1
2
3
4
5
6
7 Supplementary Information for
8
9 Sparseness and Smoothness Regularized Imaging for Improving the Resolution
10 of Cryo-EM Single Particle Reconstruction

11
12 Zhenwei Luo[1], Adam A. Campos-Acevedo[2], Longfei Lv[3], Qinghua Wang[2] & Jianpeng Ma[1,2,3,*]
13
14
15 *[1]Department of Bioengineering*
16 *Rice University*
17 *6500 Main Street*
18 *Houston, TX 77030, USA*
19
20 *[2]Verna and Marrs Mclean Department of Biochemistry and Molecular Biology*
21 *Baylor College of Medicine*
22 *One Baylor Plaza, BCM-125*
23 *Houston, TX 77030, USA*
24
25 *[3]Multiscale Research Institute of Complex Systems*
26 *Fudan University*
27 *825 Zhangheng Road, Pudong District*
28 *Shanghai, 201203, China*
29
30
31 Jianpeng Ma
32 Email: jpma@bcm.edu
33
34
35 **This PDF file includes:**
36
37   Supplementary text
38   Figures S1 to S4
39   SI References
40

41   **Supplementary Information Text**

42   **Methods**

43   **Log marginal likelihood**

44   The log marginal likelihood function in cryo-EM refinement problem can be derived as follows. For
45   simplicity, we assume the particles are from a single structure, whereas structural heterogeneity can be
46   easily incorporated into our generative model by treating the class membership of each particle as a
47   hidden variable(1). Since the images collected by cryo-EM are 2D projections of a 3D molecular structure,
48   the Fourier transform of the image has the following relation with the Fourier transform of the 3D
49   molecular structure. Let the Fourier transform of a 3D molecular structure be V, we first arrange the 3D
50   volume V into a vector with L elements. Assume an $N \times N$ image i is formed by rotating the 3D volume V
51   with the Euler angle set ϕ and projecting along with the z axis, and shifting by $[\Delta_x, \Delta_y]$ from the origin,
52   using projection-slice theorem, the Fourier transform of the image i can be expressed as

$$X_{ij} = e^{i\frac{2\pi}{N}(\Delta_x h + \Delta_y k)} \text{CTF}_{ij} \sum_{l=1}^{L} P_{jl}^{\phi} V_l,$$ (Equation S1)

53   where X_{ij} is the jth component of the Fourier transform of the image i whose corresponding 2D index is
54    $[h, k]$, CTF_{ij} is the jth component of the contrast transfer function for the image i, and P_{jl}^{ϕ} is the slice
55   operator which cuts out the plane in the 3D Fourier transform V which is rotated from the xy plane
56   according to the Euler angle set ϕ. We elaborate on the slice operator P^{ϕ} by giving its formal definition.
57   Let the index of a voxel in the 3D Fourier transform V be $[h', k', l']$, and the index of the corresponding
58   pixel of the Fourier transform of the image i is $[h, k]$, the slice operator P^{ϕ} transforms the 3D index to 2D
59   index by the following equation,

$$\begin{pmatrix} h \\ k \\ 0 \end{pmatrix} = R_{\phi} \begin{pmatrix} h' \\ k' \\ l' \end{pmatrix},$$ (Equation S2)

60   where R_{ϕ} is a rotation matrix parameterized by Euler angles ϕ. Furthermore, suppose the Fourier
61   component X_{ij} is distributed according to Gaussian with the mean defined in Equation S1 and variance
62   σ^2, and the Gaussian noise of each component is independent, the marginal probability of observing
63   image i can be obtained by integrating over all possible orientations ϕ and translations $\Delta = [\Delta_x, \Delta_y]$ as
64   follows,

$$P(X_i|V) \propto \int_{\phi,\Delta} \exp\left\{ -\frac{1}{\sigma^2} \sum_{j=1}^{J} \left(X_{ij} - e^{i\frac{2\pi}{N}(\Delta_x h + \Delta_y k)} \text{CTF}_{ij} \sum_{l=1}^{L} P_{jl}^{\phi} V_l \right)^2 \right\} d\phi d\Delta.$$ (Equation S3)

65   We omit translation factors in the squared difference term in Equation S3 to simplify expressions
66   elsewhere. The marginal probability of an image can then be leveraged to construct the log marginal
67   likelihood in Equation 1 in main text.

68   **Expectation maximization**

69   The expectation maximization algorithm works as follows. Since the difference between log
70   likelihoods of the marginal probability can be lower bounded by the difference between the sums of log
71   likelihoods of the joint probability weighted by their corresponding posterior probabilities for latent
72   variables, *i.e.*,

$$\log P(X_i|V) - \log P(X_i|V_{k-1})$$

$$\geq \sum_{\phi} -P(\phi|X_i, V_{k-1}) \left(\|X_i - \text{CTF}_i P^{\phi} V\|^2 - \|X_i - \text{CTF}_i P^{\phi} V_{k-1}\|^2 \right),$$ (Equation S4)

73   thus maximizing the lower bound improves the log likelihood of the marginal probability at least as
74   much(3). At the expectation step, we calculate the posterior probability of hidden variables conditioned on
75   a given image and the map. The method to compute the posterior probability derived in RELION(1) can
76   be applied in the context of our method without any modification.

