Structure of β-galactosidase at 3.2-Å resolution obtained by cryo-electron microscopy

Alberto Bartesaghi[1], Doreen Matthies[1], Soojay Banerjee, Alan Merk, and Sriram Subramaniam[2]

Laboratory of Cell Biology, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892

Edited by Wah Chiu, Baylor College of Medicine, Houston, TX, and approved June 20, 2014 (received for review February 14, 2014)

We report the solution structure of *Escherichia coli* β-galactosidase (~465 kDa), solved at ~3.2-Å resolution by using single-particle cryo-electron microscopy (cryo-EM). Densities for most side chains, including those of residues in the active site, and a catalytic Mg^{2+} ion can be discerned in the map obtained by cryo-EM. The atomic model derived from our cryo-EM analysis closely matches the 1.7-Å crystal structure with a global rmsd of ~0.66 Å. There are significant local differences throughout the protein, with clear evidence for conformational changes resulting from contact zones in the crystal lattice. Inspection of the map reveals that although densities for residues with positively charged and neutral side chains are well resolved, systematically weaker densities are observed for residues with negatively charged side chains. We show that the weaker densities for negatively charged residues arise from their greater sensitivity to radiation damage from electron irradiation as determined by comparison of density maps obtained by using electron doses ranging from 10 to 30 $e^{-}/Å^2$. In summary, we establish that it is feasible to use cryo-EM to determine near-atomic resolution structures of protein complexes (<500 kDa) with low symmetry, and that the residue-specific radiation damage that occurs with increasing electron dose can be monitored by using dose fractionation tools available with direct electron detector technology.

single-particle EM | frame alignment | CTF determination | 3D reconstruction | structure refinement

Rapid advances in technology for single-particle cryo-electron microscopy (cryo-EM) over the last few years have made it possible to determine high-resolution structures of large and well-ordered macromolecular assemblies such as 2D protein crystals, helical lattices, icosahedral viruses, and protein complexes with high symmetry (1–10). More recently, continuing developments in microscope hardware and image processing software have yielded near-atomic resolution information for two smaller-sized complexes with low symmetry, the 700-kDa proteasome (11) and a 300-kDa mammalian ion channel (12), and for larger assemblies with no symmetry such as ribosome complexes (13–15). These developments signal an important change in the way cryo-EM can now be used in structural biology. Rather than simply using cryo-EM maps, typically in the 6- to 20-Å resolution range, as an envelope in which to fit structures obtained by X-ray crystallography, there is the exciting prospect of using cryo-EM to derive de novo, high-resolution structural models of proteins in one or multiple functional conformational states.

The preparation of specimens for cryo-EM involves rapid plunge freezing of a thin aqueous suspension into liquid ethane cooled by liquid nitrogen. The speed of freezing is estimated to be $\sim 10^7$ K/s (16), enabling the rapid cooling of the aqueous medium into a glass-like state. This rapid cooling traps the macromolecular components in a near-native state, capturing the conformational distribution and spatial arrangement displayed at the instant of freezing. High-resolution cryo-EM can thus provide useful insights into understanding the nature of variation between solution and crystal structures, especially for proteins such as enzymes where protein dynamics is an important component of function.

Here, we report the solution structure of *Escherichia coli* β-galactosidase, determined at ~3.2-Å resolution by single-particle cryo-EM. This enzyme, which is the gene product of *lacZ*, catalyzes the hydrolysis of lactose and many other β-galactosides into monosaccharides (17). The functionally active form of this 465-kDa enzyme is a homotetramer with each monomer comprising 1,023 aa (18–20). We show that although the cryo-EM structure is in overall agreement with a 1.7-Å resolution crystallographic structure (21), there are significant differences in several segments of the backbone and in many of the side chains, some of which are involved in crystal contacts. Importantly, we assess the quality of the density map by comparing how well different side-chain types across the entire protein are resolved and how they are influenced by radiation damage. Our cryo-EM structure of β-galactosidase at 3.2-Å resolution and the recently reported structure of the TRPV1 channel (12) thus herald a new era in structural biology where de novo high-resolution "solution" structures of <500 kDa-sized proteins can be determined at near-atomic resolution by using cryo-EM.

