Sparseness and Smoothness Regularized Imaging for improving the resolution of Cryo-EM single-particle reconstruction

Zhenwei Luo (罗镇威)[a], Adam A. Campos-Acevedo[b], Longfei Lv (吕龙飞)[c], Qinghua Wang[b], and Jianpeng Ma (马剑鹏)[a,b,c,1]

[a]Department of Bioengineering, Rice University, Houston, TX 77030; [b]Verna and Marrs Mclean Department of Biochemistry and Molecular Biology, Baylor College of Medicine, Houston, TX 77030; and [c]Multiscale Research Institute of Complex Systems, Fudan University, Shanghai, 201203, China

In this paper, we present a refinement method for cryo-electron microscopy (cryo-EM) single-particle reconstruction, termed as OPUS-SSRI (Sparseness and Smoothness Regularized Imaging). In OPUS-SSRI, spatially varying sparseness and smoothness priors are incorporated to improve the regularity of electron density map, and a type of real space penalty function is designed. Moreover, we define the back-projection step as a local kernel regression and propose a first-order method to solve the resulting optimization problem. On the seven cryo-EM datasets that we tested, the average improvement in resolution by OPUS-SSRI over that from RELION 3.0, the commonly used image-processing software for single-particle cryo-EM, was 0.64 Å, with the largest improvement being 1.25 Å. We expect OPUS-SSRI to be an invaluable tool to the broad field of cryo-EM single-particle analysis. The implementation of OPUS-SSRI can be found at https://github.com/alncat/cryoem.

Cryo-EM | 3D reconstruction | ill-posed inverse problem | smoothness | sparseness

Cryo-electron microscopy (Cryo-EM) single-particle analysis is a powerful method for determining macromolecular structures. The major advantages of cryo-EM over the traditional X-ray crystallography are that it does not require crystallization and is not plagued by the phase problem. However, there remain many new challenges in this promising technique. The central problem of cryo-EM single-particle analysis is the incompleteness of experimental observations. More specifically, the information of the relative orientations and translations of all particles is missing. Furthermore, in a dataset with multiple conformations (or substates), the membership of a specific class needs to be defined. Moreover, the signal-to-noise ratio (SNR) of a cryo-EM dataset is often very low since the electron exposure of the sample needs to be strictly limited to reduce radiation damage (1). Other problems often present in cryo-EM datasets include the nonuniform angular sampling, which frequently results in inadequate sampling or even no sampling in certain orientations (2). Therefore, the problem of cryo-EM three-dimensional (3D) reconstruction is an extremely ill-posed problem. To alleviate the ill-posedness of this problem, prior assumptions must be incorporated into the reconstruction process to ensure the uniqueness of solution and the objectivity of the final maps.

Two outstanding features of 3D density maps are sparseness and smoothness. Specifically, since the atoms in macromolecules only occupy part of the 3D maps, the macromolecular maps are often sparse in space. On the other hand, because the atoms in macromolecules are connected through chemical bonds, the electron densities of macromolecules vary smoothly across the space (3). Though sparseness is a popular prior in solving inverse problems, it is a relatively novel notion to cryo-EM 3D reconstruction. In contrast, the importance of smoothness prior is widely recognized in cryo-EM 3D refinement. An early attempt to enforce the smoothness of the density map was to apply the Wiener filter (4). Later approaches improved upon the Wiener filter by using Bayesian statistics (3). Scheres et al. assumed that the Fourier components of the density map are distributed according to Gaussian distributions (3) a priori and derived a maximum a posteriori estimation for reconstruction. This approach, as implemented in REgularised LIkelihood OptimisatioN (RELION) (5), is referred to as the traditional approach in the context of this paper. Except for incorporating priors into the reconstruction process, another line of efforts aims to enhance cryo-EM 3D refinement by optimizing the defocus parameter and class membership for each particle, as exemplified by THUNDER (6). THUNDER has been shown to improve cryo-EM refinement by providing more accurate contrast transfer function and membership for each particle.

In this paper, we continued the direction used in RELION and proposed an approach to regularize the 3D maps. Our approach, named OPUS-SSRI (Sparseness and Smoothness Regularized Imaging), focuses on imposing sparseness and smoothness priors (i.e., l_1 regularization) (7) and total variation (TV) (8). To encourage sparseness and smoothness of the density map while suppressing bias, we proposed a nonconcave, nonsmooth, real-space restraint by combining l_1 regularization and TV norm. Since such target function is difficult to optimize, we designed a reweighted scheme to approximately optimize it with a sequence of weighted l_1 and TV problem. The major differences between the traditional approach in RELION and our approach in OPUS-SSRI are outlined in detail in the *Materials and Methods* section. In essence, the traditional approach in RELION can be viewed as applying a translation-invariant isotropic kernel to smooth the 3D map, whereas OPUS-SSRI applies a spatially varying anisotropic kernel. Another challenge of 3D map reconstruction in real space is the prohibitive computation cost as a result of its ultrahigh dimensionality. For example, a common $512 \times 512 \times 512$ 3D volume contains hundreds of millions of

---

Significance

Three-dimensional refinement is a critical component of cryo-EM single-particle reconstruction. In this paper, we report the development of a computational method, OPUS-SSRI, and its application to seven real cryo-EM datasets. Our data clearly demonstrated that OPUS-SSRI can improve the final resolutions and structural details in cryo-EM single-particle analysis.

variables. We addressed this challenge by leveraging a Compute Unified Device Architecture (CUDA)-accelerated implementation. In addition, we cast the back-projection as a local kernel regression problem, thus paving a way to promote the smoothness of the 3D map (9). By applying it to seven real cryo-EM datasets, we demonstrated that our OPUS-SSRI supports a significant improvement over RELION 3.0 or THUNDER, especially for systems with significant flexibility.

Results

We tested OPUS-SSRI by performing 3D refinement on a total of seven real datasets and comparing the refinement results with those obtained using RELION 3.0 or THUNDER. The detailed experimental process and optimal parameters are reported in the *SI Appendix*.

According to the gold-standard Fourier shell correlation (FSC) at 0.143, the final density maps reconstructed by OPUS-SSRI clearly have higher SNRs compared to those generated by RELION 3.0 in most resolution shells for β-galactosidase (10, 11) (Fig. 1*A*), 80S ribosome (12) (Fig. 1*B*), influenza hemagglutinin (HA) (13) (Fig. 1*C*), transient receptor potential melastatin (TRPM4) (14) (Fig. 1*D*), protein-conducting ERAD channel Hrd1/Hrd3 complex (15) (Fig. 2*A*), transient receptor

potential vanilloid 5 (TRPV5) (16) (Fig. 3*A*), and calcium-activated chloride channel (TMEM16A) in nanodisc (17) (Fig. 4*A*). The final maps refined by OPUS-SSRI have resolutions that are 0.15 to 1.25 Å better than those refined by RELION 3.0, with averaged resolution improvement of 0.64 Å for all seven systems (Table 1). The improvement of the density maps reconstructed by OPUS-SSRI is also confirmed by the model versus map FSCs. The postprocessed maps of OPUS-SSRI have much higher correlations with respect to the corresponding rigid-body fitted atomic models in most resolution shells than RELION 3.0 (*SI Appendix*, Fig. S1). Overall, for the seven systems, the improvements in resolution for the postprocessed maps of OPUS-SSRI are in the range of 0.14 to 0.73 Å, with an average of 0.30 Å, over those refined by RELION 3.0 (*SI Appendix*, Table S1).