77 **Weighted approximation**

78   The derivative of $\sum_j \log(|x_j| + \epsilon)$ at the point $x = x^i$ is $\frac{\text{sign}(x_j^i)}{|x_j^i| + \epsilon}$ for each component. Hence, the

79 tangent lines of $\sum_j \log(|x_j| + \epsilon)$ at this point are $\sum_j \frac{\text{sign}(x_j^i)}{|x_j^i| + \epsilon}(x_j - x_j^i) + \log(|x_j^i| + \epsilon)$, where $\text{sign}(x_j) =$

80 $\text{sign}(x_j^i)$. Note that x_j and x_j^i are of the same sign, the form of tangent lines can be simplified as,

$$g(x|x^i) = \sum_j \frac{|x_j|}{|x_j^i| + \epsilon} + \text{const}_j,$$ (Equation S5)

81 where $\text{const}_j = \frac{|x_j^i|}{|x_j^i| + \epsilon} + \log(|x_j^i| + \epsilon)$. We hence obtain the weighted approximation for the log norm. Since

82 the log norm is concave, its tangent line is its upper bound, namely, $g(x|x^i) \geq \sum_j \log(|x_j| + \epsilon)$, and

83 $g(x^i|x^i) = \sum_j \log(|x_j^i| + \epsilon)$. If $g(x^{i+1}|x^i) < g(x^i|x^i)$, we have $\sum_j \log(|x_j^{i+1}| + \epsilon) < g(x^{i+1}|x^i) < g(x^i|x^i) =$

84 $\sum_j \log(|x_j| + \epsilon)$. Combining the previous relation with Equation S4, we have $\log P(X_i|V) - \log P(X_i|V_{k-1}) -$

85 $\sum_{j=1}^{L}(\alpha \log(|x_j| + \epsilon) + \beta \log(\|\nabla x_j\|_2 + \epsilon')) \geq \sum_\phi -P(\phi|X_i, V_{k-1})\left(\|X_i - \text{CTF}_i P^\phi V\|^2 - \|X_i - \text{CTF}_i P^\phi V_{k-1}\|^2\right) -$

86 $g(x|x^i) - g(\nabla x|\nabla x^i)$. Thus, we can then prove that improving the lower bound will cause the left-hand

87 side of the inequality to improve at least as much by induction. The right-hand side of the preceding

88 inequality is the Equation 4 in main text ignoring constant terms and implicit gradient restraint.

89 **Nesterov smoothed TV norm**

90   To derive the gradient of Nesterov smoothed TV norm, we begin by stating the discrete form of

91 the gradient in TV norm. For a voxel $x[i, j, k]$ of a 3D map x, the gradient of the 3D map obtained by

92 discrete differentiation operator D at this voxel is of the form

$$Dx[i, j, k] = \begin{bmatrix} D_1 x[i, j, k] \\ D_2 x[i, j, k] \\ D_3 x[i, j, k] \end{bmatrix} = \begin{bmatrix} x[i, j, k] - x[i - 1, j, k] \\ x[i, j, k] - x[i, j - 1, k] \\ x[i, j, k] - x[i, j, k - 1] \end{bmatrix}.$$ (Equation S6)

93 By denoting D_i as the matrix representation of the discrete differentiation operator along the ith

94 dimension, we can express the gradient of the map x along the ith dimension as $D_i x$. By abuse of

95 notation, let $D = [D_1, D_2, D_3]^T$ be a matrix composed by concatenating D_i by rows, the TV norm can be

96 defined as,

$$\|x\|_{\text{TV}} = \max_{u \in Q_d} \langle u, Dx \rangle,$$ (Equation S7)

97 where $u = [u_1, u_2, u_3]^T \in Q_d$ is the vector of dual variables of the gradients of x along three directions, and

98 Q_d is the dual space in which each vector satisfies the inequality $u_1[i, j, k]^2 + u_2[i, j, k]^2 + u_3[i, j, k]^2 \leq 1$.

99 Using Nesterov smoothing with smoothing parameter μ, the smoothed TV norm $f_\mu(x)$ is the TV norm with

100 a quadratic restraint on the dual variables, which can be written as,

$$f_\mu(x) = \max_{u \in Q_d} \langle u, Dx \rangle - \frac{\mu}{2}\|u\|_2^2,$$ (Equation S8)

101 according to Candes *et al.*(4). The gradient of the smoothed TV norm $f_\mu(x)$ can be expressed as

$$\nabla f_\mu(x) = D^T u_\mu(x),$$ (Equation S9)

102 according to Candes *et al.*(4), where D is the discrete differentiation operator defined in Equation S6,

103 $u_\mu(x)$ is a vector of the form $[u_1, u_2, u_3]^T$ and for each dimension $a \in [1,2,3]$ and voxel $[i, j, k]$,

$$u_a[i, j, k] = \begin{cases} \mu^{-1}(D_a x)[i, j, k], \text{ if } \|\nabla x[i, j, k]\| < \mu \\ \|\nabla x[i, j, k]\|^{-1}(D_a x)[i, j, k], \text{ otherwise.} \end{cases}$$ (Equation S10)

104 For the weighted smoothed TV norm in Equation 4 in main text, its gradient $u_a[i, j, k]$ can be obtained by

105 multiplying Equation S10 with the weight $\frac{1}{\|\nabla x_{(i,j,k)}^i\| + \epsilon'}$. Equation S10 shows that the gradient of smoothed

106 TV norm can be easily obtained by first calculating the norm of discrete gradient of the volume x at each

107 voxel $[i, j, k]$, and then setting the norm of gradient at this voxel $\|\nabla x[i, j, k]\|$ with value smaller than the

108 smoothing parameter μ to μ, thus keeping the denominator of the gradient of TV norm in a valid range

109 and avoiding the non-differentiability of the non-smoothed TV norm at zero. With the form of discrete
110 differentiation operator D in Equation S6, we can write the gradient of the smoothed TV norm at a voxel
111 $[i, j, k]$ as follows,

(Equation S11)

$$\nabla f_\mu(x)_{ijk} = \sum_{a=1}^{3} u_a[i, j, k] - u_a[(i, j, k) + \Delta_a],$$

112 where Δ_a is a 1×3 vector with one on the ath entry and zeros elsewhere. Substituting Equation S10 into
113 Equation S11 leads to the complete form of the gradient in Equation 5 in main text. We then observe that
114 the gradient of TV norm at the voxel $[i, j, k]$ depends on gradients of 3D map around this voxel.

