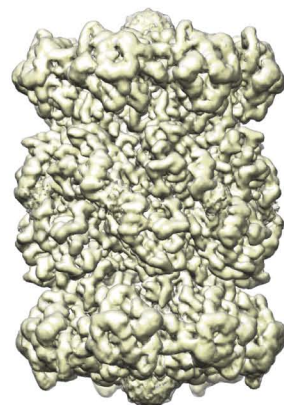
144



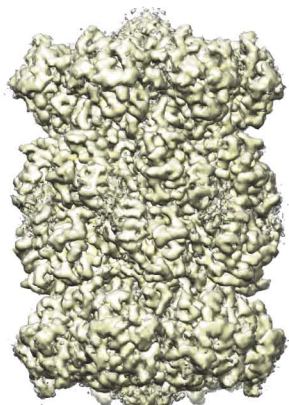
108



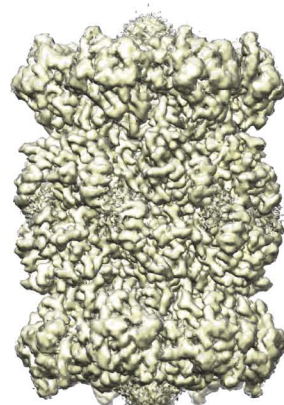
145



130



162



131

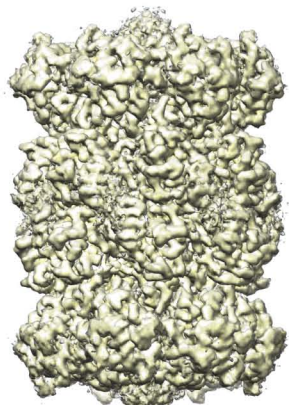
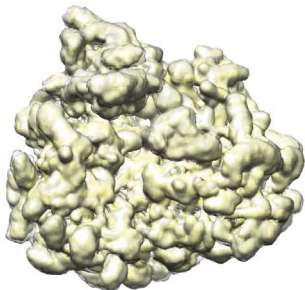
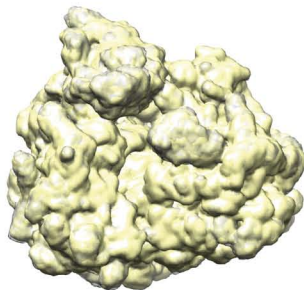


Figure S21

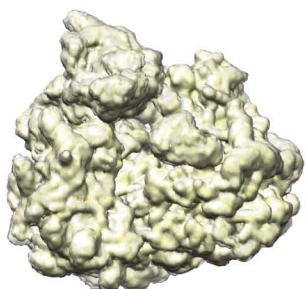
111



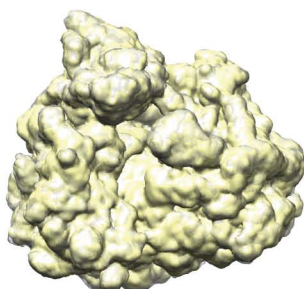
128



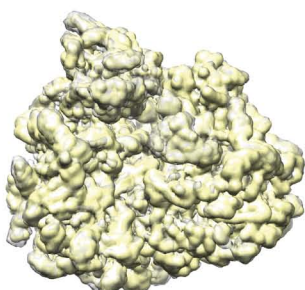
114



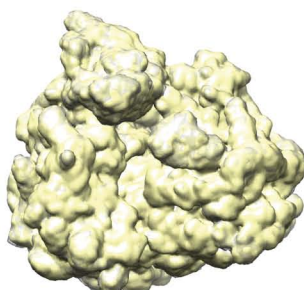
129



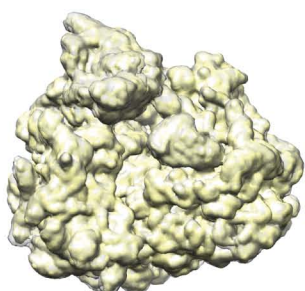
119



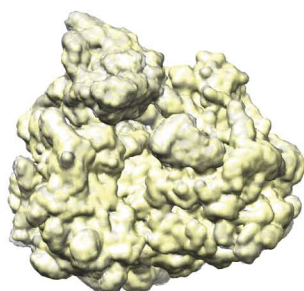
148



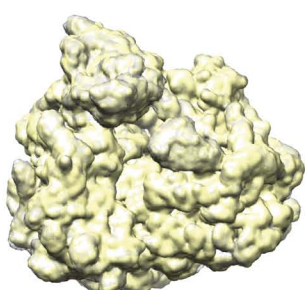
123



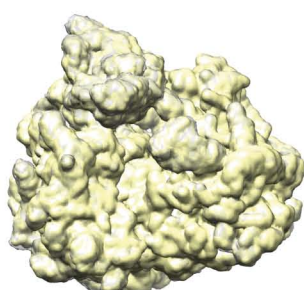
149



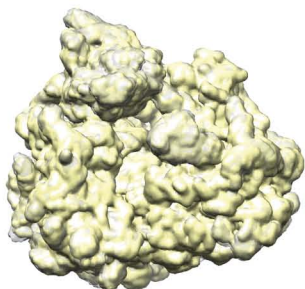
125



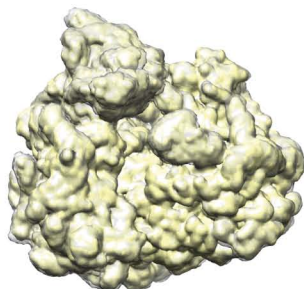
150



126



151



127

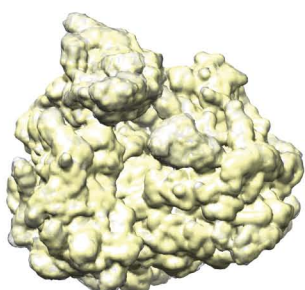
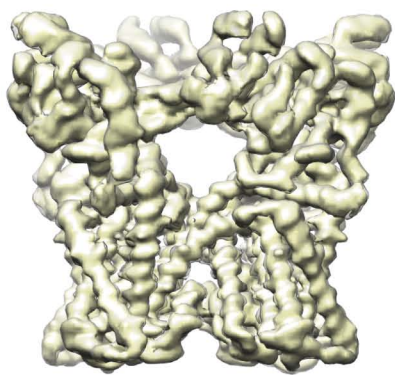
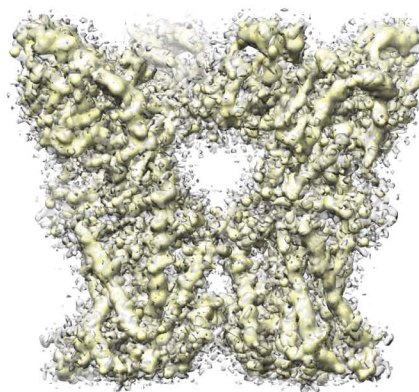