THUNDER was also run on five of these seven systems (it failed to execute on two datasets due to incompatibility with our computing facility). According to the gold-standard FSC at 0.143, the improvements in resolution by THUNDER over RELION 3.0 are in the range of −0.09 to 0.28 Å with an average of 0.07 Å (Table 1). If judged by the model versus map FSCs at 0.143, the improvements in resolution of THUNDER over RELION 3.0 are in the range of −0.18 to 0.17 Å with an average of 0.07 Å on the five systems (*SI Appendix*, Table S1). Of these

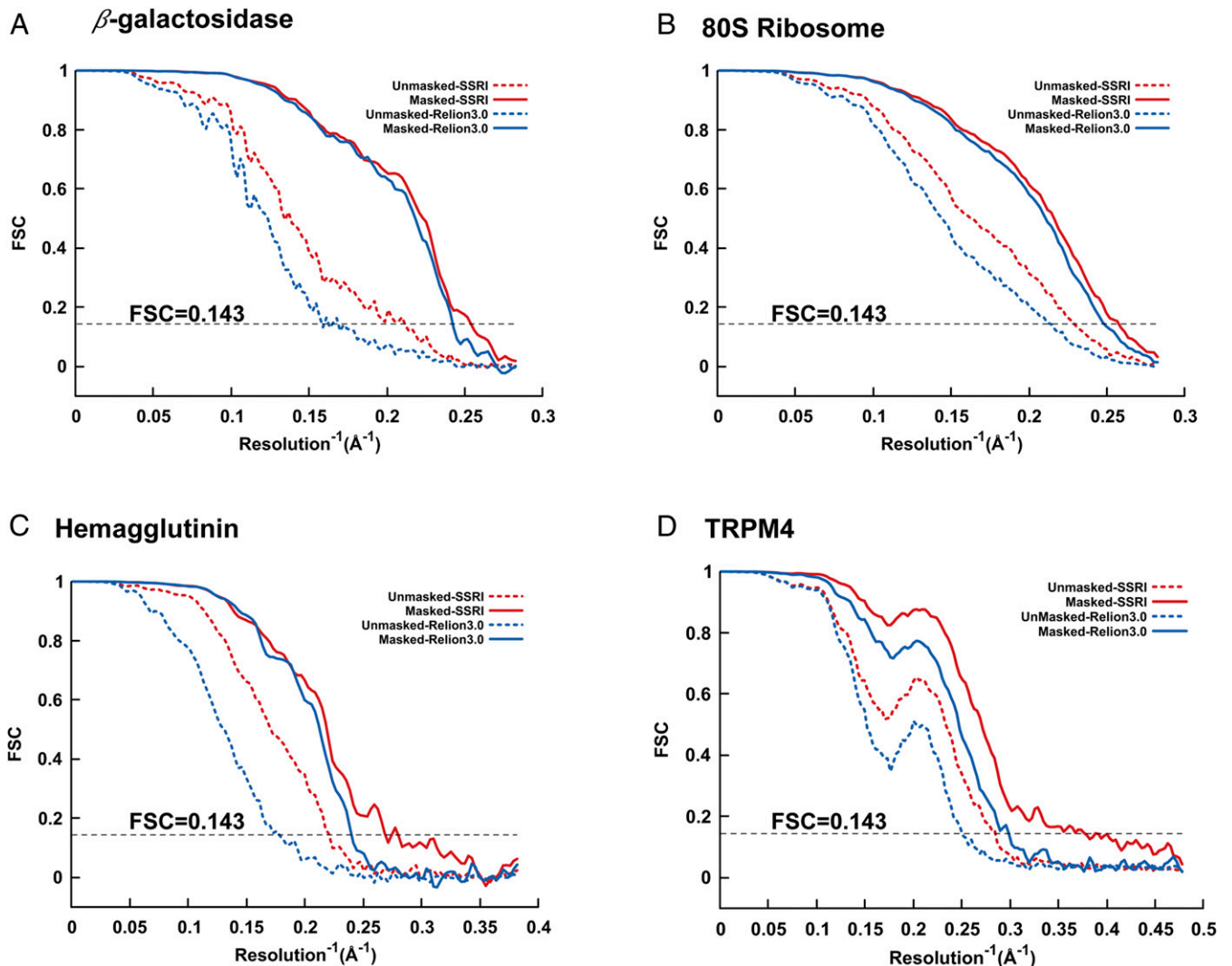**Fig. 1.** Gold-standard unmasked and masked FSC curves for the final 3D reconstructions refined by OPUS-SSRI (in red color) or RELION 3.0 (in blue color) for (*A*) β-galactosidase, (*B*) 80S ribosome, (*C*) influenza hemagglutinin, and (*D*) TRPM4. In all panels, the dashed black line represents FSC = 0.143.

Luo et al.
Sparseness and Smoothness Regularized Imaging for improving the resolution of Cryo-EM single-particle reconstruction
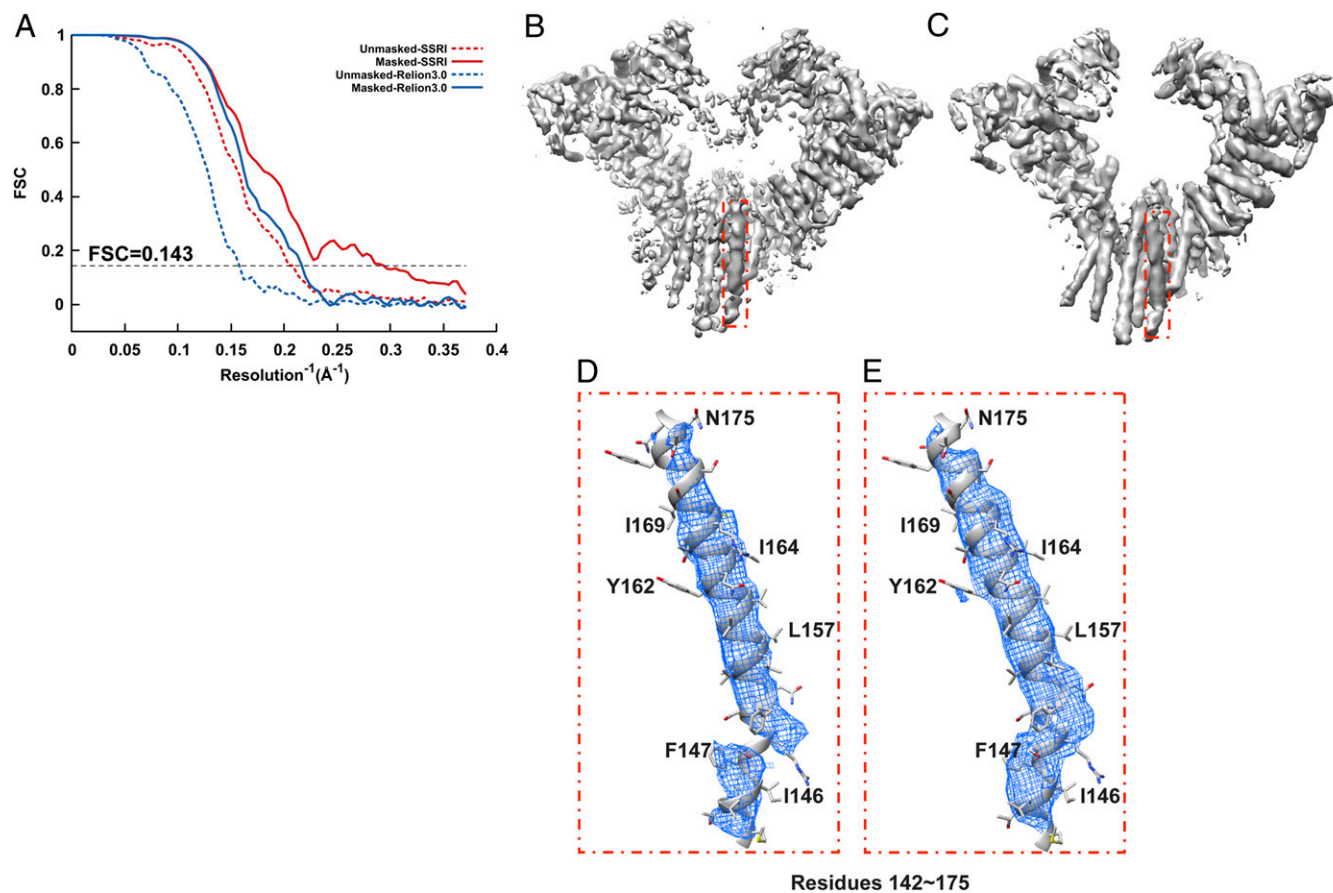
Fig. 2. Refinement of Hrd1/Hrd3 complex. (*A*) Gold-standard unmasked and masked FSC curves calculated from two independent reconstructions by OPUS-SSRI or RELION 3.0. The dashed black line represents FSC = 0.143. (*B*) Final reconstructed cryo-EM map using RELION 3.0. (*C*) Final reconstructed cryo-EM map using OPUS-SSRI. The red rectangle defines a region of the EM map to be enlarged in *D* for RELION 3.0 and *E* for OPUS-SSRI for residues 142 to 175, respectively. The EM density is represented in mesh (blue), and the atomic model is shown in a ribbons diagram with side chains in stick presentation. Both density maps are contoured at the same level.

