Supporting Information

Baker et al. 10.1073/pnas.1204935109

SI Results

Test to Determine the Accuracy of Particle Projection Matching. In single particle electron cryomicroscopy (cryo-EM), the orientation of the protein particles in the specimen must be determined relative to a 3D reference map. These orientations can be described by three Euler angles (ϕ, θ, and ψ) and two translations. Electron microscope images correspond to projections of the 3D Coulomb potential of the specimen along the direction of the electron beam. Therefore, the orientations of the protein particles in images can be estimated by projection matching, where experimental images are compared with projections of the 3D reference map of the protein complex. Initial attempts to improve maps of bovine ATP synthase by projection matching were unsuccessful, which we hypothesized to be caused by an inability to accurately determine the direction of projection, also known as the alignment, for individual particle images. Although the accuracy with which the projection direction for an individual experimental particle image has been determined cannot be assessed directly, methods have been developed to estimate the accuracy of alignment using pairs of images containing the same particles, with a rotation of the microscope goniometer applied between images (1). The microscope goniometer rotates, or "tilts," the specimen about an axis that is perpendicular to the electron beam, and consequently, these pairs of images are often referred to as tilted pairs. One such approach, the free hand test (1), consists of performing projection matching on particle images in the first micrograph and plotting the agreement of particle images in the second micrograph against all possible goniometer rotations. Other implementations use projection matching to estimate the projection direction for particles from both micrographs and calculate the goniometer rotations that best map the alignment of the first images onto the alignment of the second images (1, 2). Here, we develop a related test, referred to as the tilted pair alignment (TPA) test. In this test, the Euler angles for both particle images from the tilted pair are estimated by projection matching. These Euler angles are used to generate a 3 × 3 rotation matrix for each particle image, which is shown below for the z, $-y$, z Euler angle convention (Eq. **S1**):

$$R_{\psi\theta\phi} = \begin{bmatrix} \cos\psi\cos\theta\cos\phi - \sin\psi\sin\phi & \cos\psi\cos\theta\sin\phi + \sin\psi\cos\phi & -\cos\psi\sin\theta \\ -\sin\psi\cos\theta\cos\phi - \cos\psi\sin\phi & -\sin\psi\cos\theta\sin\phi + \cos\psi\cos\phi & \sin\psi\sin\theta \\ \sin\theta\cos\phi & \sin\theta\sin\phi & \cos\theta \end{bmatrix}. \quad \text{[S1]}$$

Given R_1 and R_2 as the rotation matrices describing the orientations of the particles in the first and second images in a tilted pair, respectively, an equation that relates the rotation matrices to the goniometer rotation can be written as (Eq. **S2**)

$$R_2 = R_\varepsilon R_g R_1, \quad \text{[S2]}$$

where R_g is the rotation applied by the microscope goniometer and R_ε is a matrix that describes any rotation between R_1 and R_2 that is not explained by R_g. The goniometer rotation, R_g, can be described as a combination of rotations about the x axis, rot_x and y axis, rot_y in the plane of the EM grid. The overall rotation magnitude, α, is given by (Eq. **S3**)

$$\alpha = \left(rot_x{}^2 + rot_y{}^2\right)^{1/2}, \quad \text{[S3]}$$

and the angle of the rotation axis with respect to the x axis, γ is $arctan(rot_y/rot_x)$ (1). R_g may be written in matrix form as (Eq. **S4**)

$$R_g = \begin{bmatrix} \cos(\alpha) + \sin(\gamma)^2(1-\cos(\alpha)) & \cos(\gamma)\sin(\gamma)(1-\cos(\alpha)) & \cos(\gamma)\sin(\alpha) \\ \cos(\gamma)\sin(\gamma)(1-\cos(\alpha)) & \cos(\alpha) + \cos(\gamma)^2(1-\cos(\alpha)) & -\sin(\gamma)\sin(\alpha) \\ -\cos(\gamma)\sin(\alpha) & \sin(\gamma)\sin(\alpha) & \cos(\alpha) \end{bmatrix}. \quad \text{[S4]}$$

The rotation matrix for R_ε can be written in terms of a rotation, ε, about an axis given by the vector (u_x, u_y, u_z) (3) (Eq. **S5**):

$$R_\varepsilon = \begin{bmatrix} \cos\varepsilon + u_x^2(1-\cos\varepsilon) & -u_z\sin\varepsilon + u_xu_y(1-\cos\varepsilon) & u_y\sin\varepsilon + u_xu_z(1-\cos\varepsilon) \\ u_z\sin\varepsilon + u_xu_y(1-\cos\varepsilon) & \cos\varepsilon + u_y^2(1-\cos\varepsilon) & -u_x\sin\varepsilon + u_yu_z(1-\cos\varepsilon) \\ -u_y\sin\varepsilon + u_xu_z(1-\cos\varepsilon) & u_x\sin\varepsilon + u_yu_z(1-\cos\varepsilon) & \cos\varepsilon + u_z^2(1-\cos\varepsilon) \end{bmatrix}. \quad \text{[S5]}$$