Local kernel regression

116   Nonparametric regression is often used to estimate the value of a point given the values of its
117 neighborhoods. Denote Y_i as the value at a certain point $x_i \in \mathbb{R}^N$, and let y be the value at the point $x \in$
118 $\mathbb{R}^N$ which is to be predicted, as in Takeda *et al.*(2), we can define y as the maximizer of

(Equation S12)

$$-\sum_{i=1}^{L} K(x_i, x)\|Y_i - y\|^2,$$

119 where $K(\cdot)$ represents a chosen kernel function. In the context of cryo-EM refinement, for an orientation
120 ϕ, let $n_j(\phi) = [h_j, k_j, l_j] \in \mathbb{R}^3$ be the back-projected voxel in the 3D volume V which is rotated from the
121 Fourier coefficient j of an image X_i (see "Back-projection as lock kernel regression"), we can define the
122 value of the target voxel $n = [h, k, l] \in \mathbb{Z}^3$ as the maximizer of the following regression problem,

(Equation S13)

$$-\sum_{i=1}^{N}\sum_{j=1}^{J}\sum_{\phi} P(\phi|X_i, V_{k-1})K(n_j(\phi), n)\|X_{ij} - \text{CTF}_{ij}V_n\|^2,$$

123 where $K(\cdot)$ is a kernel function measuring the closeness of the back-projected voxel and the target voxel.
124 A summation on losses of all voxels leads to Equation 7 in main text.

Parameter settings

126   Theoretically optimal scales of the parameters in OPUS-SSRI can be obtained by using the
127 closed form solution of our new target function. For the LASSO type problem, the closed form solution
128 can be derived from its dual form(5). The first step towards the dual form of our new target function is
129 converting our new target function to a matrix form. Assume x is the 3D volume which is rearranged into a
130 vector, that is, a voxel with index $[i, j, k]$ is mapped to the hth component x_h of x, and let A be the
131 corresponding 3D Fourier transform matrix, we can express the Fourier coefficients of the 3D volume V
132 as the result of matrix vector multiplication, namely, $V = Ax$. With Equation 7 in main text in hand, we can
133 write the matrix form of our 3D reconstruction problem as

(Equation S14)

$$\min_{x} \frac{1}{2}\|y - DAx\|^2 + \alpha \sum_{j=1}^{L} \frac{|x_j|}{|x_j^i| + \epsilon} + \beta \sum_{j=1}^{L} \frac{\|\nabla x_j\|_2}{\|\nabla x_j^i\|_2 + \epsilon'} + \gamma \|x - x^{k-1}\|_2^2,$$

134 where y is a vector representation of the 3D Fourier transform data with $y(n) =$
135 $\frac{\sum_{i=1}^{N}\sum_{j=1}^{J}\sum_{\phi} P(\phi|X_i, V_{k-1})K(n_j(\phi), n)\text{CTF}_{ij}X_{ij}}{\sqrt{N(n)}}$ for $n = [h, k, l] \in [0, L]$, D is a $L \times L$ diagonal matrix with diagonal
136 element $D(n, n) = \sqrt{N(n)}$, and Ax is the Fourier transform of the 3D map x. We can derive the dual form
137 of Equation S14 by simplifying our restraint. According to Tibshirani *et al.*(5), substituting the restraints in
138 Equation S14 by a generalized LASSO restraint, $\lambda\|Gx\|_1$, the dual of Equation S14 with new restraint is of
139 the form,

$$\min_{u}(A^H Dy - G^T u)^T (A^H D^2 A)^+ (A^H Dy - G^T u)$$

(Equation S15)

140 subject to $\|u\|_\infty \leq \lambda$, $G^T u \in \text{row}(DA)$, where $(A^H D^2 A)^+$ is the Moore-Penrose inverse of $A^H D^2 A$, $\|u\|_\infty = $
141 $\max_{u_i \in u}|u_i|$, λ is the parameter for l_1 restraint and u is the dual variable of the 3D volume x. Given u, the
142 closed form solution of x can be written as,

$$x = (A^H D^2 A)^+ (A^H Dy - G^T u). \qquad \text{(Equation S16)}$$

143   In Equation S16, $A^H Dy$ represents the inverse Fourier transform of the data, which is the unregularized
144   solution. $G^T u$ is the dual variable of restraint, which regularizes $A^H Dy$. To achieve sparseness in the
145   solution x, u needs to zero out certain components of $A^H Dy$. Since the dual variable u is bounded by λ,
146   the restraint parameter λ should be of the same scale as the average of the magnitudes of $A^H Dy$. Though
147   our restraint is of a more complex form than $\|Gx\|_1$, the dual of our restraint is in the space of a
148   combination of two domains similar to $\|u\|_\infty \le \lambda$. Detailed derivations about the dual space of the
149   combination of two norm can be found in Rockafellar *et al.*(6). Therefore, the simplified discussion drives
150   us to set the parameters of our restraint to be of the scale as the square root of the average of the
151   squares of $A^H Dy$. We denote the square root of the average of the squares of $A^H Dy$ as $\overline{\|A^H Dy\|_2}$
152   henceforth.
153   The scale of implicit gradient ascent restraint is easy to set since it is quadratic. Note that each
154   quadratic data loss term in our target function is scaled by $N(n)$ in Equation 8 in main text. Using the
155   heuristic that the penalty term should match the loss term, we can set the restraint parameter γ to be on
156   the scale of the average of $N(n)$. We denote the average of $N(n)$ as $\overline{N(n)}$ henceforth.
157   Other important parameters to be set are ϵ and ϵ', which are in the denominators of our sparsity
158   and smoothness restraint, respectively. If they are too small compared to the density values of the 3D
159   volume, the weights in our weighted norms will be very flexible and strongly depend on the magnitude of
160   the value of each voxel in the 3D volume. Such kinds of restraints might not be able to effectively remove
161   background noises and cause two independent refinements to diverge. If they are too large compared to
162   the density values of the 3D volume, the restraints degrade to the original l_1 and TV norms and leads to
163   more biased solutions. Optimal values of ϵ and ϵ' should assign large weights to background noises and
164   small weights to true molecular densities. Hence, we can set ϵ to the level of density values
165   corresponding to molecular content in the 3D volume. ϵ' can be set close to ϵ. This level can be easily
166   obtained from the intermediate volumes generated by the refinements using RELION 3.0. This is also
167   similar to choose a threshold for creating a mask when computing masked FSC.
168   In conclusion, we should set the restraint parameters α and β to be on the same scale as
169    $\overline{\|A^H Dy\|_2}$. Since the corresponding restraints are inversely weighted by quantities with two other
170   parameters ϵ and ϵ', we multiply the scale $\overline{\|A^H Dy\|_2}$ by ϵ or ϵ' to counter acting the effects of ϵ and ϵ'. For
171   zero elements, their restraint parameters are then normalized to be on the scale of $\overline{\|A^H Dy\|_2}$.