five systems, OPUS-SSRI constantly outperforms THUNDER on four systems and only slightly underperforms THUNDER on one system (80S ribosome) as gauged by the gold-standard FSC = 0.143 and model versus map FSC = 0.143. Overall, OPUS-SSRI produces an average improvement of 0.47 Å in resolution over THUNDER for all five systems if judged by the gold-standard FSC = 0.143, with the largest improvement being 1.20 Å (Table 1 and *SI Appendix*, Fig. S2), and of 0.20 Å in resolution if judged by the model versus map FSC = 0.143, with the largest improvement being 0.63 Å (*SI Appendix*, Table S1 and Fig. S3).

Fig. 2 shows some of the structural improvements for Hrd1/Hrd3 complex in more detail. Clearly, compared to the density map reconstructed by RELION 3.0 (Fig. 2*B*), the density map from OPUS-SSRI is much sharper and cleaner (Fig. 2*C*). In fact, out of the seven systems studied, OPUS-SSRI refinement on Hrd1/Hrd3 complex results in the largest improvements in resolution (Table 1 and *SI Appendix*, Table S1). For instance, in the density map from RELION 3.0, there is a gap in the main-chain density between residues 147 and 148 (Fig. 2*D*). However, in the density map from OPUS-SSRI, the density in this region becomes continuous and strong (Fig. 2*E*).

Similarly, for TRPV5, comparing to the final map obtained by RELION 3.0 (Fig. 3*B*), the density map from OPUS-SSRI becomes much sharper with improved SNRs (Fig. 3*C*). Most impressively, the density map from OPUS-SSRI even allows retracing of the structural model in the region of residues 374 to

380 that was out of the density map in the original structure (highlighted in dashed red circle in Fig. 3*D*). After the manual adjustment in the crystallographic object-oriented toolkit COOT (18) and structural refinement using Python-based Hierarchical ENvironment for Integrated Xtallography (PHENIX) (19), the match between the model and map is substantially improved (highlighted by dashed red circle in Fig. 3*E*).

In addition, for TMEM16A, in contrast to the density map from RELION 3.0 (Fig. 4*B*), the density map obtained by OPUS-SSRI (Fig. 4*C*) shows sharper and smoother densities with less noise throughout. The improvement from OPUS-SSRI is highlighted for two helices in the regions of residues 408 to 440 (Fig. 4 *D* and *E*) and 848 to 884 (Fig. 4 *F* and *G*). Most impressively, in the density map refined by OPUS-SSRI, the densities for side chains of residues F412, M416, W419, and F423 (Fig. 4*E*) and F863, I865, F867, and N869 (Fig. 4*G*) become very well separated, in marked contrast to the blobs of densities from RELION 3.0 in Fig. 4 *D* and *F*, respectively.