The magnitude of the rotation described by R_ε is given by (4) (Eq. **S6**)

$$\varepsilon = \operatorname{acos}\left(\frac{1}{2}trace(R_\varepsilon) - \frac{1}{2}\right). \quad \text{[S6]}$$

Thus, tilted pairs in which the estimated particle orientations are not well-explained by the goniometer rotation have an ε-value that is large, whereas orientations that are perfectly explained by the goniometer rotation would have an ε-value of zero. In the terminology in the work by Henderson et al. (2), ε corresponds to the distance of the second image from its expected position in a tilt pair parameter plot.

It is important to note that the TPA test provides a conservative estimate of alignment accuracy: the second image is acquired after prior irradiation of the specimen, and the accuracy with which the particle orientation can be determined from that image is expected to be worse than from an initial image. Also, for a tilted pair to produce a low ε-value, the particle orientation must be determined accurately from both images. The TPA test does not provide a way to distinguish pairs where one image has not been accurately aligned from pairs where both images have not been accurately aligned. If the alignment of the images in a pair is considered to be independent, the square root of the proportion of tilted pairs with ε-values below a selected threshold gives a conservative estimate of the proportion of individual images that can be aligned with errors less than ε.

Fig. S2*A* shows a histogram of the alignment errors from the TPA test for a set of 216 tilted pairs using the same projection matching algorithm used in initial attempts at map refinement. As seen in Fig. S2*A*, blue histogram, the majority of the tilted pairs aligned with extremely large angular errors, with a distribution of errors that was only slightly better than would be expected from random alignments (Fig. S2*A*, black line). Because this prevalence of large alignment errors was not seen when the TPA test was applied to an independent map and set of tilted pairs of the V/A-type ATPase from *Thermus thermophilus* (5, 6), it suggested that the poor alignment accuracy was responsible for degrading the map on refinement. As described in *Results*, the TPA test was used to guide the selection of alignment conditions for untilted data that could be used to improve the map. On selection of appropriate conditions, projection matching produced a distribution of alignment errors that was significantly better than random (Fig. S2*A*, red histogram).

The reasons for the poor alignment accuracy cannot be identified unambiguously. Although conformational heterogeneity of ATP synthase particles could contribute to the problem, there are several other possible explanations for the poor alignment accuracy. At less than 600 kDa, bovine ATP synthase is smaller than many other specimens studied by single particle cryo-EM, and consequently, images of ATP synthase contain less signal than images from larger specimens. Analysis of alignment accuracy using tilted image pairs, although performed for only a few different specimens, suggested that smaller protein particles are more difficult to align accurately than larger protein particles (2). A particular problem for ATP synthase is that the mass in the complex is approximately evenly distributed between the F_1 and F_O regions, making it likely that these regions are misidentified during alignment. Under some conditions, close to 15% of our tilted pairs were aligned with one image assigned an in-plane rotation angle (ψ) that was nearly 180° different from the angle assigned to its pair. Given this statistic, it is likely that, in a proportion of pairs, both images had similar but incorrect ψ-angles. Additionally, the quasi-symmetry in the F_1 region likely affects the alignment of images of bovine ATP synthase. The structures of the α- and β-subunits are similar at low resolution, unlike the structures of the related B- and A-subunits from the A/V-type ATP synthase from *T. thermophilus* that can be distinguished by the presence of a non-homologous domain (5), increasing the likelihood of errors in the estimation of the ϕ-angle by projection matching.

Exclusion of Misaligned Images. After the final iteration of refinement with untilted data, particle images were excluded from the map using three criteria. First, a map of the background noise and diffuse density from misaligned particle images was created by subtracting the map after tight masking of the protein, detergent, and lipid density from the overall map. Any particle image that had an amplitude-weighted phase residual with the background map at its assigned orientation that was better than 93° was excluded from the calculation of the final map. Second, any image with a SD over the last three iterations of alignment in assigned ϕ- or ψ-angles, rms translation, or phase residual that was greater than 5° or 3°, 0.3 Å, and 0.3°, respectively, was excluded from the calculation of the final map. Finally, any image with an amplitude-weighted phase residual against the penultimate map at the assigned orientation that was worse than 57° was excluded. As judged by Fourier shell correlation and the features visible in the map, the map was optimal when 10,461 of the initial 57,885 images (comprising ~18% of the dataset) were included in the final map. Whereas this data exclusion method was applied to all of the untilted images to improve the 3D map, it could also be applied to the tilted images pairs to show that misaligned images are more likely to be excluded than aligned images (Fig. S2*B*). Slices through maps built with all particle images, with the 10,461 systematically selected particle images and 10,461 randomly selected particle images, are shown in Fig. S2*C*.