Results and Discussion

Projection electron micrographs of β-galactosidase were recorded by using the K2 direct electron detector mounted at the end of a Quantum energy filter on a Titan Krios electron microscope.

Significance

Atomic resolution models for proteins and protein complexes are usually obtained using X-ray crystallography or NMR spectroscopy, and in selected instances, by cryo-electron microscopy (cryo-EM) of ordered protein assemblies. The vast majority of high-resolution structures obtained using cryo-EM have been typically restricted to large, well-ordered entities such as helical or icosahedral assemblies or two-dimensional crystals. We show here that emerging methods in single-particle cryo-EM now allow structure determination at near-atomic resolution, even for much smaller protein complexes with low symmetry, by determining the structure of the 465-kDa enzyme β-galactosidase. In addition, by quantitative comparison of density maps obtained at different electron dosages, we demonstrate preferential sensitivity of residues such as Asp and Glu to damage upon irradiation with electrons.

Author contributions: A.B., D.M., S.B., A.M., and S.S. designed research; A.B., D.M., S.B., A.M., and S.S. performed research; A.B., D.M., S.B., A.M., and S.S. analyzed data; and A.B., D.M., and S.S. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

Data deposition: Density map derived by cryo-electron microscopy and the fitted atomic model have been deposited in the EMDataBank (EMDB), www.emdatabank.org (EMDB ID code EMD-5995), and the Protein Data Bank (PDB), www.pdb.org (PDB ID code 3J7H), respectively.

[1]A.B. and D.M. contributed equally to this work.

[2]To whom correspondence should be addressed. Email: ss1@nih.gov.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1402809111/-/DCSupplemental.

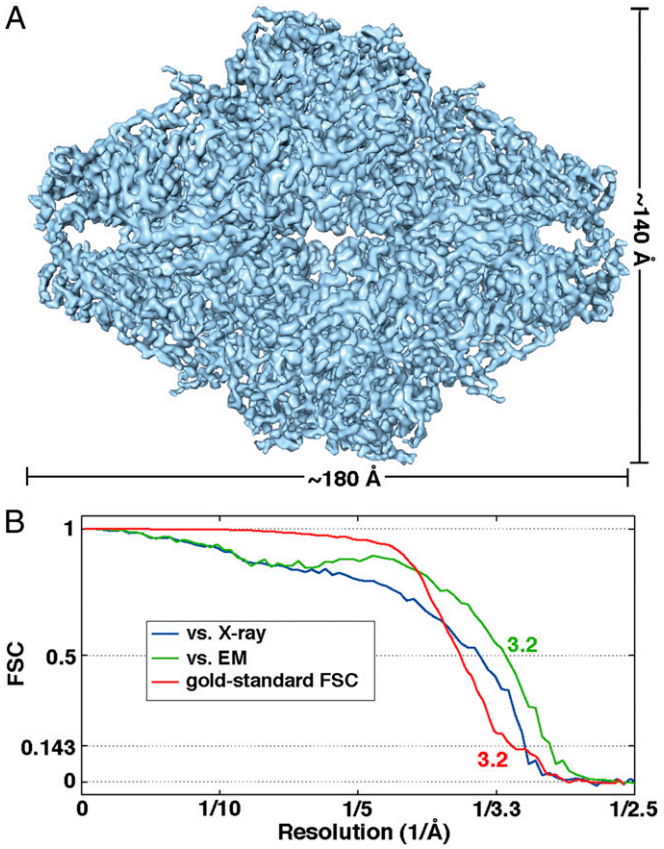

Fig. 1. Cryo-EM density map of *E. coli* β-galactosidase at 3.2-Å resolution. (*A*) Surface representation of the density map derived by cryo-EM. The dimensions of the β-galactosidase tetramer are ~180 Å x 140 Å x 87 Å. (*B*) Three different FSC plots to estimate resolution of the map shown in *A*. The curve in red shows the gold-standard FSC (24), with a value of 0.143 at 3.2-Å resolution, calculated between two independently refined halves of the dataset; the curve in blue shows the FSC (value of 0.5 at 3.6-Å resolution) calculated between the map in *A* and the map derived from the structure derived by X-ray crystallography PDB ID code 1DP0 (21); and the curve in green shows the FSC (value of 0.5 at 3.2-Å resolution) calculated between the map in *A* and the map derived from the cryo-EM–derived atomic model.