Discussion

In this paper, we proposed OPUS-SSRI, a 3D refinement method for cryo-EM single-particle analysis. The improvement of our method in gold-standard FSC of the final reconstructions is the most noticeable, which can be largely attributed to the superior denoising effect of the sparseness and smoothness priors that we introduced. By setting relatively small components in the 3D map to zero and filtering components to be more
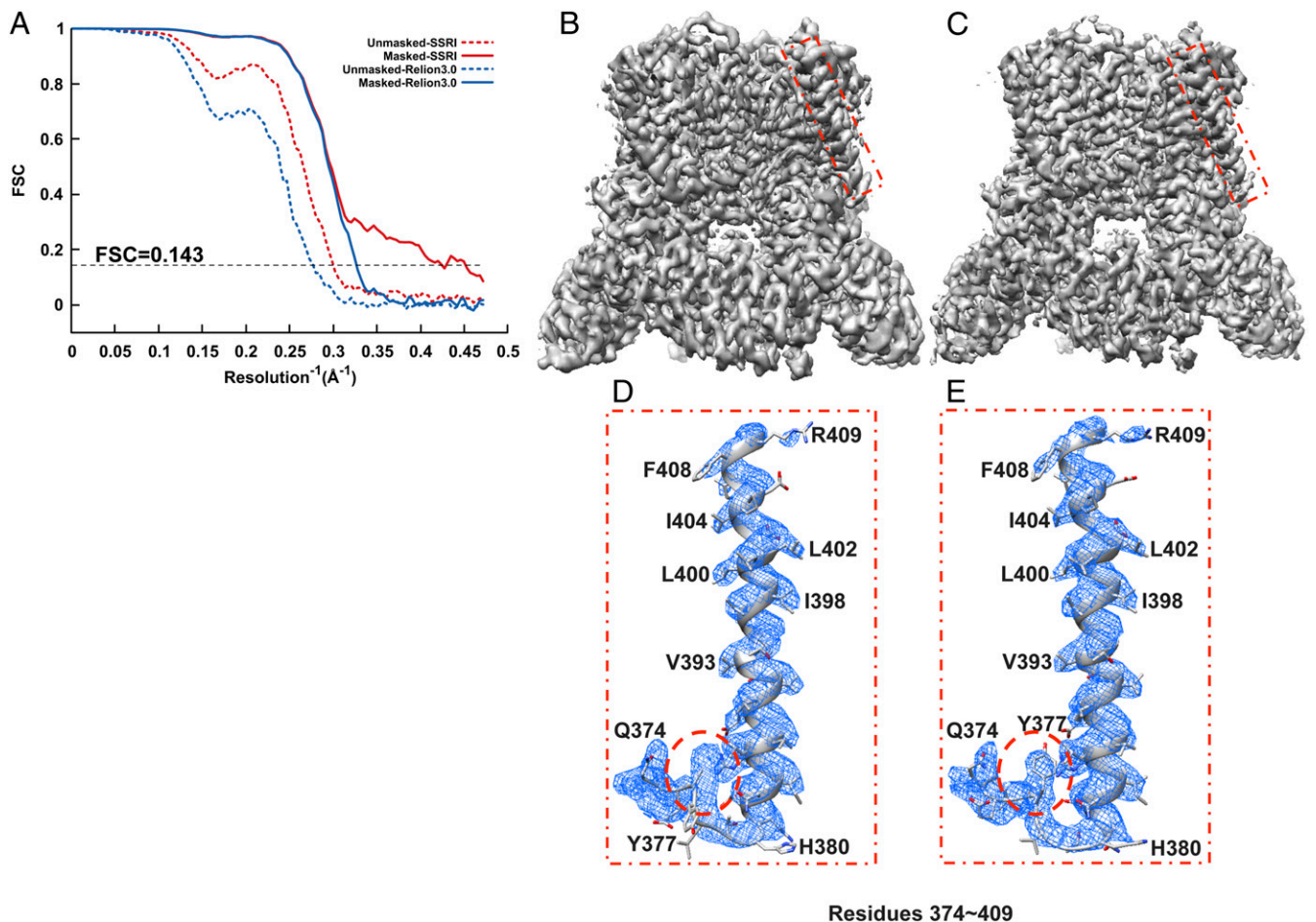
Fig. 3. Refinement of TRPV5. (*A*) Gold-standard unmasked and masked FSC curves calculated from two independent reconstructions by OPUS-SSRI or RELION 3.0. The dashed black line represents FSC = 0.143. (*B*) Final reconstructed cryo-EM map using RELION 3.0. (*C*) Final reconstructed cryo-EM map using OPUS-SSRI. The red rectangles in *B* and *C* define a region of the EM map to be enlarged in *D* for RELION 3.0 and *E* for OPUS-SSRI for residues 374 to 409, respectively. The dash red circles highlight a region in the model before (*D*) and after (*E*) the manual adjustments in COOT and structural refinement using PHENIX. The EM density is represented in mesh (blue), and the structural model is represented by a ribbons diagram with side chains in stick presentation. Both density maps are contoured at the same level.

consistent with their neighbors, the sparseness and smoothness restraints can suppress the noisy densities that do not belong to the molecules in the map, thus producing cleaner reconstructions. The cleaner map in turn leads to more accurate pose estimation for each particle. These improvements brought about by our method result in an overall much-improved final reconstruction. Furthermore, the relatively large improvements for structures with heterogenous flexibility such as Hrd1/Hrd3 and TMEM16A confirm the theoretical difference between the traditional smoothness prior in RELION and our smoothness prior in OPUS-SSRI. For structures with heterogeneous flexibilities in different regions, the traditional approach in RELION enforces translation-invariant isotropic smoothness to the 3D maps, thus smearing the rigid regions and creating large biases in the reconstructions. In contrast, OPUS-SSRI can adapt to different flexibilities in different regions in the maps, thus greatly reducing biases and improving the final reconstructions. Another approach we explored to promote smoothness is by casting the back-projection as a local kernel regression problem. This formulation enables us to embed the 3D maps in a reproducing kernel Hilbert space (RKHS) with specific smoothness.

Although our method introduces five more parameters, their optimal values can be easily determined. First of all, we can set ϵ to the level of density values corresponding to molecular content in the 3D volume. This level can be easily obtained from the intermediate volumes generated by the refinement using RELION 3.0. The optimal values of α, β, γ, and ϵ' can be found by grid search as detailed in *SI Appendix, Experiment process* and exemplified in *SI Appendix,* Fig. S4. The parameters were searched in the order $\alpha \to \beta \to \gamma \to \epsilon'$. The complexity of the grid search depends linearly on the number of parameters. The ranges of parameters where optimal settings were found in our tests are summarized as follows: α are in the range of $[0.4,\ 0.8]\overline{\|A^H Dy\|_2}\epsilon$, β are in the range of $[0.5,\ 4]\overline{\|A^H Dy\|_2}\epsilon$, γ are in the range of $[0.05,\ 0.2]\overline{N(n)}$, and ϵ' are in the range of $\left[\frac{1}{3},\ 2\right]\epsilon$.

These ranges can serve as useful guidance for future applications of OPUS-SSRI to other systems. In practice, a reasonable set of parameters can be obtained with just a few trials.