1. Rosenthal PB, Henderson R (2003) Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. *J Mol Biol* 333:721–745.
2. Henderson R, et al. (2011) Tilt-pair analysis of images from a range of different specimens in single-particle electron cryomicroscopy. *J Mol Biol* 413:1028–1046.
3. Euler L (1776) Formulae generales pro translatione quacunque corporum rigidorum [General formulas for the translation of arbitrary rigid bodies]. *Novi Commentarii Academiae Scientiarum Petropolitanae 20*, pp 189–207.
4. Santoro A, Mighell AD (1973) Coincidence-site lattices. *Acta Crystallogr A* 29: 169–175.
5. Lau WC, Rubinstein JL (2010) Structure of intact Thermus thermophilus V-ATPase by cryo-EM reveals organization of the membrane-bound V(O) motor. *Proc Natl Acad Sci USA* 107:1367–1372.
6. Lau WCY, Rubinstein JL (2012) Subnanometer-resolution structure of the intact *Thermus thermophilus* H+-driven ATP synthase. *Nature* 481:214–218.

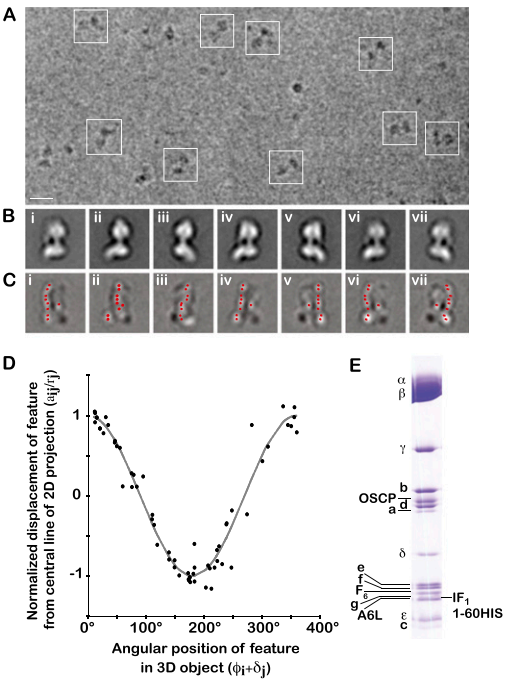

Fig. S1. Cryo-EM of the bovine F-type ATP synthase. (*A*) A representative micrograph shows particles of protein (examples outlined with white rectangles), mostly presenting side views of the complex. (Scale bar: 200 Å.) (*B*) Class average images showing the F_1 and F_O regions. Rotational analysis (1–4) was used to determine the viewing angle for class averages in *B* and produce an initial 3D map of the complex. (*C*) The positions of asymmetric features in the class averages were identified from difference maps between each class average and the average of all other class averages. The features used for rotational analysis are marked with red circles. (*D*) The displacements (points) of these features in the 2D class averages show good agreement with an ideal single-axis rotation series in 3D (line). (*E*) A 12–22% polyacrylamide SDS/PAGE gel of the preparation of bovine ATP synthase used for imaging.

1. Lau WC, Rubinstein JL (2010) Structure of intact Thermus thermophilus V-ATPase by cryo-EM reveals organization of the membrane-bound V(O) motor. *Proc Natl Acad Sci USA* 107: 1367–1372.
2. Baker LA, Rubinstein JL (2008) Angle determination for side views in single particle electron microscopy. *J Struct Biol* 162:260–270.
3. Lau WC, Baker LA, Rubinstein JL (2008) Cryo-EM structure of the yeast ATP synthase. *J Mol Biol* 382:1256–1264.
4. Rubinstein JL, Walker JE, Henderson R (2003) Structure of the mitochondrial ATP synthase by electron cryomicroscopy. *EMBO J* 22:6182–6192.