Figure S22

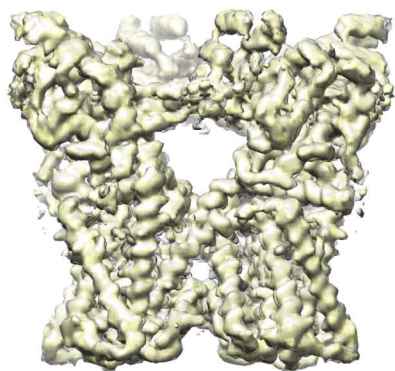
101



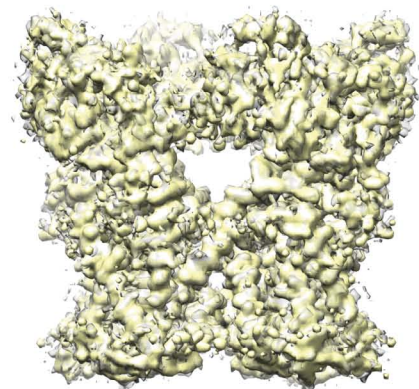
146



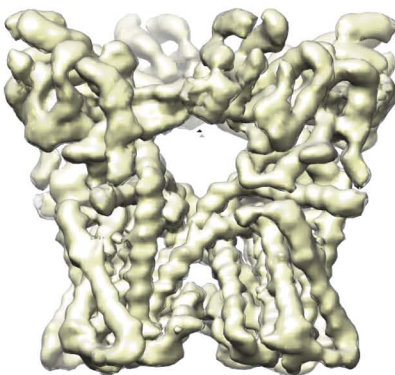
115



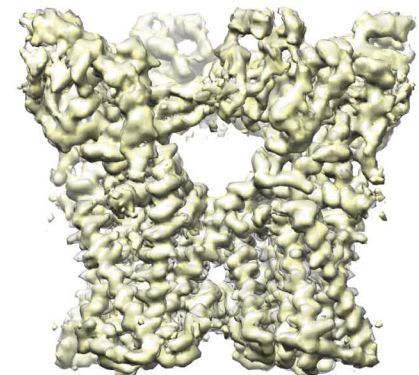
156



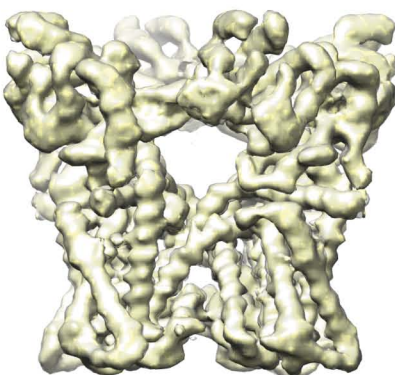
133



161



135



163

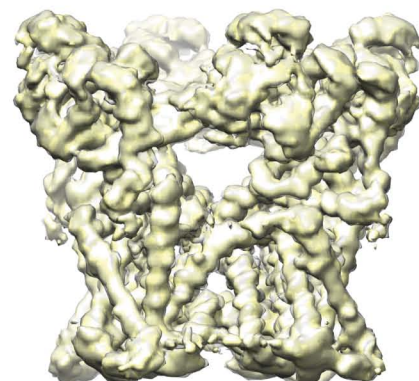
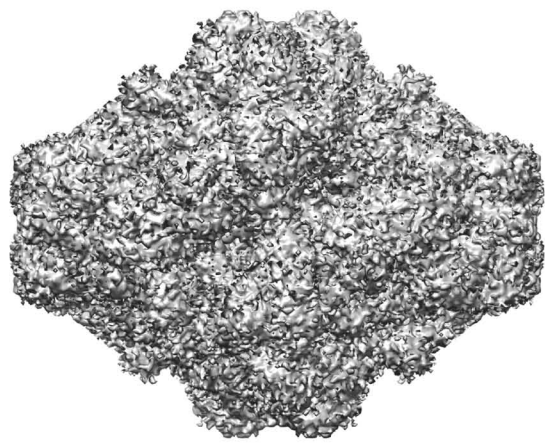