It is worth noting that OPUS-SSRI focuses on improving accuracies of pose parameters for each particle in the maximization step, which is complementary to the approach explored by THUNDER that targets other latent variables, such as defocus parameters and class membership. Hence, these two approaches can be readily combined. In fact, accurate determination of pose parameters are the prerequisites for a better per-particle defocus
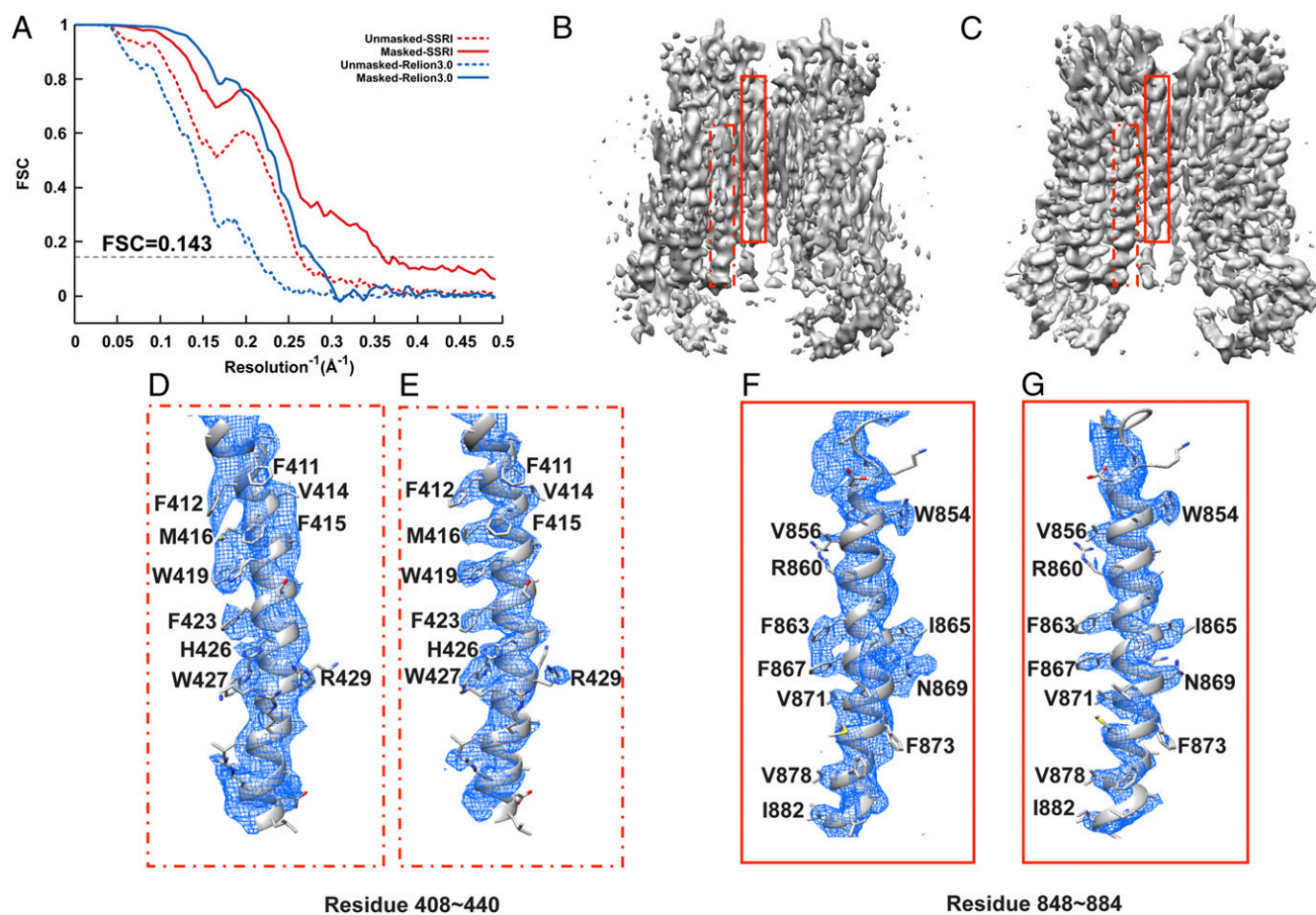
Fig. 4. Refinement of TMEM16A in nanodisc. (*A*) Gold-standard unmasked and masked FSC curves calculated from two independent reconstructions by OPUS-SSRI or RELION 3.0. The dashed black line represents FSC = 0.143. (*B*) Final reconstructed cryo-EM map using RELION 3.0. (*C*) Final reconstructed cryo-EM map using OPUS-SSRI. The dashed red rectangles in *B* and *C* define a region of EM map to be enlarged in *D* for RELION 3.0 and *E* for OPUS-SSRI for residues 408 to 440, respectively. The solid red rectangles in *B* and *C* define a region of the EM map to be enlarged in *F* for RELION 3.0 and *G* for OPUS-SSRI for residues 848 to 884, respectively. The EM density is represented in mesh (blue), and the structural model is represented by a ribbons diagram with side chains in stick presentation. Both density maps are contoured at the same level.

parameter refinement. This is exemplified by the limited improvement of THUNDER on the highly noisy dataset Hrd1/Hrd3, in which the pose of each particle was of large errors (*SI Appendix*, Figs. S2 and S3), yielding inaccurate reference two-dimensional (2D) projections and adversely affecting the per-particle contrast transfer function (CTF) refinement. Therefore,

Table 1. Comparison of the final reconstructions refined by RELION 3.0, THUNDER, or OPUS-SSRI

	RELION	THUNDER		OPUS-SSRI		
	Gold-standard FSC = 0.143					
Proteins	Resolution (Å)	Resolution (Å)	ΔÅ over RELION*	Resolution (Å)	ΔÅ over RELION*	ΔÅ over THUNDER†
β-galactosidase (EMPIAR-10017)	4.16	4.25	−0.09	3.93	0.23	0.33
80S ribosome (EMPIAR-10002)	4.08	3.80	0.28	3.93	0.15	−0.13
Hemagglutinin (EMPIAR-10097)	4.19	4.11	0.08	3.77	0.42	0.34
TRPM4 (EMPIAR-10126)	3.48	/	/	2.74	0.74	/
Hrd1/Hrd3 (EMPIAR-10099)	4.80	4.75	0.05	3.55	1.25	1.20
TRPV5 (EMPIAR-10254)	3.12	3.09	0.03	2.47	0.65	0.62
TMEM16A (EMPIAR-10123)	3.90	/	/	2.84	1.06	/
Average improvement			0.07		0.64	0.47

/ indicates that the comparison was unavailable in two cases in which THUNDER failed to execute due to computer incompatibility.
*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.
†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.

Luo et al.
Sparseness and Smoothness Regularized Imaging for improving the resolution of Cryo-EM single-particle reconstruction

PNAS | 5 of 8
https://doi.org/10.1073/pnas.2013756118

our OPUS-SSRI might enhance the per-particle CTF refinement on some noisy datasets by improving the pose determination of these datasets.

Finally, our tests of OPUS-SSRI on seven real datasets support that OPUS-SSRI can greatly improve the resolution of the final density map, thus allowing more accurate building of atomic models. We expect OPUS-SSRI to be an invaluable tool to the general field of cryo-EM single-particle analysis.

Materials and Methods

We clarify some notations here. For a vector $x \in \mathbb{R}^N$, we use $\|x\|_p$ to represent the l_p norm of the vector x, which is defined as $\|x\|_p = (\sum_{i=1}^{N} |x_i|^p)^{1/p}$. $\mathbb{Z}$ represents the set of integers, while $\mathbb{R}$ stands for the set of reals. We use x to represent the 3D map, V to represent the Fourier transform (FT) of x, and X_i to represent the FT of the ith image. For simplicity, we often use vector representation for multidimensional data (i.e., a point with multidimensional index $[i, \ldots, k]$ is mapped to the hth component x_h of x).

Introduction to Cryo-EM Refinement. Formally, the FT of 3D map V to be reconstructed in cryo-EM refinement can be defined as the maximizer of the penalized log marginal likelihood function (3) $\sum_{i=1}^{N} \log P(X_i|V) - J(x)$, where $P(X_i|V)$ is the marginal probability (see *SI Appendix, Log marginal likelihood* for derivation) of X_i given map V by marginalizing over all hidden variables, such as orientation and translation (pose), and $J(x)$ is the penalty for x, which serves to reduce overfitting and guarantee the feasibility of solution.

As 3D molecular maps are both sparse and smooth, in order to incorporate these priors into refinement, a mathematical formulation for them must be developed. Conventionally, the smoothness of a function is associated with the norm of its gradient, and sparseness is referred to as the number of zeros in the values of function (20). In the following subsections, we will formulate different smoothness priors and reveal their differences. The key equations illustrating the effects of the traditional smoothness restraint and our smoothness restraint are Eqs. **2** and **5**, respectively.

Traditional Smoothness Prior. The traditional method (5) enforces smoothness by applying a quadratic restraint on the magnitudes of FTs based on the assumption that they are distributed according to Gaussian. Since the traditional method is an instance of Wiener filtering (21), the restraint strength depends on the SNR. The 3D map reconstructed by the traditional method can be defined as the maximizer of

$$\sum_{i=1}^{N} \log P(X_i|V) - \sum_{hkl} \overline{\frac{N(r)}{SNR(r)}} |V_{hkl}|^2, \qquad [1]$$

where $r = \sqrt{h^2 + k^2 + l^2}$ is the modulus of the spatial frequency vector of the Fourier coefficient V_{hkl}, $\overline{N(r)}$ is the average of weights $N([h, k, l])$ in Eq. **8**, and $SNR(r)$ is the SNR estimated by gold-standard FSC (22).