172 **Experiment process**

173   β-galactosidase (EMPIAR-10017) : Since there was no ready-to-use particle stack for model
174   building, our test began with extracting particles from micrographs using the coordinates manually picked
175   by Richard Henderson(7). We carried out 3D refinement in RELION 3.0 with a 50 Å low-pass filtered initial
176   map while enforcing D2 symmetry. Then, we performed further 3D refinement using OPUS-SSRI with a
177   grid search to determine the possible ranges of parameters α, β, γ and ε. We started by setting ϵ to 0.1,
178   which is higher than the level of density values of the EM map from RELION 3.0. We considered setting ϵ'
179   to be $\epsilon/3$ since the magnitude of gradient is often smaller than the density value of molecule. The initial
180   guesses for β and γ were $\beta = 2/3\overline{\|A^H Dy\|_2}\epsilon$, $\gamma = 0.05\overline{N(n)}$. We scanned through $\alpha \in [0.4,0.6]\overline{\|A^H Dy\|_2}\epsilon$
181   with as a step size of $0.1\overline{\|A^H Dy\|_2}\epsilon$. The final resolutions based on gold standard FSC=0.143 for different
182   α is shown in **Fig.S4a**. The best resolution was obtained at $\alpha = 0.5\overline{\|A^H Dy\|_2}\epsilon$. We then set $\alpha =$
183    $0.5\overline{\|A^H Dy\|_2}\epsilon$ and scanned through $\beta \in [1/3, 2.2/3]\overline{\|A^H Dy\|_2}\epsilon$ with a step size of $0.2/3\overline{\|A^H Dy\|_2}\epsilon$. The
184   best resolution was obtained at $\beta = 1.6/3\overline{\|A^H Dy\|_2}\epsilon$ or $1.8/3\overline{\|A^H Dy\|_2}\epsilon$ (**Fig.S4b**). To calculate the model
185   vs. map FSCs, we fitted the atomic coordinates of an *E. coli* β-galactosidase structure (PDB code
186   3I3E)(8) into the post-processed density maps reconstructed by different methods.
187   80S ribosome (EMPIAR-10002): We extracted particles from the micrographs using the semi-
188   automated selection process in RELION 3.0(9). The particles were pruned by one round of 2D
189   classification where only the particles classified to major classes were kept. We then constructed an *ab*
190   *initio* map in RELION 3.0 through 3D classification. 3D refinements continued from the 70 Å low pass
191   filtered initial map. For OPUS-SSRI, we used the same ϵ, ϵ' and γ as β-galactosidase and scanned
192   through $\alpha \in [0.3,0.5]\overline{\|A^H Dy\|_2}\epsilon$ with a step size of $0.1\overline{\|A^H Dy\|_2}\epsilon$. The resolutions with different α are

193 shown in **Fig.S4c**. The best result is obtained at $\alpha = 0.4\overline{\|A^HDy\|_2}\epsilon$. We then scanned through $\beta \in$
194 $[1.2,2.4]$ with a step size of $0.4/3\overline{\|A^HDy\|_2}\epsilon$ for $\alpha = 0.4\overline{\|A^HDy\|_2}\epsilon$ and $\alpha = 0.5\overline{\|A^HDy\|_2}\epsilon$. The results are
195 shown in **Fig.S4d**. The best results were obtained at $\alpha = 0.4\overline{\|A^HDy\|_2}\epsilon$ and $\beta = 1.6/3\overline{\|A^HDy\|_2}\epsilon$ or
196 $2/3\overline{\|A^HDy\|_2}\epsilon$. To assess the cryo-EM maps determined using different methods, we fitted 80S crystal
197 structure (PDB code 3U5B(10)) using a simple rigid-body fit into post-processed maps. The 40S and 60S
198 subunits were fitted separately.
199   Hemagglutinin (EMPIAR-10097): The 3D refinements were performed by using a 40 Å low-
200 passed filtered initial map and enforcing the C3 symmetry. For OPUS-SSRI, the optimal parameters are
201 $\epsilon = 0.015$ and $\epsilon' = \epsilon/3$, $\alpha = 0.4\overline{\|A^HDy\|_2}\epsilon$, $\beta = 2.6/3\overline{\|A^HDy\|_2}\epsilon$ and $\gamma = 0.2\overline{N(n)}$. To compare the post-
202 processed density maps obtained from different refinement methods, we used a crystal structure of HA
203 trimer (PDB code 3WHE)(11) as a reference following the original publication(12).
204   TRPM4 (EMPIAR-10126): The 3D refinement rounds were performed with a 50 Å low-passed
205 filtered initial map and C4 symmetry. The best parameters of our method were $\epsilon = 0.01$, $\epsilon' = \epsilon$, $\alpha =$
206 $0.8\overline{\|A^HDy\|_2}\epsilon$, $\beta = 4\overline{\|A^HDy\|_2}\epsilon$ and $\gamma = 0.1\overline{N(n)}$. The cryo-EM maps were compared with the atomic
207 model in PDB code 6BQR(13).
208   Hrd1/Hrd3 (EMPIAR-10099): Due to the heterogeneity of this dataset, the 3D classification was
209 used to classify the particles, and generate the corresponding initial maps of different complexes for 3D
210 refinements. The particles classified as Hrd1/Hrd3 dimer were selected, and then subject to 3D
211 refinements. We performed 3D refinements using different methods and the same 20 Å low-passed
212 filtered initial map. The best parameters of our method are $\epsilon = 0.01$, $\epsilon' = \frac{\epsilon}{2}$, $\alpha = 0.4\overline{\|A^HDy\|_2}\epsilon$, $\beta =$
213 $0.6\overline{\|A^HDy\|_2}\epsilon$ and $\gamma = 0.1\overline{N(n)}$. The final results were compared with the atomic models of Hrd1 dimer
214 (PDB code 5V6P) and Hrd3 monomer (PDB code 5V7V)(14). 5V6P and 5V7V were fitted into density
215 map separately.
216   TRPV5 (EMPIAR-10254): The 3D refinement rounds were performed by using a 40 Å low-pass
217 filtered 3D initial map and enforcing C4 symmetry. The best parameters of our method were $\epsilon = 0.01$, $\epsilon' =$
218 2ϵ, $\alpha = 0.5\overline{\|A^HDy\|_2}\epsilon$, $\beta = 2\overline{\|A^HDy\|_2}\epsilon$ and $\gamma = 0.1\overline{N(n)}$. The cryo-EM maps were compared to the atomic
219 model in PDB code 6O1P(15).
220   TMEM16A (EMPIAR-10123): The 3D refinements were performed by using a 40 Å low-passed
221 filtered initial map and enforcing the C2 symmetry. The best parameters of our method are $\epsilon = 0.01$, $\epsilon' =$
222 ϵ, $\alpha = 0.5\overline{\|A^HDy\|_2}\epsilon$, $\beta = 2\overline{\|A^HDy\|_2}\epsilon$ and $\gamma = 0.1\overline{N(n)}$. The final results were compared to the atomic
223 model in PDB code 6BGI(16) by calculating model vs. map FSC.
224
225