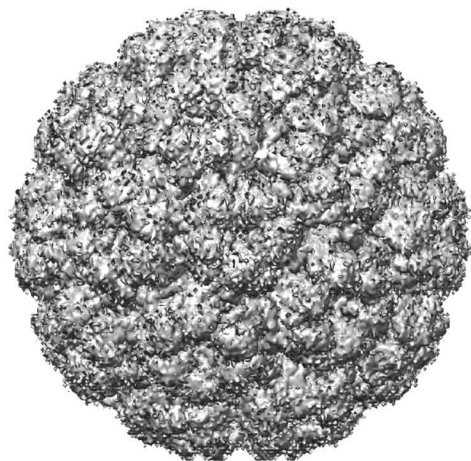
Figure S23

A



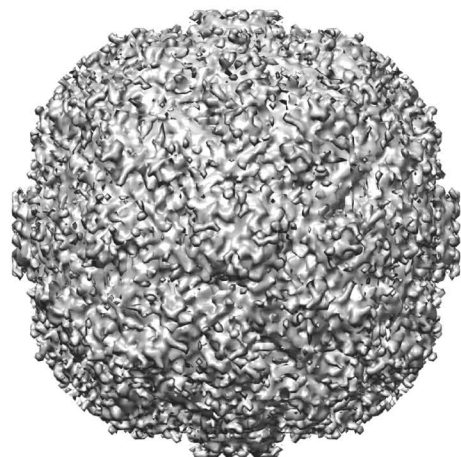
106

B



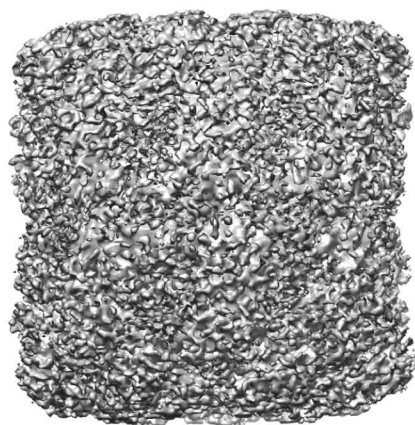
110

C



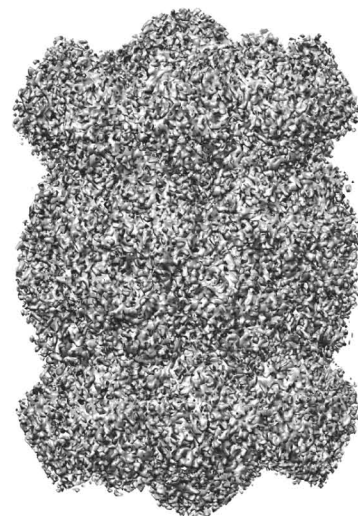
155

D



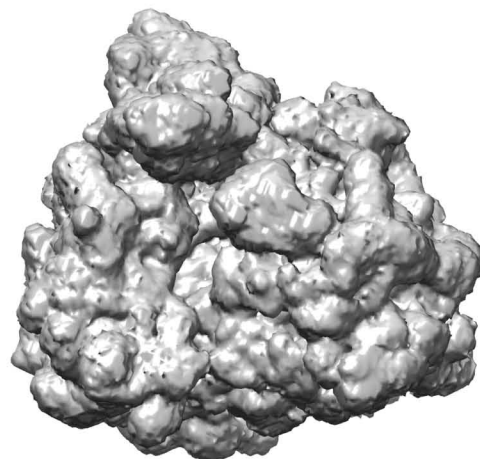
169

E



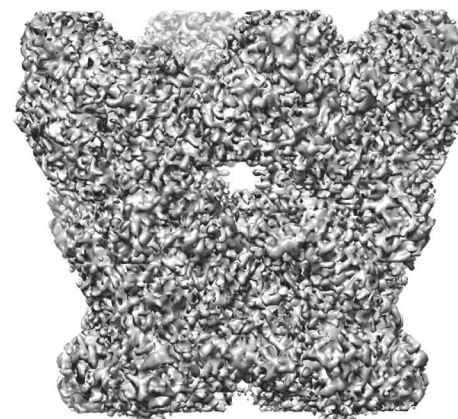
103

F



114

G



146

Figure S24

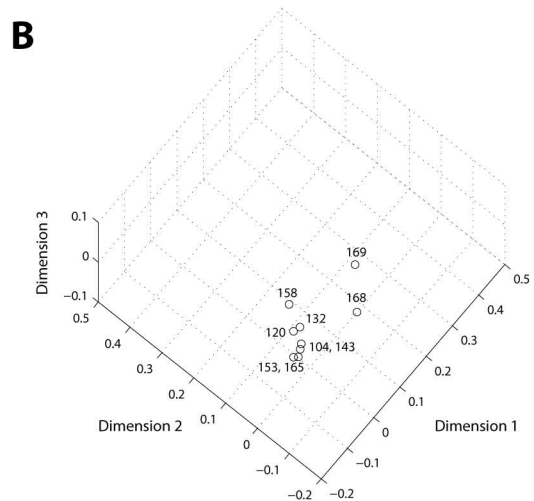
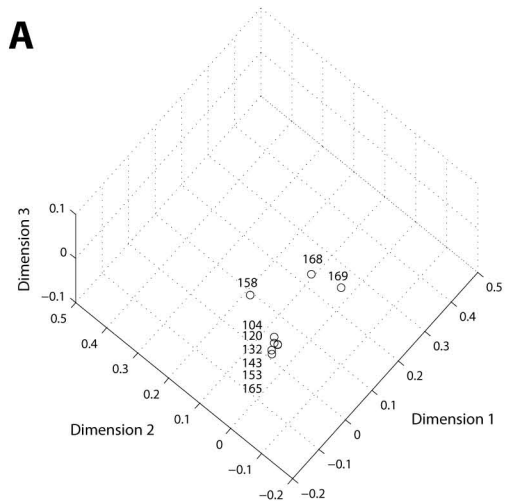


Figure S25

| β-Galactosidase density map | Number of 3D Gaussian functions | Approximation error [%] |
|------------------------------------|--|--------------------------------|
| 106 | 56252 | 13.5 |
| 113 | 86139 | 9.3 |
| 116 | 28203 | 44.4 |
| 134 | 60610 | 16.6 |
| 138 | 83932 | 9.8 |
| 139 | 83716 | 9.8 |
| 154 | 69551 | 12.3 |
| 157 | 50187 | 16.1 |
| 159 | 80876 | 11.8 |
| 160 | 51015 | 15.5 |
| 164 | 77262 | 12.3 |
| 167 | 85703 | 10.1 |

Table S1: Number of 3D Gaussian functions and density-map approximation error obtained using these Gaussian functions for β-Galactosidase density maps. The parameters used for the approximation are provided in **Table 1**.

| Brome Mosaic Virus density map | Number of 3D Gaussian functions | Approximation error [%] |
|---------------------------------------|--|--------------------------------|
| 102 | 11093 | 16.6 |
| 110 | 4977 | 17.9 |
| 136 | 13558 | 10.0 |
| 137 | 12949 | 12.1 |
| 140 | 9389 | 17.4 |
| 142 | 13106 | 13.0 |
| 152 | 3480 | 16.3 |

Table S2: Number of 3D Gaussian functions and density-map approximation error obtained using these Gaussian functions for Brome Mosaic Virus density maps. The parameters used for the approximation are provided in **Table 1**.

| Apo-ferritin density map | Number of 3D Gaussian functions | Approximation error [%] |
|---------------------------------|--|--------------------------------|
| 112 | 3348 | 9.9 |
| 118 | 4230 | 9.8 |
| 121 | 4019 | 10.0 |
| 122 | 7206 | 13.3 |
| 124 | 4943 | 9.7 |
| 147 | 5628 | 9.9 |
| 155 | 3621 | 21.0 |
| 166 | 5076 | 9.9 |

Table S3: Number of 3D Gaussian functions and density-map approximation error obtained using these Gaussian functions for Apo-ferritin density maps. The parameters used for the approximation are provided in **Table 1**.

| GroEL <i>in silico</i> density map | Number of 3D Gaussian functions | Approximation error [%] |
|---------------------------------------|---------------------------------|----------------------------|
| 104 | 7023 | 10.0 |
| 120 | 7425 | 10.3 |
| 132 | 7292 | 11.6 |
| 143 | 6315 | 9.9 |
| 153 | 6508 | 10.0 |
| 158 | 4832 | 22.7 |
| 165 | 6298 | 9.8 |
| 168 | 3378 | 30.0 |
| 169 | 2452 | 44.7 |

Table S4: Number of 3D Gaussian functions and density-map approximation error obtained using these Gaussian functions for GroEL *in silico* density maps. The parameters used for the approximation are provided in **Table 1**.

| T20S proteasome density map | Number of 3D Gaussian functions | Approximation error [%] |
|--------------------------------|---------------------------------|----------------------------|
| 103 | 32503 | 62.5 |
| 107 | 66985 | 23.8 |
| 108 | 50578 | 37.2 |
| 130 | 77555 | 21.3 |
| 131 | 77763 | 21.5 |
| 141 | 88315 | 18.1 |
| 144 | 87594 | 14.8 |
| 145 | 86520 | 14.9 |
| 162 | 72821 | 21.6 |

Table S5: Number of 3D Gaussian functions and density-map approximation error obtained using these Gaussian functions for T20S proteasome density maps. The parameters used for the approximation are provided in **Table 1**.

| 80S ribosome density map | Number of 3D Gaussian functions | Approximation error [%] |
|-----------------------------|---------------------------------|----------------------------|
| 111 | 11092 | 9.3 |
| 114 | 16137 | 8.3 |
| 119 | 7628 | 13.3 |
| 123 | 16057 | 7.7 |
| 125 | 16883 | 6.8 |
| 126 | 16184 | 7.3 |
| 127 | 17210 | 6.3 |
| 128 | 17751 | 6.0 |
| 129 | 19981 | 5.0 |
| 148 | 18744 | 5.3 |
| 149 | 17012 | 6.8 |
| 150 | 16566 | 7.2 |
| 151 | 16083 | 7.7 |

Table S6: Number of 3D Gaussian functions and density-map approximation error obtained using these Gaussian functions for 80S ribosome density maps. The parameters used for the approximation are provided in **Table 1**.

| TRPV1 channel density map | Number of 3D Gaussian functions | Approximation error [%] |
|---------------------------|---------------------------------|-------------------------|
| 101 | 24209 | 34.1 |
| 115 | 16031 | 61.3 |
| 133 | 26915 | 27.0 |
| 135 | 26204 | 28.7 |
| 146 | 6109 | 303.4 |
| 156 | 12755 | 77.2 |
| 161 | 10600 | 96.7 |
| 163 | 13141 | 661.8 |

Table S7: Number of 3D Gaussian functions and density-map approximation error obtained using these Gaussian functions for TRPV1 channel density maps. The parameters used for the approximation are provided in **Table 1**.

| GroEL density map | Gaussian standard deviation: 1.5 voxel Target approximation error: 10 % | | Gaussian standard deviation: 1 voxel Target approximation error: 5 % | |
|-------------------|--|-------------------------|---|-------------------------|
| | Number of Gaussian functions | Approximation error [%] | Number of Gaussian functions | Approximation error [%] |
| 104 | 7023 | 10.0 | 31855 | 8.5 |
| 120 | 7425 | 10.3 | 31586 | 8.9 |
| 132 | 7292 | 11.6 | 35758 | 8.4 |
| 143 | 6315 | 9.9 | 37124 | 7.6 |
| 153 | 6508 | 10.0 | 37666 | 7.7 |
| 158 | 4832 | 22.7 | 32025 | 10.7 |
| 165 | 6298 | 9.8 | 44398 | 6.8 |
| 168 | 3378 | 30.0 | 24585 | 13.4 |
| 169 | 2452 | 44.7 | 11028 | 30.9 |

Table S8: Number of 3D Gaussian functions and density-map approximation error obtained using these Gaussian functions for GroEL density maps for the following two sets of values of the Gaussian-function standard deviation and the target approximation error: 1) 1.5 voxel and 10 % (the parameters used to get the results shown in **Table S4**); and 2) 1 voxel and 5 % (the Gaussian-function standard deviation and the target approximation error, respectively).