To understand the effect of the smoothness restraint of the traditional method, we consider the role of the restraint in the gradient ascent iteration, which is of the form

$$V'_{hkl} = V_{hkl} + \eta \left(\sum_{i=1}^{N} \frac{\partial}{\partial V_{hkl}} \log P(X_i|V) - \lambda(r) V_{hkl} \right), \qquad [2]$$

where $\lambda(r) = \overline{\frac{N(r)}{SNR(r)}}$ is the damping weight for V_{hkl}, and $\eta > 0$ is the learning rate. The gradient of the traditional prior $\lambda(r) V_{hkl}$ can be viewed as a smoothed map in real space. By convolution theorem, let the inverse FT of $\lambda(r)$ be $K(\|u\|_2)$, which is a radial function. The gradient of traditional prior $x'(u)$ is a convolution in real space, namely, $x'(u) = \int_{\mathbb{R}^3} K(\|u - v\|_2) x(v) dv$, which represents a map smoothed by the translation-invariant isotropic kernel $K(\|u - v\|_2)$. At each step, the old solution is modified by a linear combination of the gradient of log marginal likelihood function and the radial-kernel–smoothed old solution. We thus postulate that the traditional method in RELION biases the solution toward the 3D map with homogeneous smoothness across space.

Sparseness and Smoothness Priors in OPUS-SSRI. Sparseness resembles the idea of masking in the calculation of masked FSC, where the voxels which are below a certain threshold are setting to 0. The similar effects can be achieved by restraining the sum of the absolute values of densities, namely, the l_1 norm

of the density map (7) during reconstruction. Hence, we can encode the sparseness using the l_1 norm (7) and the smoothness of 3D map using the TV norm (8). Though these two priors can effectively guarantee both sparseness and smoothness of the map, they also introduce biases to the final solution (23). In particular, l_1 regularization tends to underestimate true nonzero elements (23) since the corresponding soft-threshold operator shrinks the volume by a global threshold. Fan et al. discovered that nonconcave penalty can prevent true nonzero elements from being overly shrunk while preserving sparseness (23). Therefore, we employ a nonconcave penalty log norm (24) to reduce biases in the solution. The penalized log likelihood function in OPUS-SSRI has the form

$$\sum_{i=1}^{N} \log P(X_i|V) - \sum_{j=1}^{L} (\alpha \log(|x_j| + \epsilon) + \beta \log(\|\nabla x_j\|_2 + \epsilon')), \qquad [3]$$

where α and β are positive, and ϵ and ϵ' are positive parameters to guard against the singularity of logarithm function at zero.

Optimization Methods in OPUS-SSRI. This subsection presents the algorithm to optimize the penalized log likelihood in Eq. **3**. First, the log marginal likelihood function can be optimized by the expectation–maximization method (25) (see *SI Appendix, Expectation maximization* for derivation). The reconstruction process alternates between the expectation step in which the distribution of pose parameters for each particle is determined and the maximization step in which the 3D map is reconstructed. Secondly, to address the nonconcavity of log norm, we approximate the logarithm function by concave function and iteratively improve the approximation (24) at each maximization step (see *SI Appendix, Weighted approximation* for derivation). Lastly, to average 3D maps reconstructed in consecutive maximization steps, we consider leveraging implicit gradient ascent (26), which is a widely used technique to improve the stability of optimization method. The implicit gradient ascent restrains the Euclidean distance between the new solution and the 3D map of previous maximization step x^{k-1}. These choices yield a target function in Eq. **4** by optimizing which can improve Eq. **3** (24). Formally, at the $i + 1$th iteration of the kth maximization step, let the solution obtained in the previous iteration be x^i. We approximate Eq. **3** with the expected log likelihood (the first term in Eq. **4**) and the weighted l_1 and TV norms and define the solution at iteration $i + 1$ as the maximizer of the following equation:

$$\max_x \sum_{i=1}^{N} \sum_{\phi} -P(\phi|X_i, V_{k-1}) \left\| X_i - CTF_i P^{\phi} V \right\|_2^2 - \sum_{j=1}^{L} \left(\alpha \frac{|x_j|}{|x_j^i| + \epsilon} + \beta \frac{\|\nabla x_j\|_2}{\|\nabla x_j^i\|_2 + \epsilon'} \right) - \frac{\gamma}{2} \left\| x - x^{k-1} \right\|_2^2, \qquad [4]$$

where $P(\phi|X_i, V_{k-1})$ is the conditional probability of the pose parameters given the observation X_i and the map V_{k-1} from the previous maximization step, and $\gamma > 0$ is the weight of the implicit gradient ascent penalty. The effect of log norm becomes evident after converting it to the weighted norm in Eq. **4**. Reweighting each voxel yields a spatially varying threshold (Eq. **6**), which can reduce the bias of prior. Next, we will demonstrate that the gradient of TV norm enforces heterogenous smoothness to the 3D map, and the l_1 norm achieves sparseness by the soft-thresholding rule while presenting the optimization algorithm.

Though TV norm is nondifferentiable at zero, we can approximate its gradient by Nesterov smoothing (27). The approximate gradient of TV norm at a voxel $[i, j, k]$ (see *SI Appendix, Nesterov smoothed TV norm* for derivation) is of the form

$$\nabla f_\mu(x)_{ijk} = c_{(i,j,k)} x[i,j,k] - \sum_{a=1}^{3} c'_{(i, j, k) \pm \Delta_a} x[(i, j, k) \pm \Delta_a], \qquad [5]$$

where Δ_a is a 1×3 vector with 1 on the ath entry and zeros elsewhere,

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

$c'_{(i,j,k)-\Delta_a} = \|\nabla x[i, j, k]\|_\mu^{-1}$ and $c'_{(i,j,k)+\Delta_a} = \|\nabla x[(i, j, k) + \Delta_a]\|_\mu^{-1}$, where $\|\nabla x[i, j, k]\|_\mu = (\|\nabla x_{(i,j,k)}^i\| + \epsilon') \max(\|\nabla x[i, j, k]\|, \mu)$, and μ is the Nesterov smoothing parameter. The gradient of the smoothed TV norm at the voxel $[i, j, k]$ depends on the gradients at this voxel and its neighboring voxels $[(i, j, k) \pm \Delta_a]$. Specifically, the voxel on the direction with smaller gradient has a larger contribution to the gradient of TV norm. We can write the continuous form of Eq. **5** as $x'(u) = \int_{\mathbb{R}^3} K(u, v) x(v) dv$, where the kernel $K(u, v)$

depends on the voxel u and its neighboring voxel v simultaneously. Therefore, the weighted TV norm smooths the map by a spatially varying anisotropic kernel $K(u, v)$, which adapts to heterogenous smoothness in the 3D map.