Images were collected over a time period of 15 s, and the integrated exposure obtained after alignment of individual frames show well-dispersed particles and a variety of distinct views (Figs. S1*A* and S2). Processing of the individual frames recorded over the course of the exposure allows correction for beam-induced motion, which frequently leads to significant improvements in image quality, as in the example presented in Fig. S3 (see also Figs. S4 and S5). The information in the aligned images extends to resolutions close to ~3 Å, as illustrated by the Thon rings visible in the Fourier transform of images and in the 1D power spectrum profile (Figs. S1 and S2). Automated particle selection followed by reference-free 2D classification (Fig. S6) was used to obtain class averages suitable for de novo determination of an initial map at ~20-Å resolution by using EMAN2 (22). This map was used as a starting model for refinement of the structure by using FREALIGN (23). The final map, obtained from 11,726 particles, following 16 cycles of refinement (the first 8 by using binned, and the last 8 by using unbinned images) is shown in Fig. 1*A*, and has a resolution of ~3.2 Å, as estimated by the gold standard (24) Fourier shell correlation (FSC) (Fig. 1*B*). Further, the resolution at which an FSC value of 0.5 is obtained between the experimental cryo-EM map and the map computed from the corresponding atomic model is also ~3.2 Å (Fig. 1*B*), providing independent confirmation of map resolution. The FSC curve is different when the cryo-EM density map is compared with the map derived form the crystal structure (3.6 Å vs. 3.2 Å; Fig. 1*B*), providing a measure of the differences between the atomic models derived by cryo-EM and X-ray crystallography.

There are numerous structures for β-galactosidase determined by X-ray crystallography, the majority of which use constructs where the N-terminal 9 or 12 residues are replaced with a different sequence to facilitate crystallization (Table S1). This region of β-galactosidase is thought to be more flexible than the rest of the N-terminal segment and almost all of the best-resolved crystal structures [e.g., PDB ID code 1DP0 (21) at 1.7-Å resolution] do not contain the native N-terminal sequence. Only two of the 57 available structures, solved at lower resolutions (2.5 and 2.8 Å) (21, 25) include 10 of the 12 N-terminal residues. Inspection of the cryo-EM map shows clearly resolved density for the N-terminal residues (Fig. 2*A*, boxed region) and enabled us to build these residues into the density map solely from the amino acid sequence and refine both the backbone and side-chain conformation in real space (Fig. 2*B*). The rest of the map was at similar or better quality, allowing us to obtain a complete