| | Gaussian-based EM map approximation | | Non-Gaussian-based EM map approximation | |
|------------|---------------------------------------|---------------------------------------|---|---------------------------------------|
| | Approximation error for mdscale in 3D | Approximation error for mdscale in 2D | Approximation error for mdscale in 3D | Approximation error for mdscale in 2D |
| BetaGal | 0.034 | 0.042 | 0.039 | 0.462 |
| BMV | 0.034 | 0.079 | 0.028 | 0.028 |
| Ferritin | 0.008 | 0.010 | 0.012 | 0.011 |
| GroEL | 0.009 | 0.026 | 0.013 | 0.020 |
| Proteasome | 0.009 | 0.013 | 0.036 | 0.037 |
| Ribosome | 0.022 | 0.027 | 0.013 | 0.017 |
| TRPVI | 0.013 | 0.044 | 0.012 | 0.034 |

Table S9: Approximation error for mdscale in 3D and in 2D with and without Gaussian-based EM density map approximation. The mdscale approximation error is determined by the root-mean-square error of the distances among maps in 3D or 2D with respect to these distances in the original space of higher dimension.

| | 106 | 113 | 116 | 134 | 138 | 139 | 154 | 157 | 159 | 160 | 164 | 167 |
|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| 106 | 0 | 0.24 | 0.34 | 0.09 | 0.12 | 0.12 | 0.09 | 0.12 | 0.41 | 0.13 | 0.41 | 0.41 |
| 113 | | 0 | 0.35 | 0.25 | 0.10 | 0.10 | 0.23 | 0.24 | 0.21 | 0.22 | 0.21 | 0.21 |
| 116 | | | 0 | 0.35 | 0.31 | 0.31 | 0.34 | 0.34 | 0.46 | 0.33 | 0.46 | 0.42 |
| 134 | | | | 0 | 0.12 | 0.12 | 0.11 | 0.14 | 0.39 | 0.13 | 0.39 | 0.38 |
| 138 | | | | | 0 | 0.04 | 0.11 | 0.14 | 0.26 | 0.13 | 0.26 | 0.25 |
| 139 | | | | | | 0 | 0.11 | 0.14 | 0.26 | 0.12 | 0.26 | 0.26 |
| 154 | | | | | | | 0 | 0.14 | 0.38 | 0.14 | 0.38 | 0.39 |
| 157 | | | | | | | | 0 | 0.40 | 0.08 | 0.40 | 0.35 |
| 159 | | | | | | | | | 0 | 0.40 | 0.08 | 0.15 |
| 160 | | | | | | | | | | 0 | 0.37 | 0.32 |
| 164 | | | | | | | | | | | 0 | 0.15 |
| 167 | | | | | | | | | | | | 0 |

Table S10: Distances among Gaussian-based approximations of β -Galactosidase maps before projecting the maps to 3D space. Map index is in bold. Only the asymmetric part of the table is shown.

| | 106 | 113 | 116 | 134 | 138 | 139 | 154 | 157 | 159 | 160 | 164 | 167 |
|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| 106 | 0 | 0.22 | 0.27 | 0.05 | 0.13 | 0.13 | 0.05 | 0.13 | 0.37 | 0.15 | 0.37 | 0.37 |
| 113 | | 0 | 0.31 | 0.21 | 0.09 | 0.09 | 0.19 | 0.21 | 0.17 | 0.17 | 0.17 | 0.17 |
| 116 | | | 0 | 0.32 | 0.26 | 0.26 | 0.25 | 0.25 | 0.46 | 0.26 | 0.46 | 0.42 |
| 134 | | | | 0 | 0.13 | 0.13 | 0.09 | 0.15 | 0.35 | 0.16 | 0.35 | 0.35 |
| 138 | | | | | 0 | 0.00 | 0.10 | 0.16 | 0.26 | 0.13 | 0.26 | 0.26 |
| 139 | | | | | | 0 | 0.09 | 0.16 | 0.26 | 0.13 | 0.26 | 0.26 |
| 154 | | | | | | | 0 | 0.16 | 0.35 | 0.16 | 0.35 | 0.35 |
| 157 | | | | | | | | 0 | 0.37 | 0.05 | 0.37 | 0.31 |
| 159 | | | | | | | | | 0 | 0.32 | 0.00 | 0.15 |
| 160 | | | | | | | | | | 0 | 0.32 | 0.26 |
| 164 | | | | | | | | | | | 0 | 0.15 |
| 167 | | | | | | | | | | | | 0 |

Table S11: Distances among Gaussian-based approximations of β -Galactosidase maps after projecting the maps to 3D space. Map index is in bold. Only the asymmetric part of the table is shown.

| | 106 | 113 | 116 | 134 | 138 | 139 | 154 | 157 | 159 | 160 | 164 | 167 |
|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| 106 | 0 | 0.23 | 0.46 | 0.11 | 0.10 | 0.10 | 0.08 | 0.13 | 0.39 | 0.13 | 0.39 | 0.41 |
| 113 | | 0 | 0.47 | 0.26 | 0.06 | 0.07 | 0.23 | 0.23 | 0.15 | 0.21 | 0.15 | 0.17 |
| 116 | | | 0 | 0.48 | 0.43 | 0.43 | 0.46 | 0.45 | 0.55 | 0.44 | 0.55 | 0.51 |
| 134 | | | | 0 | 0.13 | 0.13 | 0.13 | 0.18 | 0.39 | 0.17 | 0.39 | 0.40 |
| 138 | | | | | 0 | 0.01 | 0.10 | 0.13 | 0.21 | 0.11 | 0.21 | 0.23 |
| 139 | | | | | | 0 | 0.09 | 0.13 | 0.21 | 0.11 | 0.21 | 0.24 |
| 154 | | | | | | | 0 | 0.15 | 0.37 | 0.15 | 0.37 | 0.40 |
| 157 | | | | | | | | 0 | 0.38 | 0.07 | 0.38 | 0.32 |
| 159 | | | | | | | | | 0 | 0.34 | 0.00 | 0.09 |
| 160 | | | | | | | | | | 0 | 0.34 | 0.28 |
| 164 | | | | | | | | | | | 0 | 0.09 |
| 167 | | | | | | | | | | | | 0 |

Table S12: Distances among original β -Galactosidase maps before projecting the maps to 3D space. Map index is in bold. Only the asymmetric part of the table is shown.

| | 106 | 113 | 116 | 134 | 138 | 139 | 154 | 157 | 159 | 160 | 164 | 167 |
|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| 106 | 0 | 0.18 | 0.42 | 0.15 | 0.15 | 0.15 | 0.02 | 0.17 | 0.34 | 0.18 | 0.34 | 0.36 |
| 113 | | 0 | 0.43 | 0.21 | 0.03 | 0.03 | 0.18 | 0.19 | 0.17 | 0.18 | 0.17 | 0.19 |
| 116 | | | 0 | 0.55 | 0.43 | 0.43 | 0.41 | 0.42 | 0.54 | 0.42 | 0.54 | 0.54 |
| 134 | | | | 0 | 0.18 | 0.18 | 0.16 | 0.18 | 0.33 | 0.18 | 0.33 | 0.33 |
| 138 | | | | | 0 | 0.00 | 0.15 | 0.17 | 0.19 | 0.16 | 0.19 | 0.21 |
| 139 | | | | | | 0 | 0.15 | 0.17 | 0.19 | 0.16 | 0.19 | 0.21 |
| 154 | | | | | | | 0 | 0.19 | 0.33 | 0.19 | 0.33 | 0.37 |
| 157 | | | | | | | | 0 | 0.34 | 0.02 | 0.34 | 0.29 |
| 159 | | | | | | | | | 0 | 0.32 | 0.00 | 0.15 |
| 160 | | | | | | | | | | 0 | 0.32 | 0.27 |
| 164 | | | | | | | | | | | 0 | 0.15 |
| 167 | | | | | | | | | | | | 0 |