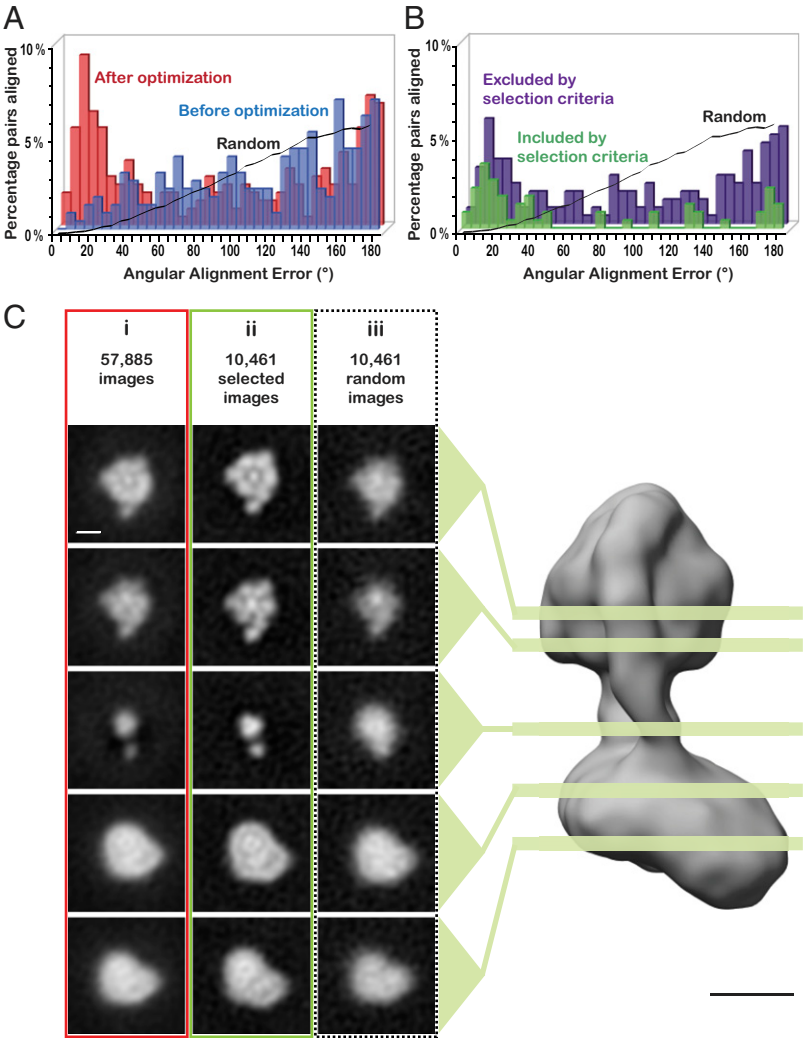

Fig. S2. Particle image alignment accuracy. (*A*) A histogram of the alignment errors for a set of 231 tilted pairs aligned with the conditions used in initial attempts at map refinement (blue). The majority of the tilted pairs aligned with extremely large angular errors. For comparison, the distribution of angular alignment errors for a set of 50,000 random alignments is plotted (black line). After optimization of alignment conditions, the alignment accuracy improved significantly (red). (*B*) Poorly aligning particle images were excluded from construction of the final map according to the three types of criteria described in *SI Results*. Application of these criteria to the first image in each pair enriched the included particles for well-aligning images (green histogram). However, some of the images excluded by these criteria (purple histogram) also aligned well. The histogram for randomly aligned images (black curve) is shown for comparison. (*C*) Systematically excluding particle images improves the map quality. Slices are shown through the map built with all images (i), with the 10,461 images after systematic exclusion of data (ii), and with 10,461 randomly selected images from the dataset (iii). The map built after systematic exclusion of poorly aligned particle images shows internal features, such as the $\alpha_3\beta_3$-hexamer in F_1 and the c-ring in F_O, more clearly than the other two maps. From top to bottom, each column of slices show the F_1 region ~50 Å from OSCP, the F_1 region close to the membrane, the central and peripheral stalks, the F_O region near the mitochondrial matrix, and the F_O region near the intermembrane space. Each slice is normal to the rotation axis of the c-ring. (Scale bar: 50 Å.)

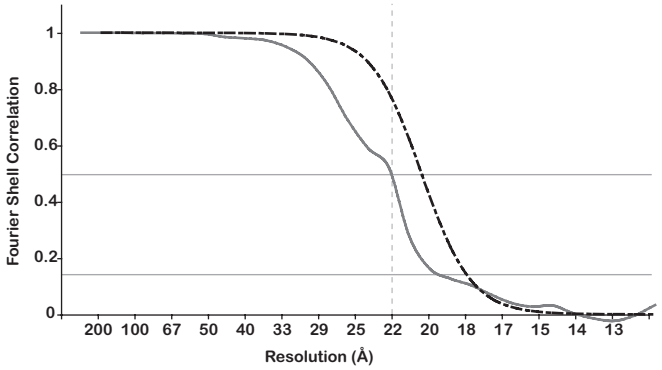

Fig. S3. Fourier shell and Fourier neighbor correlation. The resolution of the final map was estimated by Fourier shell correlation (FSC) (gray line) and supported by measurement of the Fourier neighbor correlation with Rmeasure (1) (dashed line). The 0.143 and 0.5 FSC thresholds are marked by horizontal gray lines. The highest spatial frequency used during alignment is indicated with a vertical dashed line.

1. Sousa D, Grigorieff N (2007) Ab initio resolution measurement for single particle structures. *J Struct Biol* 157:201–210.