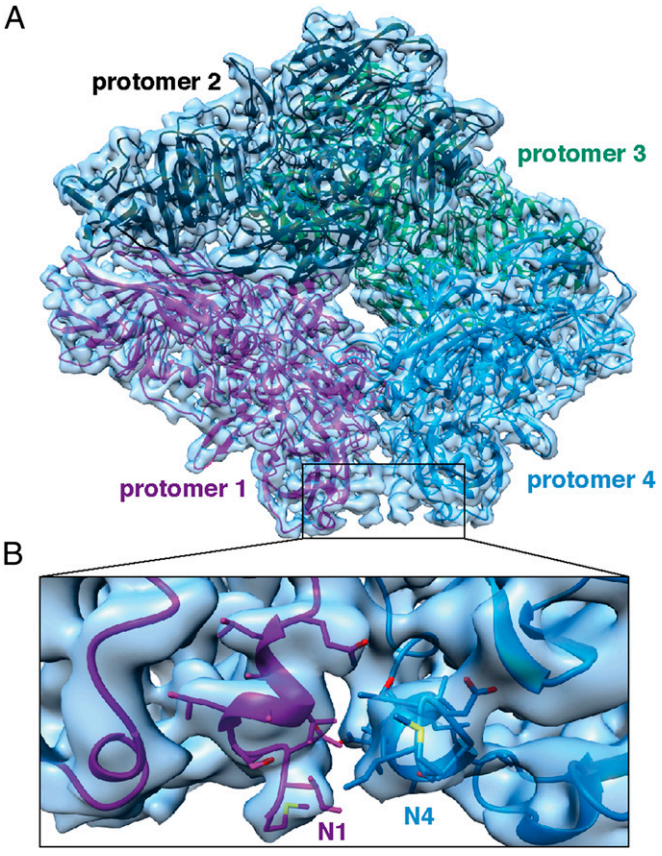

Fig. 2. Atomic model for β-galactosidase tetramer derived by cryo-EM. (*A*) Cryo-EM map shown in surface representation with fitted atomic model of β-galactosidase. The four labeled protomers are colored magenta, black, green, and blue, respectively. The density corresponding to the 12 N-terminal residues of protomers 1 and 4 are shown in the boxed region to highlight the absence of these residues in 96% of the 57 available X-ray structures. (*B*) Zoomed in region of the dimer interface highlighted in *A* showing a superposition of the cryo-EM–derived atomic model derived de novo from the experimental density map (N1 and N4 shown in magenta and blue, respectively); the cryo-EM derived atomic model closely matches that derived in the X-ray models that report coordinates for this region (only 2 of the 57 reported structures).

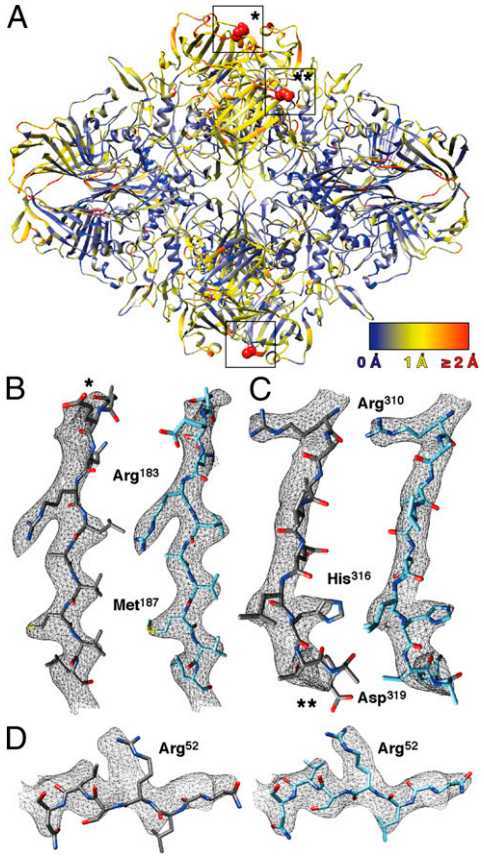

Fig. 3. Comparison of atomic models of *E. coli* β-galactosidase derived by X-ray crystallography at 1.7-Å resolution and by cryo-EM at 3.2-Å resolution. (*A*) X-ray structure of *E. coli* β-galactosidase (PDB ID code 1DP0; ref. 21) after alignment to the cryo-EM structure shown in ribbon representation colored according to rmsd values [ranging from blue (low) to red (high)]. Each chain was aligned separately in Chimera, using the default algorithm (Needleman–Wunsch in BLOSUM-62 Matrix). Highlighted areas shown as red spheres correspond to regions marked with asterisks (* and **) with rmsd values for Cα over 2 Å that are involved in crystal contacts (Fig. S9). (*B–D*) Local deviations between the atomic model derived from X-ray crystallography (*Left* in gray sticks) and the atomic model derived from cryo-EM (*Right* in light blue sticks) superimposed on the cryo-EM density map (mesh representation). Zones of protein-protein contacts in the crystal lattice including residues 179–189 (marked * in *A*) are shown in *B*, and the region including residues 310–320 (marked ** in *A*) is shown in *C*. Example of local deviations near residues 50–55, which are not involved in crystal contacts but located at the periphery of the enzyme is shown in *D*. In all examples, displacement of side chains and the backbone regions is visible.

atomic structural model of full-length β-galactosidase (Fig. 2 and Fig. S7).