Table S13: Distances among original β -Galactosidase maps after projecting the maps to 3D space. Map index is in bold. Only the asymmetric part of the table is shown.

| | 102 | 110 | 136 | 137 | 140 | 142 | 152 |
|------------|------------|------------|------------|------------|------------|------------|------------|
| 102 | 0 | 0.49 | 0.23 | 0.17 | 0.16 | 0.16 | 0.54 |
| 110 | | 0 | 0.55 | 0.47 | 0.48 | 0.49 | 0.41 |
| 136 | | | 0 | 0.12 | 0.25 | 0.14 | 0.61 |
| 137 | | | | 0 | 0.19 | 0.09 | 0.55 |
| 140 | | | | | 0 | 0.18 | 0.52 |
| 142 | | | | | | 0 | 0.57 |
| 152 | | | | | | | 0 |

Table S14: Distances among Gaussian-based approximations of Brome Mosaic Virus maps before projecting the maps to 3D space. Map index is in bold. Only the asymmetric part of the table is shown.

| | 102 | 110 | 136 | 137 | 140 | 142 | 152 |
|------------|------------|------------|------------|------------|------------|------------|------------|
| 102 | 0 | 0.52 | 0.19 | 0.12 | 0.08 | 0.11 | 0.56 |
| 110 | | 0 | 0.56 | 0.51 | 0.51 | 0.52 | 0.43 |
| 136 | | | 0 | 0.08 | 0.27 | 0.08 | 0.61 |
| 137 | | | | 0 | 0.19 | 0.02 | 0.56 |
| 140 | | | | | 0 | 0.19 | 0.52 |
| 142 | | | | | | 0 | 0.58 |
| 152 | | | | | | | 0 |

Table S15: Distances among Gaussian-based approximations of Brome Mosaic Virus maps after projecting the maps to 3D space. Map index is in bold. Only the asymmetric part of the table is shown.

| | 102 | 110 | 136 | 137 | 140 | 142 | 152 |
|------------|------------|------------|------------|------------|------------|------------|------------|
| 102 | 0 | 0.19 | 0.10 | 0.05 | 0.02 | 0.03 | 0.42 |
| 110 | | 0 | 0.35 | 0.26 | 0.20 | 0.24 | 0.47 |
| 136 | | | 0 | 0.03 | 0.08 | 0.03 | 0.49 |
| 137 | | | | 0 | 0.03 | 0.01 | 0.42 |
| 140 | | | | | 0 | 0.01 | 0.42 |
| 142 | | | | | | 0 | 0.43 |
| 152 | | | | | | | 0 |

Table S16: Distances among original Brome Mosaic Virus maps before projecting the maps to 3D space. Map index is in bold. Only the asymmetric part of the table is shown.

| | 102 | 110 | 136 | 137 | 140 | 142 | 152 |
|------------|------------|------------|------------|------------|------------|------------|------------|
| 102 | 0 | 0.20 | 0.15 | 0.10 | 0.04 | 0.07 | 0.43 |
| 110 | | 0 | 0.35 | 0.29 | 0.22 | 0.26 | 0.48 |
| 136 | | | 0 | 0.07 | 0.13 | 0.09 | 0.49 |
| 137 | | | | 0 | 0.07 | 0.03 | 0.43 |
| 140 | | | | | 0 | 0.04 | 0.43 |
| 142 | | | | | | 0 | 0.44 |
| 152 | | | | | | | 0 |

Table S17: Distances among original Brome Mosaic Virus maps after projecting the maps to 3D space. Map index is in bold. Only the asymmetric part of the table is shown.

| | 112 | 118 | 121 | 122 | 124 | 147 | 155 | 166 |
|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| 112 | 0 | 0.06 | 0.05 | 0.05 | 0.05 | 0.15 | 0.08 | 0.08 |
| 118 | | 0 | 0.05 | 0.07 | 0.06 | 0.10 | 0.12 | 0.04 |
| 121 | | | 0 | 0.06 | 0.07 | 0.12 | 0.10 | 0.06 |
| 122 | | | | 0 | 0.10 | 0.17 | 0.08 | 0.09 |
| 124 | | | | | 0 | 0.08 | 0.15 | 0.06 |
| 147 | | | | | | 0 | 0.23 | 0.09 |
| 155 | | | | | | | 0 | 0.16 |
| 166 | | | | | | | | 0 |

Table S18: Distances among Gaussian-based approximations of Apo-ferritin maps before projecting the maps to 3D space. Map index is in bold. Only the asymmetric part of the table is shown.

| | 112 | 118 | 121 | 122 | 124 | 147 | 155 | 166 |
|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| 112 | 0 | 0.07 | 0.04 | 0.06 | 0.07 | 0.15 | 0.09 | 0.09 |
| 118 | | 0 | 0.04 | 0.07 | 0.07 | 0.10 | 0.14 | 0.03 |
| 121 | | | 0 | 0.07 | 0.08 | 0.13 | 0.10 | 0.07 |
| 122 | | | | 0 | 0.10 | 0.17 | 0.09 | 0.09 |
| 124 | | | | | 0 | 0.09 | 0.16 | 0.07 |
| 147 | | | | | | 0 | 0.23 | 0.09 |
| 155 | | | | | | | 0 | 0.16 |
| 166 | | | | | | | | 0 |

Table S19: Distances among Gaussian-based approximations of Apo-ferritin maps after projecting the maps to 3D space. Map index is in bold. Only the asymmetric part of the table is shown.

| | 112 | 118 | 121 | 122 | 124 | 147 | 155 | 166 |
|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| 112 | 0 | 0.03 | 0.01 | 0.08 | 0.06 | 0.12 | 0.10 | 0.04 |
| 118 | | 0 | 0.02 | 0.09 | 0.02 | 0.06 | 0.15 | 0.00 |
| 121 | | | 0 | 0.08 | 0.05 | 0.10 | 0.11 | 0.03 |
| 122 | | | | 0 | 0.11 | 0.18 | 0.17 | 0.10 |
| 124 | | | | | 0 | 0.05 | 0.17 | 0.02 |
| 147 | | | | | | 0 | 0.23 | 0.06 |
| 155 | | | | | | | 0 | 0.16 |
| 166 | | | | | | | | 0 |

Table S20: Distances among original Apo-ferritin maps before projecting the maps to 3D space. Map index is in bold. Only the asymmetric part of the table is shown.

| | 112 | 118 | 121 | 122 | 124 | 147 | 155 | 166 |
|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| 112 | 0 | 0.05 | 0.02 | 0.09 | 0.07 | 0.13 | 0.12 | 0.05 |
| 118 | | 0 | 0.04 | 0.10 | 0.03 | 0.08 | 0.16 | 0.01 |
| 121 | | | 0 | 0.10 | 0.06 | 0.12 | 0.12 | 0.04 |
| 122 | | | | 0 | 0.12 | 0.18 | 0.17 | 0.11 |
| 124 | | | | | 0 | 0.06 | 0.17 | 0.02 |
| 147 | | | | | | 0 | 0.23 | 0.08 |
| 155 | | | | | | | 0 | 0.17 |
| 166 | | | | | | | | 0 |

Table S21: Distances among original Apo-ferritin maps after projecting the maps to 3D space. Map index is in bold. Only the asymmetric part of the table is shown.