The differentiable function with l_1 penalty can be optimized by a joint application of gradient ascent and soft-thresholding operator (28). Denote the implicit gradient-restrained expected log likelihood $-\sum_{i=1}^{N} P(\phi|X_i, V_{k-1})\left\|X_i - CTF_iP^\phi V\right\|_2^2 - \frac{\gamma}{2}\left\|x - x^{k-1}\right\|_2^2$ as $h(x)$ and the weight $\frac{1}{|x_j^i| + \epsilon}$ as w_j, and let the learning rate be η for the jth component of x at $i + 1$ iteration. Eq. 4 can be optimized by the following equation:

$$x_j^{i+1} = \begin{cases} 0, & |x_j^{i\prime}| < \eta\alpha w_j \\ x_j^{i\prime} - \eta\alpha w_j sign\left(x_j^{i\prime}\right), & |x_j^{i\prime}| \geq \eta\alpha w_j \end{cases}, \quad \text{[6]}$$

where $x^{i\prime} = x^i + \eta(\nabla h(x^i) - \beta\nabla f_\mu(x^i))$ is the 3D map updated by gradient ascent, and sign() extracts the sign of a voxel value. Eq. 6 is referred to as the soft-thresholding operator. The sparseness of the volume is preserved by the soft-thresholding operator since it sets the voxels with relatively small values to zeros. Moreover, the threshold $\eta\alpha w_j$ is inversely proportional to the strength of the voxel. Therefore, the voxels with small values are more likely to be shrunk to 0 s. Since the values of electron densities for molecules are often higher than the background noises, this soft-thresholding operator can suppress noises while leaving the electron densities of molecules intact. We hence demonstrate how our l_1 restraint yields an unbiased cleaner 3D map.

In summary, Eq. 6 is applied iteratively in the maximization step. The gradient of the TV norm enforces spatially varying smoothness in gradient ascent, while the soft-thresholding operator induced by the weighted l_1 norm guarantees the sparseness of the 3D map. In all of our experiments, convergences were achieved in a maximization step after 100 iterations.

Back-Projection as Local Kernel Regression. To reconstruct the 3D map, cryo-EM researchers introduced a back-projection operator, which puts the 2D FT of the image into the 3D map. As the inverse of slice operator P^ϕ (*SI Appendix*, Eq. S2), back-projection puts the data with 2D index $[i, j]$ to the 3D index $n(\phi)$ according to $n(\phi) = R_\phi^{-1}[i, j, 0]^T$, where R_ϕ^{-1} is the inverse of the rotation matrix parameterized by orientation parameter ϕ. Back-projection prompts us to create a 3D volume to store the data on nonintegral indices $[h, k, l] \in \mathbb{R}^3$ if the orientation ϕ is randomly sampled, while the indices of the 3D map operated in computer memory are integral, namely, $[h, k, l] \in \mathbb{Z}^3$. Therefore, the limitation of computer memory prevents us from storing the voxel of the back-projected 2D slice $P^{-\phi}X_i$ in the 3D map V. Nonparametric regression can be leveraged to eliminate this discrepancy by storing weighted back-projected data on integral grids (29). Using local kernel regression (see *SI Appendix, Local kernel regression* for derivation), the expected log likelihood can be written as

$$-\sum_{hkl}\sum_{i=1}^{N}\sum_{j=1}^{J}\sum_{\phi}P(\phi|X_i, V_{k-1})K(n_j(\phi), n)\left\|X_{ij} - CTF_{ij}V\right\|_n^2, \quad \text{[7]}$$

where K is the kernel, $n_j(\phi)$ is a voxel back-projected with parameter ϕ, and $n = [h, k, l] \in \mathbb{Z}^3$ is the voxel storing the kernel-weighted back-projected data. Hence, in contrast to the first term in Eq. 4, this new formulation decouples the slice operator from V and has variables V_{hkl} on integral grids.

OPUS-SSRI used Gaussian kernel, which is of the form $K(n_j(\phi), n) = \exp\left(-\frac{(h_j-h)^2+(k_j-k)^2+(l_j-l)^2}{\sigma^2}\right)$. We set the bandwidth σ of the Gaussian kernel to be $\sqrt{2}$ in our implementation of OPUS-SSRI. For practical consideration, a Fourier coefficient $n_j(\phi)$ cannot be scattered to every voxel. Since the Gaussian kernel is quickly diminishing, we let $[\![\cdot]\!]$ be the round operator, and we consider storing X_{ij} with the index $[h_j, k_j, l_j]$ on the neighboring voxels $[h_j + \Delta_h, [\![k_j]\!] + \Delta_k, [\![l_j]\!] + \Delta_l]$, where $\Delta_a = \pm 1$ or 0. In the original implementation of RELION (5), they chose a kernel with trilinear interpolation–like weight. Using the formalism introduced before, their kernel functions are of the form $K(n_j(\phi), n) = w(h_j, h)w(k_j, k)w(l_j, l)$ with $w(x_j, x) = \begin{cases} \Delta x + 1 - x_j, & x \leq x_j \\ \Delta x_j + 1 - x, & x > x_j \end{cases}$, and $x \in [\lfloor x_j \rfloor, \lceil x_j \rceil]$, where $\lfloor \cdot \rfloor$ and $\lceil \cdot \rceil$ are the floor and ceiling operator, respectively. Hence, a Fourier coefficient is

scattered to eight neighboring voxels in the original implementation of RELION (5).

Eq. 7 has a closed form solution,

$$V_n = \frac{\sum_{i=1}^{N}\sum_{j=1}^{J}\sum_{\phi}P(\phi|X_i, V_{k-1})K(n_j(\phi), n)CTF_{ij}X_{ij}}{N(n)}, \quad \text{[8]}$$

with $N(n) = \sum_{i=1}^{N}\sum_{j=1}^{J}\sum_{\phi}P(\phi|X_i, V_{k-1})K(n_j(\phi), n)CTF_{ij}^2$. Ignoring regularization terms, local kernel regression estimates the value of a voxel $[h, k, l]$ by an integral between the kernel K and neighboring experimental data $V_{h'k'l'} = P(\phi|X_i, V_{k-1})CTF_{ij}X_{ij}$, $V_{hkl} = \Pi \int_{h'k'l'} K(n, n')V_{h'k'l'}dh'dk'dl'$, where Π is the normalizing factor. This convolution projects V'_{hkl} into the RKHS associated with the given kernel according to the representer theorem (9), thus enforcing V_{hkl} to exhibit the smoothness of the RKHS induced by that kernel.

OPUS-SSRI Implementation. The implementation of OPUS-SSRI is based on RELION. The 3D refinement program in RELION consists of two modules, expectation and maximization. We implemented our method as a new routine in its maximization module. Therefore, when performing refinement, the expectation steps of RELION (5) and OPUS-SSRI use exactly the same settings. The gradient calculation and soft-thresholding operators are implemented with CUDA, thus allowing fast maximization.

Gold-Standard FSC. The gold-standard FSC is the FSC between two independently refined half maps F and G (22). The gold-standard FSC of Fourier coefficients at shell k relates to the SNR through $FSC(k) = \frac{SNR(k)}{SNR(k)+1}$, as noises inside two maps are independent (30). The frequency where the gold-standard FSC passes through 0.143 is denoted as the estimated resolution of the reconstruction. A high-resolution noise substitution method is often used to correct effects of masking in the masked FSC (30).

Model versus Map FSC. If there exists a high-resolution atomic structural model, we can validate the cryo-EM map by comparing it to this atomic model. The first step in calculating the model versus map FSC is fitting the atomic model into the cryo-EM density map. The model map is constructed from the fitted atomic model by sampling on the same grid as the experimental map. The model versus map FSC (31) is the correlation between the FT of the model map and the FT of cryo-EM map. The point where the model versus map FSC approaches 0.143 can be regarded as the resolution of the experimental map.