Detailed comparison of the atomic model derived by cryo-EM with several X-ray structures shows small, but significant differences (Fig. 3). For example, the rmsd of all Cα atoms relative to the 1.7-Å resolution X-ray structure (PDB ID code 1DP0) is only ∼0.66 Å (Fig. S8), but it is worth noting that individual Cα atoms differ by as much as 4.6 Å. As expected, the best agreement is in the more central regions, and the greatest differences are in the periphery of the tetramer (Fig. 3*A*). Interestingly, some of the largest deviations occur precisely at zones of crystal contacts (see also Figs. S9 and S10). Comparison of the crystal structure with the structure of the protein in solution in the absence of lattice contacts shows a measurable shift of the Cα trace and an alteration of the local side-chain conformations at these zones (Fig. 3 *B* and *C*), although significant deviations are also observed elsewhere in the protein (Fig. 3*D*). These differences show that contact with neighboring molecules in the crystal results both in stabilization and local perturbation of the protein structure. Overall, this comparison confirms the potential of cryo-EM for de novo protein structure determination, especially in instances in which deletions, mutations, and crystal contacts can result in differences in conformation between the solution structure and the crystallized form.

Inspection of the cryo-EM density map reveals that, as expected from the estimated resolution, side-chain densities in β-strand and α-helical regions are well resolved (Fig. 4 *A* and *B*) with densities comparable to those observed for the same regions in an X-ray crystallographic structure of β-galactosidase (Fig. S11). The structure of the active site, including the central His and Glu residues and the catalytic Mg^{2+} ion, can also be clearly discerned (Fig. 4*C*). The coordinating water molecules and Na^{+} ions could not be identified at our present resolution. Nevertheless, the quality of the map indicates that single-particle cryo-EM methods can now furnish sufficiently high resolution to provide mechanistic information on the structure of the active site under conditions where the enzyme is dispersed in solution and is not in the environment of a 3D crystal. Clearly resolved densities are observed not merely for bulky side chains such as tryptophan and tyrosine, but also for extended side chains such as arginine and lysine and smaller side chains such as Leu, Ile, and Val (Figs. 3–5). Interestingly, densities for positively charged and neutral residues are clearly resolved in virtually all instances (see selected Arg, Lys, His, and Gln residues shown in Fig. 5). However, densities for negatively charged residues, Glu and Asp, are typically weaker (Fig. 5*B* and Fig. S12). The residues chosen in Fig. 5*B* are from different regions of the protein and include those that are buried (e.g., Glu-979 and Asp-172) as well as those exposed to solvent (e.g., Glu-198 and Asp-130), indicating that these trends are independent of residue location. One way to quantitatively explore the trends in density values for different residue types in the map is to analyze the variation in map values as shown in Fig. 5*C*. These measurements add more depth to the observations already presented in the isosurface

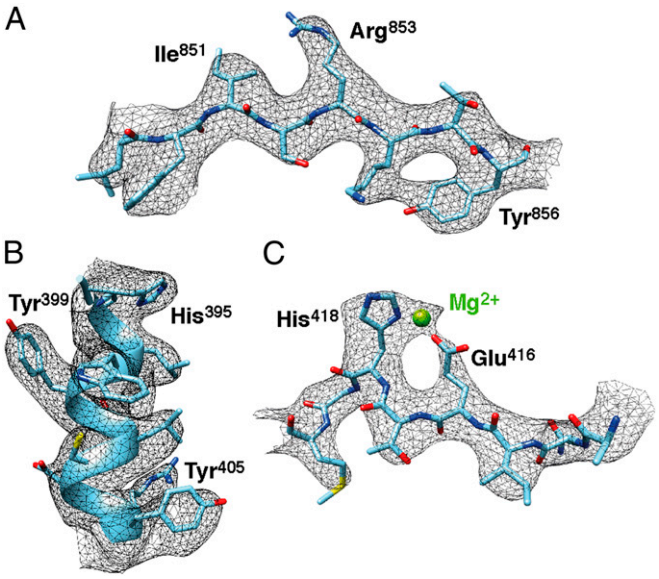

Fig. 4. Quality of cryo-EM–derived density map. Selected regions showing the fit of the derived atomic model to the cryo-EM density map (black mesh). Residues 849–856 in a β-strand (*A*), residues 395–405 in an α-helix (*B*), and residues 413–420 in the active site along with density for a Mg^{2+} ion partly coordinated by His-418 and Glu-416 (*C*).