226

227 **Table S1. Comparison of the final reconstructions refined by RELION 3.0, THUNDER or OPUS-SSRI**

Proteins	Model *vs.* Map FSC=0.143					
	RELION	THUNDER		OPUS-SSRI		
	Resolution (Å)	Resolution (Å)	ΔÅ over RELION[1]	Resolution (Å)	ΔÅ over RELION[1]	ΔÅ over THUNDER[2]
β-galactosidase (EMPIAR-10017)	4.05	4.08	-0.03	3.91	0.14	0.17
80S ribosome (EMPIAR-10002)	4.04	3.70	0.34	3.89	0.15	-0.19
Hemagglutinin (EMPIAR-10097)	4.06	3.89	0.17	3.72	0.34	0.17
TRPM4 (EMPIAR-10126)	3.34	/	/	2.93	0.41	/
Hrd1/Hrd3 (EMPIAR-10099)	4.70	4.88	-0.18	4.25	0.45	0.63
TRPV5 (EMPIAR-10254)	3.05	2.98	0.07	2.76	0.29	0.22
TMEM16A (EMPIAR-10123)	3.87	/	/	3.14	0.73	/
Average improvement			**0.07**		**0.30**	**0.20**

228
229 [1]: the value in negative indicates the resultant resolution is worse than that from RELION, while the value in positive indicates the resultant resolution is better than that from RELION.
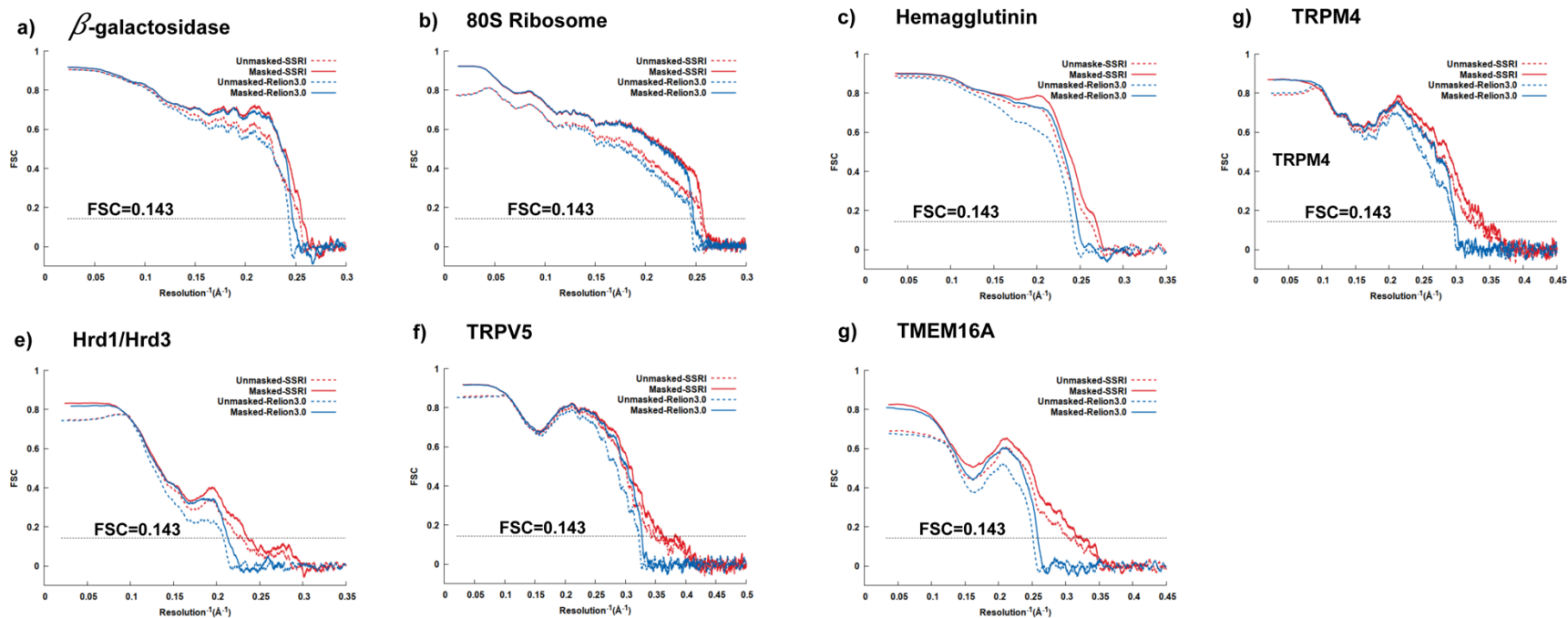
230
231 [2]: the value in negative indicates the resultant resolution is worse than that from THUNDER, while the value in positive indicates the resultant resolution is better than that from THUNDER.