| | 104 | 120 | 132 | 143 | 153 | 158 | 165 | 168 | 169 |
|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| 104 | 0 | 0.03 | 0.03 | 0.01 | 0.02 | 0.13 | 0.02 | 0.15 | 0.24 |
| 120 | | 0 | 0.02 | 0.04 | 0.04 | 0.14 | 0.05 | 0.18 | 0.26 |
| 132 | | | 0 | 0.03 | 0.05 | 0.11 | 0.05 | 0.14 | 0.22 |
| 143 | | | | 0 | 0.02 | 0.12 | 0.02 | 0.15 | 0.24 |
| 153 | | | | | 0 | 0.15 | 0.02 | 0.18 | 0.27 |
| 158 | | | | | | 0 | 0.13 | 0.17 | 0.23 |
| 165 | | | | | | | 0 | 0.16 | 0.25 |
| 168 | | | | | | | | 0 | 0.16 |
| 169 | | | | | | | | | 0 |

Table S22: Distances among Gaussian-based approximations of GroEL maps before projecting the maps to 3D space. Map index is in bold. Only the asymmetric part of the table is shown.

| | 104 | 120 | 132 | 143 | 153 | 158 | 165 | 168 | 169 |
|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| 104 | 0 | 0.02 | 0.02 | 0.00 | 0.02 | 0.14 | 0.02 | 0.15 | 0.24 |
| 120 | | 0 | 0.02 | 0.02 | 0.02 | 0.14 | 0.04 | 0.17 | 0.25 |
| 132 | | | 0 | 0.02 | 0.04 | 0.12 | 0.04 | 0.15 | 0.23 |
| 143 | | | | 0 | 0.02 | 0.14 | 0.02 | 0.15 | 0.24 |
| 153 | | | | | 0 | 0.15 | 0.02 | 0.17 | 0.27 |
| 158 | | | | | | 0 | 0.14 | 0.16 | 0.23 |
| 165 | | | | | | | 0 | 0.15 | 0.25 |
| 168 | | | | | | | | 0 | 0.15 |
| 169 | | | | | | | | | 0 |

Table S23: Distances among Gaussian-based approximations of GroEL maps after projecting the maps to 3D space. Map index is in bold. Only the asymmetric part of the table is shown.

| | 104 | 120 | 132 | 143 | 153 | 158 | 165 | 168 | 169 |
|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| 104 | 0 | 0.03 | 0.02 | 0.00 | 0.01 | 0.09 | 0.01 | 0.19 | 0.33 |
| 120 | | 0 | 0.01 | 0.03 | 0.03 | 0.10 | 0.04 | 0.23 | 0.37 |
| 132 | | | 0 | 0.02 | 0.03 | 0.07 | 0.03 | 0.19 | 0.32 |
| 143 | | | | 0 | 0.01 | 0.09 | 0.00 | 0.18 | 0.33 |
| 153 | | | | | 0 | 0.11 | 0.01 | 0.22 | 0.36 |
| 158 | | | | | | 0 | 0.09 | 0.18 | 0.26 |
| 165 | | | | | | | 0 | 0.19 | 0.33 |
| 168 | | | | | | | | 0 | 0.16 |
| 169 | | | | | | | | | 0 |

Table S24: Distances among original GroEL maps before projecting the maps to 3D space. Map index is in bold. Only the asymmetric part of the table is shown.

| | 104 | 120 | 132 | 143 | 153 | 158 | 165 | 168 | 169 |
|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| 104 | 0 | 0.03 | 0.03 | 0.02 | 0.03 | 0.11 | 0.03 | 0.20 | 0.35 |
| 120 | | 0 | 0.03 | 0.04 | 0.04 | 0.11 | 0.05 | 0.23 | 0.37 |
| 132 | | | 0 | 0.03 | 0.04 | 0.09 | 0.04 | 0.21 | 0.34 |
| 143 | | | | 0 | 0.02 | 0.11 | 0.01 | 0.20 | 0.35 |
| 153 | | | | | 0 | 0.12 | 0.02 | 0.22 | 0.37 |
| 158 | | | | | | 0 | 0.11 | 0.18 | 0.27 |
| 165 | | | | | | | 0 | 0.21 | 0.35 |
| 168 | | | | | | | | 0 | 0.18 |
| 169 | | | | | | | | | 0 |

Table S25: Distances among original GroEL maps after projecting the maps to 3D space. Map index is in bold. Only the asymmetric part of the table is shown.

| | 103 | 107 | 108 | 130 | 131 | 141 | 144 | 145 | 162 |
|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| 103 | 0 | 0.13 | 0.14 | 0.22 | 0.22 | 0.16 | 0.23 | 0.22 | 0.16 |
| 107 | | 0 | 0.07 | 0.10 | 0.10 | 0.06 | 0.09 | 0.09 | 0.05 |
| 108 | | | 0 | 0.14 | 0.14 | 0.09 | 0.14 | 0.13 | 0.08 |
| 130 | | | | 0 | 0.02 | 0.09 | 0.09 | 0.09 | 0.10 |
| 131 | | | | | 0 | 0.09 | 0.09 | 0.09 | 0.10 |
| 141 | | | | | | 0 | 0.07 | 0.06 | 0.06 |
| 144 | | | | | | | 0 | 0.03 | 0.08 |
| 145 | | | | | | | | 0 | 0.08 |
| 162 | | | | | | | | | 0 |

Table S26: Distances among Gaussian-based approximations of T20S proteasome maps before projecting the maps to 3D space. Map index is in bold. Only the asymmetric part of the table is shown.

| | 103 | 107 | 108 | 130 | 131 | 141 | 144 | 145 | 162 |
|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| 103 | 0 | 0.14 | 0.14 | 0.23 | 0.23 | 0.16 | 0.23 | 0.22 | 0.16 |
| 107 | | 0 | 0.07 | 0.10 | 0.10 | 0.03 | 0.09 | 0.09 | 0.03 |
| 108 | | | 0 | 0.15 | 0.15 | 0.10 | 0.14 | 0.14 | 0.07 |
| 130 | | | | 0 | 0.00 | 0.09 | 0.09 | 0.09 | 0.10 |
| 131 | | | | | 0 | 0.09 | 0.09 | 0.09 | 0.10 |
| 141 | | | | | | 0 | 0.07 | 0.07 | 0.03 |
| 144 | | | | | | | 0 | 0.00 | 0.07 |
| 145 | | | | | | | | 0 | 0.07 |
| 162 | | | | | | | | | 0 |

Table S27: Distances among Gaussian-based approximations of T20S proteasome maps after projecting the maps to 3D space. Map index is in bold. Only the asymmetric part of the table is shown.

| | 103 | 107 | 108 | 130 | 131 | 141 | 144 | 145 | 162 |
|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| 103 | 0 | 0.13 | 0.12 | 0.28 | 0.28 | 0.19 | 0.32 | 0.31 | 0.19 |
| 107 | | 0 | 0.05 | 0.10 | 0.10 | 0.03 | 0.07 | 0.07 | 0.03 |
| 108 | | | 0 | 0.16 | 0.16 | 0.08 | 0.13 | 0.13 | 0.07 |
| 130 | | | | 0 | 0.00 | 0.09 | 0.09 | 0.09 | 0.11 |
| 131 | | | | | 0 | 0.09 | 0.09 | 0.09 | 0.11 |
| 141 | | | | | | 0 | 0.05 | 0.04 | 0.05 |
| 144 | | | | | | | 0 | 0.00 | 0.07 |
| 145 | | | | | | | | 0 | 0.06 |
| 162 | | | | | | | | | 0 |

Table S28: Distances among original T20S proteasome maps before projecting the maps to 3D space. Map index is in bold. Only the asymmetric part of the table is shown.