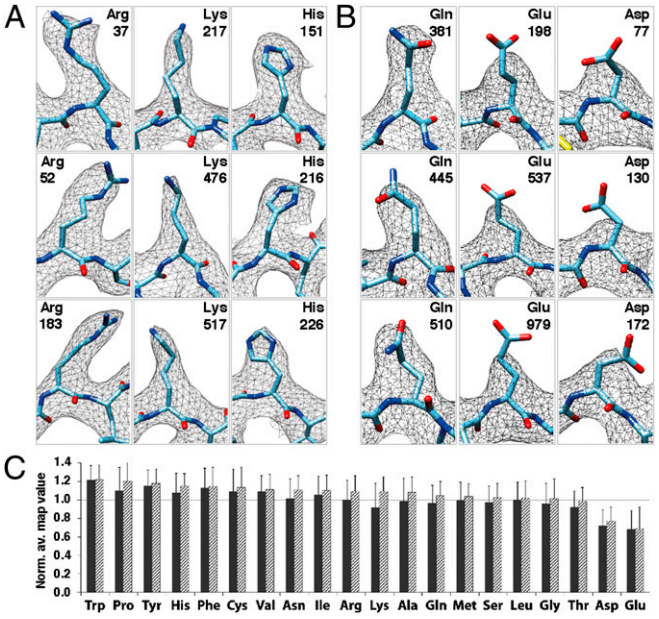

Fig. 5. Variation in observed side-chain densities between different types of residues. (*A*) Densities observed for a set of Arg, Lys, and His residues (shown in stick representation). (*B*) Comparison of densities observed for a set of Gln and Glu residues, as well as Asp residues to indicate preferential loss of density for the negatively charged side chains in comparison with the similarly sized, but neutral side chains. (*C*) Plot of normalized averaged map values (as detailed in Fig. S12), shown for each residue type when only buried residues are considered (hatched bars) or when all residues including buried and exposed varieties are considered (filled bars). Buried residues are defined as those where <30% of their maximum possible surface area is exposed to solvent.

representations of the density map and show that Glu and Asp residues display, on average, ∼30% and ∼29% less density than the similarly sized neutral Gln and Asn residues, respectively. Further, we find that the pattern of preference for radiation damage is independent of the solvent exposure of the residues (Fig. 5*C*).

It has been proposed that a possible origin of the weaker densities for Glu and Asp residues is that the differences between X-ray and electron scattering factors result in a lower signal at medium resolutions from negatively charged, deprotonated carboxylate moieties (26). An alternative, more likely reason is that Glu and Asp side chains are more sensitive to damage that occurs with electron irradiation (27–30). Selective sensitivity of negatively charged residues such as Glu and Asp, and disulfide bonds (of which there are none in the *E. coli* β-galactosidase), has also been observed for structures determined by X-ray crystallography (31–33). To evaluate this possibility, we compared the density maps reconstructed from different fractions of the total exposure (10, 20, or 30 $e^-/Å^2$), taking advantage of the dose-fractionation capabilities of current direct detector technology (Fig. 6). This comparison reveals that there is a systematic loss of the side-chain density for Glu and Asp residues with increasing dose, and that this loss is not observed for positively charged or neutral residues in the same degree (selected examples are shown in Fig. 6). The preferential sensitivity of Asp and Glu residues to electron irradiation is also presented quantitatively in Fig. S12, which shows that increasing the dose from 10 $e^-/Å^2$ to 45 $e^-/Å^2$ results in an ∼twofold increase in preferential damage of Asp and Glu residues compared with Asn and Gln residues, respectively. The trends we observe for preferential radiation damage of Asp and Glu residues can also be demonstrated by similar analysis of map value variations in a recently reported cryo-EM density map for Frh (10), an unrelated protein, at a resolution of 3.36 Å (Fig. S12*A*).