Refinement Protocol. The single-particle datasets used in this paper were obtained from either the deposited particle stack or the coordinate files. In all experiments, we built the initial maps ab initio in RELION 3.0 and refined those initial maps using the three methods to be compared. The initial map building began with one round of 2D classification in RELION 3.0. The particles belonging to the major classes were then selected to build the initial map ab initio using the 3D classification procedure in RELION 3.0. The same low-pass–filtered initial maps were subsequently supplied into the three methods, RELION 3.0, THUNDER, and OPUS-SSRI, for refinement. For the datasets with specific symmetry, the symmetry was enforced throughout the refinement process. For RELION 3.0 and OPUS-SSRI, we also used the same convergence criteria [i.e., no resolution improvement and pose changes for the last two iterations (5)]. In THUNDER, the particle grading and CTF search options were set as "True" for better results. Finally, the gold-standard FSC calculations and density map postprocessing of the refinement results of all methods were carried out in RELION 3.0. In the postprocessing step, the mask was created from the final reconstruction using all particles in the 3D refinement procedure. Using *relion_postprocess* (30), we obtained gold-standard FSCs and the postprocessed map from independent maps by correcting the modulation transfer function of the detector and sharpening with automatically estimated B-factors. We then compared the postprocessed map with respect to the corresponding published atomic model(s) by calculating model versus map FSC using *Phenix.Mtriage* (32). Before comparison, the atomic model was fitted into the postprocessed density maps reconstructed by different methods using the rigid-body fit in *Chimera* (33).

Data Availability. The implementation of OPUS-SSRI can be found at GitHub (https://github.com/alncat/cryoem). All other study data are included in the article and supporting information.

Luo et al.
Sparseness and Smoothness Regularized Imaging for improving the resolution of Cryo-EM single-particle reconstruction

PNAS | 7 of 8
https://doi.org/10.1073/pnas.2013756118

1. R. M. Glaeser, Limitations to significant information in biological electron microscopy as a result of radiation damage. *J. Ultrastruct. Res.* **36**, 466–482 (1971).
2. R. M. Glaeser, How good can cryo-EM become? *Nat. Methods* **13**, 28–32 (2016).
3. S. H. Scheres, A Bayesian view on cryo-EM structure determination. *J. Mol. Biol.* **415**, 406–418 (2012).
4. C. V. Sindelar, N. Grigorieff, An adaptation of the Wiener filter suitable for analyzing images of isolated single particles. *J. Struct. Biol.* **176**, 60–74 (2011).
5. S. H. Scheres, RELION: Implementation of a bayesian approach to cryo-EM structure determination. *J. Struct. Biol.* **180**, 519–530 (2012).
6. M. Hu *et al.*, A particle-filter framework for robust cryo-EM 3D reconstruction. *Nat. Methods* **15**, 1083–1089 (2018).
7. R. Tibshirani, Regression shrinkage and selection via the lasso. *J. R. Stat. Soc. B* **58**, 267–288 (1996).
8. L. I. Rudin, S. Osher, E. Fatemi, Nonlinear total variation based noise removal algorithms. *Phys. Nonlinear Phenom.* **60**, 259–268 (1992).
9. B. Schölkopf, R. Herbrich, A. J. Smola, "A generalized representer theorem" in *International Conference on Computational Learning Theory*, D. Helmbold, B. Williamson, Eds. (Springer, 2001), pp. 416–426.
10. S. H. Scheres, Semi-automated selection of cryo-EM particles in RELION-1.3. *J. Struct. Biol.* **189**, 114–122 (2015).
11. K. R. Vinothkumar, G. McMullan, R. Henderson, Molecular mechanism of antibody-mediated activation of β-galactosidase. *Structure* **22**, 621–627 (2014).
12. X. Bai, I. S. Fernandez, G. McMullan, S. H. Scheres, Ribosome structures to near-atomic resolution from thirty thousand cryo-EM particles. *elife* **2**, e00461 (2013).
13. Y. Z. Tan *et al.*, Addressing preferred specimen orientation in single-particle cryo-EM through tilting. *Nat. Methods* **14**, 793–796 (2017).
14. H. E. Autzen *et al.*, Structure of the human TRPM4 ion channel in a lipid nanodisc. *Science* **359**, 228–232 (2018).
15. S. Schoebel *et al.*, Cryo-EM structure of the protein-conducting ERAD channel Hrd1 in complex with Hrd3. *Nature* **548**, 352–355 (2017).
16. S. Dang *et al.*, Structural insight into TRPV5 channel function and modulation. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 8869–8878 (2019).
17. S. Dang *et al.*, Cryo-EM structures of the TMEM16A calcium-activated chloride channel. *Nature* **552**, 426–429 (2017).
18. P. Emsley, K. Cowtan, Coot: Model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004).
19. P. D. Adams *et al.*, PHENIX: A comprehensive python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 (2010).
20. T. Hastie, R. Tibshirani, M. Wainwright, *Statistical Learning with Sparsity: The Lasso and Generalizations* (CRC press, 2015).
21. E. J. Candes, Modern statistical estimation via oracle inequalities. *Acta Numer.* **15**, 257–325 (2006).
22. S. H. Scheres, S. Chen, Prevention of overfitting in cryo-EM structure determination. *Nat. Methods* **9**, 853–854 (2012).
23. J. Fan, R. Li, Variable selection via nonconcave penalized likelihood and its oracle properties. *J. Am. Stat. Assoc.* **96**, 1348–1360 (2001).
24. E. J. Candes, M. B. Wakin, S. P. Boyd, Enhancing sparsity by reweighted ℓ 1 minimization. *J. Fourier Anal. Appl.* **14**, 877–905 (2008).
25. A. P. Dempster, N. M. Laird, D. B. Rubin, Maximum likelihood from incomplete data via the EM algorithm. *J. R. Stat. Soc. B* **39**, 1–22 (1977).
26. P. Toulis, "Implicit methods for iterative estimation with large data sets," Doctoral dissertation, Graduate School of Arts & Sciences, Harvard University, Cambridge, MA (2016).
27. Y. Nesterov, Smooth minimization of non-smooth functions. *Math. Program.* **103**, 127–152 (2005).
28. N. Parikh, S. Boyd, Proximal algorithms. *Found. Trends Optim.* **1**, 127–239 (2014).
29. H. Takeda, S. Farsiu, P. Milanfar, Kernel regression for image processing and reconstruction. *IEEE Trans. Image Process.* **16**, 349–366 (2007).
30. S. Chen *et al.*, High-resolution noise substitution to measure overfitting and validate resolution in 3D structure determination by single particle electron cryomicroscopy. *Ultramicroscopy* **135**, 24–35 (2013).
31. C.-I. Bränd'en, T. Alwyn Jones, Between objectivity and subjectivity. *Nature* **343**, 687–689 (1990).
32. P. V. Afonine *et al.*, New tools for the analysis and validation of cryo-EM maps and atomic models. *Acta Crystallogr. D Struct. Biol.* **74**, 814–840 (2018).
33. E. F. Pettersen *et al.*, UCSF Chimera–A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).

8 of 8 | PNAS
https://doi.org/10.1073/pnas.2013756118

Luo et al.
Sparseness and Smoothness Regularized Imaging for improving the resolution of Cryo-EM single-particle reconstruction