| | 103 | 107 | 108 | 130 | 131 | 141 | 144 | 145 | 162 |
|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| 103 | 0 | 0.20 | 0.13 | 0.30 | 0.30 | 0.23 | 0.32 | 0.31 | 0.23 |
| 107 | | 0 | 0.09 | 0.13 | 0.13 | 0.04 | 0.12 | 0.11 | 0.06 |
| 108 | | | 0 | 0.22 | 0.22 | 0.12 | 0.20 | 0.18 | 0.12 |
| 130 | | | | 0 | 0.00 | 0.12 | 0.12 | 0.12 | 0.13 |
| 131 | | | | | 0 | 0.12 | 0.12 | 0.12 | 0.13 |
| 141 | | | | | | 0 | 0.09 | 0.08 | 0.08 |
| 144 | | | | | | | 0 | 0.02 | 0.10 |
| 145 | | | | | | | | 0 | 0.09 |
| 162 | | | | | | | | | 0 |

Table S29: Distances among original T20S proteasome maps after projecting the maps to 3D space. Map index is in bold. Only the asymmetric part of the table is shown.

| | 111 | 114 | 119 | 123 | 125 | 126 | 127 | 128 | 129 | 148 | 149 | 150 | 151 |
|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| 111 | 0 | 0.22 | 0.10 | 0.23 | 0.22 | 0.22 | 0.24 | 0.24 | 0.31 | 0.25 | 0.26 | 0.25 | 0.23 |
| 114 | | 0 | 0.27 | 0.04 | 0.04 | 0.05 | 0.05 | 0.05 | 0.12 | 0.05 | 0.06 | 0.06 | 0.04 |
| 119 | | | 0 | 0.28 | 0.28 | 0.28 | 0.30 | 0.30 | 0.37 | 0.30 | 0.31 | 0.30 | 0.28 |
| 123 | | | | 0 | 0.04 | 0.04 | 0.04 | 0.04 | 0.10 | 0.05 | 0.05 | 0.05 | 0.04 |
| 125 | | | | | 0 | 0.04 | 0.04 | 0.04 | 0.10 | 0.04 | 0.05 | 0.05 | 0.04 |
| 126 | | | | | | 0 | 0.04 | 0.05 | 0.11 | 0.05 | 0.05 | 0.05 | 0.04 |
| 127 | | | | | | | 0 | 0.04 | 0.10 | 0.04 | 0.05 | 0.05 | 0.04 |
| 128 | | | | | | | | 0 | 0.09 | 0.04 | 0.05 | 0.05 | 0.05 |
| 129 | | | | | | | | | 0 | 0.08 | 0.09 | 0.10 | 0.11 |
| 148 | | | | | | | | | | 0 | 0.05 | 0.06 | 0.05 |
| 149 | | | | | | | | | | | 0 | 0.03 | 0.05 |
| 150 | | | | | | | | | | | | 0 | 0.05 |
| 151 | | | | | | | | | | | | | 0 |

Table S30: Distances among Gaussian-based approximations of 80S ribosome maps before projecting the maps to 3D space. Map index is in bold. Only the asymmetric part of the table is shown.

| | 111 | 114 | 119 | 123 | 125 | 126 | 127 | 128 | 129 | 148 | 149 | 150 | 151 |
|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| 111 | 0 | 0.22 | 0.10 | 0.23 | 0.23 | 0.23 | 0.23 | 0.24 | 0.32 | 0.24 | 0.25 | 0.24 | 0.23 |
| 114 | | 0 | 0.27 | 0.01 | 0.01 | 0.02 | 0.02 | 0.02 | 0.11 | 0.03 | 0.03 | 0.03 | 0.02 |
| 119 | | | 0 | 0.28 | 0.28 | 0.28 | 0.29 | 0.29 | 0.37 | 0.29 | 0.30 | 0.29 | 0.28 |
| 123 | | | | 0 | 0.01 | 0.01 | 0.01 | 0.01 | 0.10 | 0.02 | 0.02 | 0.02 | 0.01 |
| 125 | | | | | 0 | 0.01 | 0.01 | 0.01 | 0.10 | 0.01 | 0.02 | 0.03 | 0.01 |
| 126 | | | | | | 0 | 0.01 | 0.02 | 0.10 | 0.02 | 0.02 | 0.02 | 0.01 |
| 127 | | | | | | | 0 | 0.01 | 0.09 | 0.01 | 0.02 | 0.02 | 0.01 |
| 128 | | | | | | | | 0 | 0.09 | 0.01 | 0.02 | 0.02 | 0.02 |
| 129 | | | | | | | | | 0 | 0.08 | 0.09 | 0.10 | 0.10 |
| 148 | | | | | | | | | | 0 | 0.03 | 0.03 | 0.02 |
| 149 | | | | | | | | | | | 0 | 0.01 | 0.02 |
| 150 | | | | | | | | | | | | 0 | 0.02 |
| 151 | | | | | | | | | | | | | 0 |

Table S31: Distances among Gaussian-based approximations of 80S ribosome maps after projecting the maps to 3D space. Map index is in bold. Only the asymmetric part of the table is shown.

| | 111 | 114 | 119 | 123 | 125 | 126 | 127 | 128 | 129 | 148 | 149 | 150 | 151 |
|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| 111 | 0 | 0.13 | 0.04 | 0.13 | 0.11 | 0.12 | 0.13 | 0.13 | 0.20 | 0.13 | 0.14 | 0.14 | 0.13 |
| 114 | | 0 | 0.14 | 0.01 | 0.01 | 0.01 | 0.02 | 0.02 | 0.09 | 0.02 | 0.02 | 0.02 | 0.01 |
| 119 | | | 0 | 0.14 | 0.13 | 0.14 | 0.15 | 0.15 | 0.25 | 0.16 | 0.16 | 0.15 | 0.14 |
| 123 | | | | 0 | 0.01 | 0.00 | 0.01 | 0.01 | 0.08 | 0.02 | 0.01 | 0.01 | 0.00 |
| 125 | | | | | 0 | 0.01 | 0.01 | 0.01 | 0.07 | 0.02 | 0.02 | 0.02 | 0.01 |
| 126 | | | | | | 0 | 0.01 | 0.01 | 0.08 | 0.02 | 0.01 | 0.01 | 0.01 |
| 127 | | | | | | | 0 | 0.01 | 0.07 | 0.01 | 0.02 | 0.02 | 0.01 |
| 128 | | | | | | | | 0 | 0.06 | 0.01 | 0.02 | 0.02 | 0.01 |
| 129 | | | | | | | | | 0 | 0.06 | 0.07 | 0.07 | 0.08 |
| 148 | | | | | | | | | | 0 | 0.02 | 0.03 | 0.02 |
| 149 | | | | | | | | | | | 0 | 0.00 | 0.01 |
| 150 | | | | | | | | | | | | 0 | 0.01 |
| 151 | | | | | | | | | | | | | 0 |

Table S32: Distances among original 80S ribosome maps before projecting the maps to 3D space. Map index is in bold. Only the asymmetric part of the table is shown.