232 /: indicates that the comparison was unavailable in two cases where THUNDER failed to execute due to computer incompatibility.
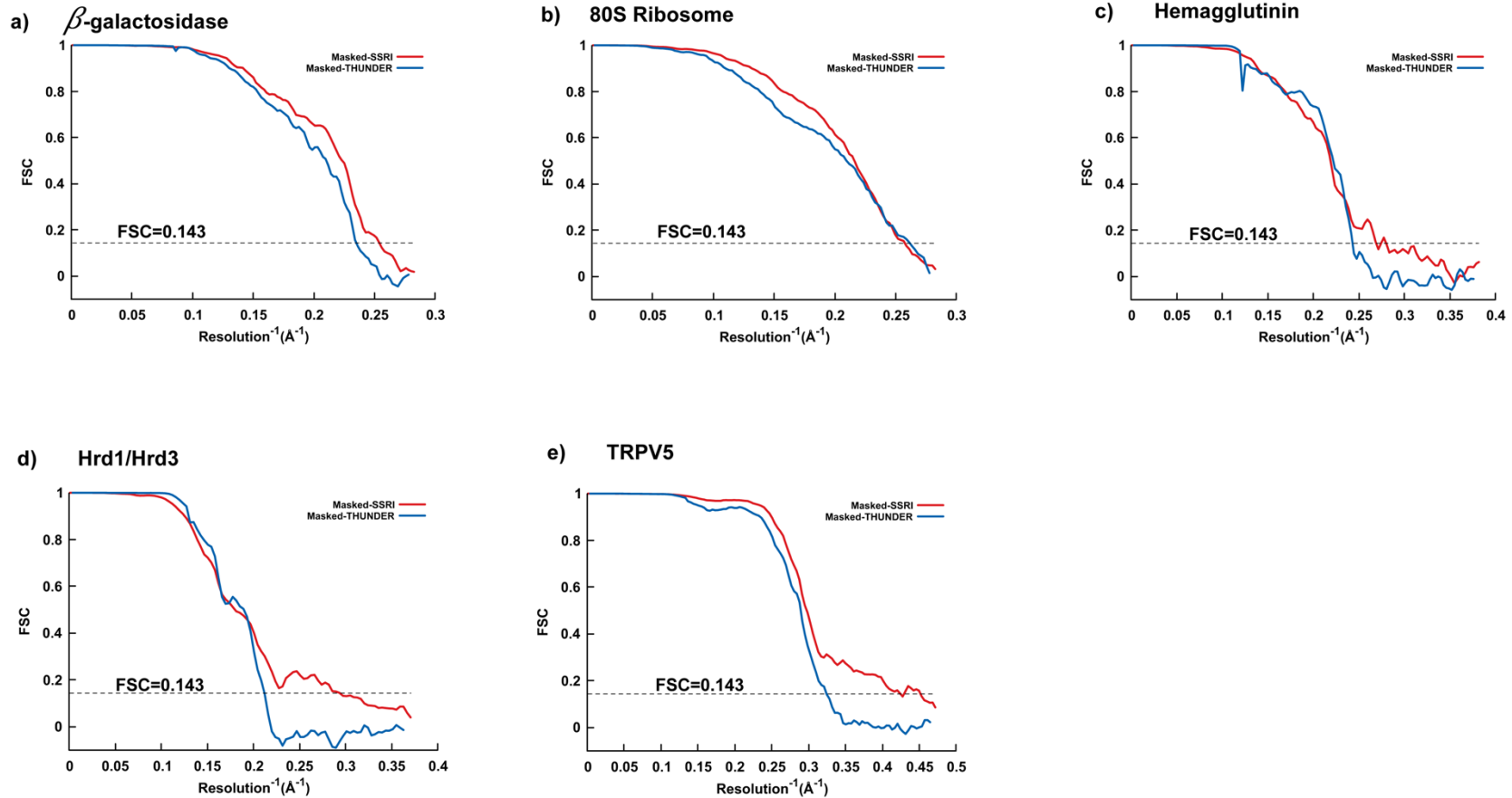
233
234