| | 111 | 114 | 119 | 123 | 125 | 126 | 127 | 128 | 129 | 148 | 149 | 150 | 151 |
|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| 111 | 0 | 0.14 | 0.07 | 0.14 | 0.14 | 0.14 | 0.14 | 0.14 | 0.23 | 0.14 | 0.15 | 0.15 | 0.14 |
| 114 | | 0 | 0.15 | 0.01 | 0.01 | 0.01 | 0.01 | 0.02 | 0.11 | 0.02 | 0.01 | 0.01 | 0.01 |
| 119 | | | 0 | 0.15 | 0.15 | 0.15 | 0.15 | 0.16 | 0.25 | 0.16 | 0.16 | 0.16 | 0.15 |
| 123 | | | | 0 | 0.00 | 0.00 | 0.01 | 0.01 | 0.10 | 0.01 | 0.01 | 0.01 | 0.00 |
| 125 | | | | | 0 | 0.00 | 0.01 | 0.01 | 0.10 | 0.01 | 0.01 | 0.01 | 0.01 |
| 126 | | | | | | 0 | 0.01 | 0.01 | 0.10 | 0.01 | 0.01 | 0.01 | 0.00 |
| 127 | | | | | | | 0 | 0.01 | 0.10 | 0.01 | 0.01 | 0.01 | 0.01 |
| 128 | | | | | | | | 0 | 0.10 | 0.01 | 0.01 | 0.01 | 0.01 |
| 129 | | | | | | | | | 0 | 0.10 | 0.10 | 0.10 | 0.10 |
| 148 | | | | | | | | | | 0 | 0.02 | 0.02 | 0.02 |
| 149 | | | | | | | | | | | 0 | 0.00 | 0.01 |
| 150 | | | | | | | | | | | | 0 | 0.01 |
| 151 | | | | | | | | | | | | | 0 |

Table S33: Distances among original 80S ribosome maps after projecting the maps to 3D space. Map index is in bold. Only the asymmetric part of the table is shown.

| | 101 | 115 | 133 | 135 | 146 | 156 | 161 | 163 |
|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| 101 | 0 | 0.04 | 0.05 | 0.04 | 0.22 | 0.14 | 0.09 | 0.11 |
| 115 | | 0 | 0.07 | 0.07 | 0.22 | 0.13 | 0.08 | 0.11 |
| 133 | | | 0 | 0.01 | 0.21 | 0.15 | 0.09 | 0.08 |
| 135 | | | | 0 | 0.21 | 0.15 | 0.09 | 0.08 |
| 146 | | | | | 0 | 0.29 | 0.20 | 0.22 |
| 156 | | | | | | 0 | 0.13 | 0.16 |
| 161 | | | | | | | 0 | 0.11 |
| 163 | | | | | | | | 0 |

Table S34: Distances among Gaussian-based approximations of TRPV1 channel maps before projecting the maps to 3D space. Map index is in bold. Only the asymmetric part of the table is shown.

| | 101 | 115 | 133 | 135 | 146 | 156 | 161 | 163 |
|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| 101 | 0 | 0.03 | 0.03 | 0.03 | 0.22 | 0.14 | 0.07 | 0.10 |
| 115 | | 0 | 0.05 | 0.05 | 0.22 | 0.12 | 0.05 | 0.10 |
| 133 | | | 0 | 0.00 | 0.21 | 0.15 | 0.07 | 0.07 |
| 135 | | | | 0 | 0.21 | 0.15 | 0.07 | 0.07 |
| 146 | | | | | 0 | 0.29 | 0.18 | 0.22 |
| 156 | | | | | | 0 | 0.12 | 0.15 |
| 161 | | | | | | | 0 | 0.10 |
| 163 | | | | | | | | 0 |

Table S35: Distances among Gaussian-based approximations of TRPV1 channel maps after projecting the maps to 3D space. Map index is in bold. Only the asymmetric part of the table is shown.

| | 101 | 115 | 133 | 135 | 146 | 156 | 161 | 163 |
|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| 101 | 0 | 0.05 | 0.04 | 0.04 | 0.38 | 0.20 | 0.10 | 0.10 |
| 115 | | 0 | 0.10 | 0.08 | 0.38 | 0.19 | 0.09 | 0.13 |
| 133 | | | 0 | 0.01 | 0.37 | 0.22 | 0.11 | 0.07 |
| 135 | | | | 0 | 0.37 | 0.22 | 0.10 | 0.07 |
| 146 | | | | | 0 | 0.46 | 0.37 | 0.40 |
| 156 | | | | | | 0 | 0.17 | 0.22 |
| 161 | | | | | | | 0 | 0.13 |
| 163 | | | | | | | | 0 |

Table S36: Distances among original TRPV1 channel maps before projecting the maps to 3D space. Map index is in bold. Only the asymmetric part of the table is shown.

| | 101 | 115 | 133 | 135 | 146 | 156 | 161 | 163 |
|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| 101 | 0 | 0.04 | 0.04 | 0.03 | 0.38 | 0.20 | 0.07 | 0.08 |
| 115 | | 0 | 0.07 | 0.06 | 0.39 | 0.20 | 0.06 | 0.12 |
| 133 | | | 0 | 0.01 | 0.37 | 0.22 | 0.10 | 0.06 |
| 135 | | | | 0 | 0.37 | 0.21 | 0.08 | 0.06 |
| 146 | | | | | 0 | 0.46 | 0.36 | 0.40 |
| 156 | | | | | | 0 | 0.15 | 0.21 |
| 161 | | | | | | | 0 | 0.12 |
| 163 | | | | | | | | 0 |

Table S37: Distances among original TRPV1 channel maps after projecting the maps to 3D space. Map index is in bold. Only the asymmetric part of the table is shown.

| | 104 | 120 | 132 | 143 | 153 | 158 | 165 | 168 | 169 |
|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| 104 | 0 | 0.03 | 0.03 | 0.02 | 0.02 | 0.12 | 0.02 | 0.16 | 0.23 |
| 120 | | 0 | 0.02 | 0.04 | 0.05 | 0.13 | 0.05 | 0.18 | 0.24 |
| 132 | | | 0 | 0.03 | 0.05 | 0.10 | 0.05 | 0.15 | 0.21 |
| 143 | | | | 0 | 0.02 | 0.12 | 0.01 | 0.16 | 0.23 |
| 153 | | | | | 0 | 0.15 | 0.02 | 0.20 | 0.26 |
| 158 | | | | | | 0 | 0.13 | 0.16 | 0.18 |
| 165 | | | | | | | 0 | 0.17 | 0.24 |
| 168 | | | | | | | | 0 | 0.10 |
| 169 | | | | | | | | | 0 |

Table S38: Distances among Gaussian-based approximations of GroEL maps before projecting the maps to 3D space, for the standard deviation of Gaussian functions of 1 voxel and the target approximation error of 5 %. For comparison with **Table S22** obtained for a larger standard deviation of Gaussian functions (1.5 voxel) and a larger target approximation error (10 %). Map index is in bold. Only the asymmetric part of the table is shown.

| | 104 | 120 | 132 | 143 | 153 | 158 | 165 | 168 | 169 |
|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| 104 | 0 | 0.03 | 0.03 | 0.02 | 0.03 | 0.12 | 0.03 | 0.17 | 0.23 |
| 120 | | 0 | 0.03 | 0.05 | 0.05 | 0.12 | 0.06 | 0.18 | 0.24 |
| 132 | | | 0 | 0.04 | 0.06 | 0.10 | 0.05 | 0.16 | 0.21 |
| 143 | | | | 0 | 0.03 | 0.12 | 0.02 | 0.17 | 0.23 |
| 153 | | | | | 0 | 0.14 | 0.03 | 0.19 | 0.26 |
| 158 | | | | | | 0 | 0.12 | 0.17 | 0.18 |
| 165 | | | | | | | 0 | 0.17 | 0.24 |
| 168 | | | | | | | | 0 | 0.10 |
| 169 | | | | | | | | | 0 |

Table S39: Distances among Gaussian-based approximations of GroEL maps after projecting the maps to 3D space, for the standard deviation of Gaussian functions of 1 voxel and the target approximation error of 5 %. For comparison with **Table S23** obtained for a larger standard deviation of Gaussian functions (1.5 voxel) and a larger target approximation error (10 %). Map index is in bold. Only the asymmetric part of the table is shown.