Fig. S1. Model *vs.* map FSC curves between the post processed density maps and the corresponding rigid-body fitted atomic models for seven proteins. **(a)** β-galactosidase with atomic model in PDB code 3I3E; **(b)** 80S ribosome with atomic model in PDB code 3U5B, for which the 40S and 60S subunits were fitted separately; **(c)** Influenza hemagglutinin with atomic model in PDB code 3WHE; **(d)** TRPM4 with atomic model in PDB code 6BQR; **(e)** Hrd1/Hrd3 complex, for which the Hrd1 dimer (PDB code 5V6P) and Hrd3 monomer (PDB code 5V7V) were fitted separately; **(f)** TRPV5 with atomic model in PDB code 6O1P; **(g)** TMEM16A with atomic model in PDB code 6BGI. Dashed black line represents FSC=0.143 in all panels.
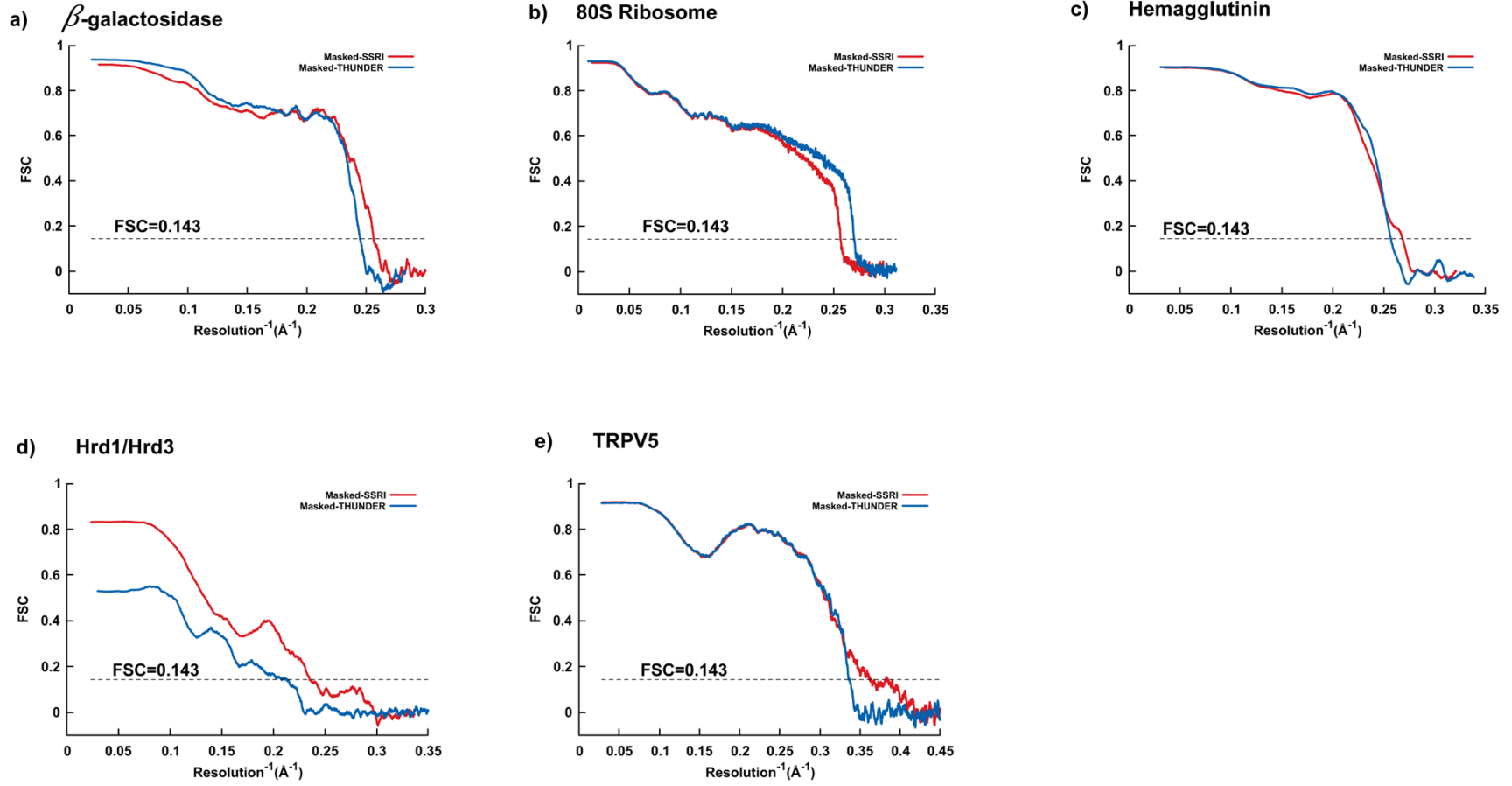
Fig. S2. Gold standard masked FSC curves for the final 3D reconstructions refined by OPUS-SSRI or THUNDER for **(a)** β-galactosidase, **(b)** 80S ribosome, **(c)** influenza hemagglutinin, **(d)** Hrd1/Hrd3 and **(e)** TRPV5. In all panels, dashed black line represents FSC=0.143.

Fig. S3. Model *vs*. map FSC curves between the post processed density maps and the corresponding rigid-body fitted atomic models for five proteins. **(a)** β-galactosidase with atomic model in PDB code 3I3E; **(b)** 80S ribosome with atomic model in PDB code 3U5B, for which the 40S and 60S subunits were fitted separately; **(c)** Influenza hemagglutinin with atomic model in PDB code 3WHE; **(d)** Hrd1/Hrd3 complex, for which the Hrd1 dimer (PDB code 5V6P) and Hrd3 monomer (PDB code 5V7V) were fitted separately; **(e)** TRPV5 with atomic model in PDB code 6O1P. Dashed black line represents FSC=0.143 in all panels.

9
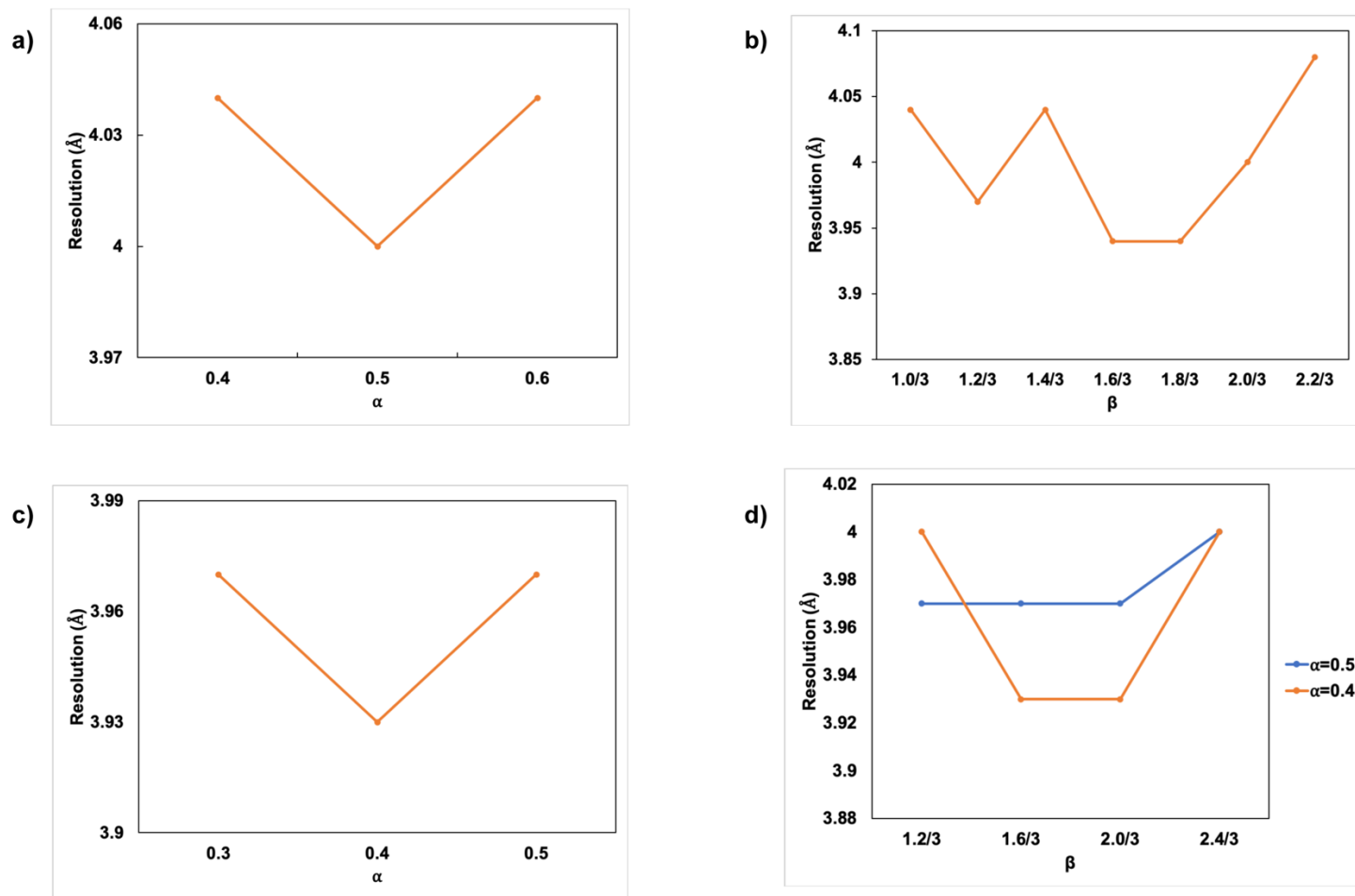
Fig. S4. Grid search for optimal parameters in OPUS-SSRI. **(a)** Resolutions of EM maps refined with different α for β-galactosidase. **(b)** Resolutions of EM maps refined with different β for β-galactosidase. **(c)** Resolutions of EM maps refined with different α for 80S ribosome. **(d)** Resolutions of EM maps refined with different β and α for 80S ribosome.

SI References

1. S. H. Scheres, A Bayesian view on cryo-EM structure determination. *J. Mol. Biol.* **415**, 406–418 (2012).
2. H. Takeda, S. Farsiu, P. Milanfar, Kernel regression for image processing and reconstruction. *IEEE Trans. Image Process.* **16**, 349–366 (2007).
3. Little, J. A. Roderick, Rubin, B. Donald, Statistical analysis with missing data. *Technometrics* **45**, 364–365 (2002).
4. S. Becker, J. Bobin, E. J. Candès, NESTA: A fast and accurate first-order method for sparse recovery. *SIAM J. Imaging Sci.* **4**, 1–39 (2011).
5. R. J. Tibshirani, J. Taylor, others, The solution path of the generalized lasso. *Ann. Stat.* **39**, 1335–1371 (2011).
6. R. T. Rockafellar, *Convex analysis* (Princeton university press, 1970).
7. K. R. Vinothkumar, G. McMullan, R. Henderson, Molecular mechanism of antibody-mediated activation of β-galactosidase. *Structure* **22**, 621–627 (2014).
8. M. L. Dugdale, D. L. Dymianiw, B. K. Minhas, I. d'Angelo, R. E. Huber, Role of Met-542 as a guide for the conformational changes of Phe-601 that occur during the reaction of β-galactosidase (Escherichia coli). *Biochem. Cell Biol.* **88**, 861–869 (2010).
9. S. H. Scheres, Semi-automated selection of cryo-EM particles in RELION-1.3. *J. Struct. Biol.* **189**, 114–122 (2015).
10. A. Ben-Shem, *et al.*, The structure of the eukaryotic ribosome at 3.0 \AA resolution. *Science* **334**, 1524–1529 (2011).
11. Y. Iba, *et al.*, Conserved neutralizing epitope at globular head of hemagglutinin in H3N2 influenza viruses. *J. Virol.*, JVI–00420 (2014).
12. Y. Z. Tan, *et al.*, Addressing preferred specimen orientation in single-particle cryo-EM through tilting. *Nat. Methods* **14**, 793 (2017).
13. H. E. Autzen, *et al.*, Structure of the human TRPM4 ion channel in a lipid nanodisc. *Science* **359**, 228–232 (2018).
14. S. Schoebel, *et al.*, Cryo-EM structure of the protein-conducting ERAD channel Hrd1 in complex with Hrd3. *Nature* **548**, 352 (2017).
15. S. Dang, *et al.*, Structural insight into TRPV5 channel function and modulation. *Proc. Natl. Acad. Sci.* **116**, 8869–8878 (2019).
16. S. Dang, *et al.*, Cryo-EM structures of the TMEM16A calcium-activated chloride channel. *Nature* **552**, 426–429 (2017).