In principle, there is no reason why the success in achieving near-atomic resolution structures for the proteasome at 3.3 Å (11), mammalian TPRV1 channel with regions at 3.4 Å (12), and now β-galactosidase at 3.2-Å resolution, cannot be extended to other proteins of similar, or even smaller size. One likely prerequisite for this approach to be applied successfully is that the protein sample is structurally homogenous even in the absence of stabilization provided by a crystal lattice. NMR spectroscopic studies of proteins suggest that most proteins dispersed in aqueous solution will not be uniformly ordered in all regions of the polypeptide, and are likely to be more ordered at their core and less ordered at the periphery, where there may be greater contact with the solvent. This type of variability does not necessarily preclude structure determination by cryo-EM, but simply results in a gradient in resolution from the center to the periphery of the molecule, as seen in the case of the density map of the TRPV1 channel, where the resolution declines from 3.4 Å at the center to lower values on the periphery, preventing the visualization of side-chain densities (12). Nevertheless, the presence of this type of gradient in resolution may be informative in its own and could identify regions of the protein that are more flexible than others. Additionally, the use of Fab fragments, interacting proteins, or small molecule ligands to stabilize conformations could help in some instances to stabilize the conformation to enable cryo-EM analyses (34, 35). Further, as methods for image processing and 3D classification continue to evolve (36–38), single-particle cryo-EM analyses can be expected to routinely yield not just a single averaged structure, but an ensemble of conformations present in solution (39), thus complementing higher resolution structural studies of specific, stabilized conformations by X-ray crystallography.

The methods used to prepare cryo-EM specimens and obtain raw electron micrographs are also likely to have a significant impact on the quality of the final map(s) obtained. Dispersion of protein complexes in a layer of ice as thin as possible, to

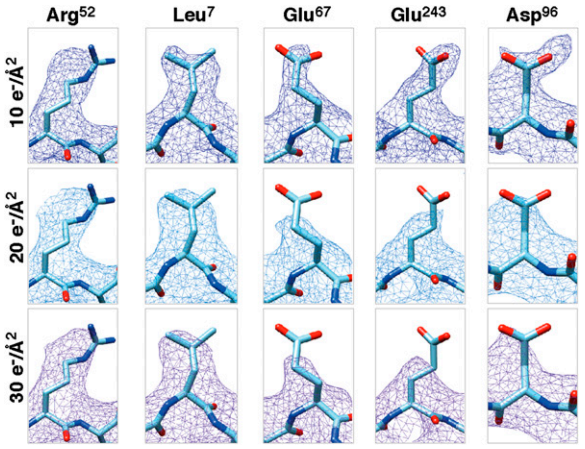

Fig. 6. Differential dose-dependent radiation damage. Side-chain densities for selected Glu and Asp residues are shown to illustrate the effects of radiation damage by comparison of maps obtained with progressively higher total electron doses. Density maps were derived by using the first 8, 16, or 24 frames (*Top*, *Middle*, and *Bottom*), corresponding to 10, 20, or 30 $e^-/Å^2$ total dose, respectively. In comparison with Arg and Leu side chains (illustrated using Arg-52 and Leu-7 as examples), Glu and Asp side chains (illustrated by using Glu-67, Glu-243, and Asp-96) are much more sensitive to radiation damage showing increased loss of density (26%, 22%, and 10% for Glu-67, Glu-243, and Asp-96, respectively) with an increase in dose from 10 to 30 $e^-/Å^2$ (see also Fig. S12 for a quantitative analysis of residue-specific radiation damage at different dose levels).

minimize background contribution from the vitrified buffer, and the use of electron optical conditions that allow recording of images at the highest possible resolution, are prerequisites for obtaining high-resolution maps. In addition, other factors unique to the protein may be especially relevant. Some proteins tend to accumulate at the air-water interface, with potential for partial denaturation. In turn, this nonrandom distribution can lead to aberrant projection images, preferential orientations relative to the vitreous ice layer, and, ultimately, to lower resolution. Some protein complexes that are relatively stable in bulk solution may also be more prone to dissociation when it is subject to interfacial forces during the blotting process that precedes plunge freezing.

Beam-induced specimen movement has long been recognized as an important factor in degradation of image quality (29). The development of direct electron detectors that allow dose fractionation now provides a way to begin to resolve this problem (40, 41). By collecting dose fractionated images, the total exposure can be divided into a large number of images, each at a very small dose, enabling correction for drift resulting from beam-induced movement and also from thermal drifts of the specimen stage. The emerging consensus from a number of recent studies (10, 11, 15, 40), which is confirmed in our present analysis, is that, on average, the largest movements tend to occur early in the exposure (Fig. S3); these early frames can be either discarded or down-weighted during image processing. Similarly, later frames collected during the exposure are expected to display greater radiation damage effects and can also be discarded to preserve high-resolution features, as shown in our analysis. Our finding that the use of just the first 10 e^-/A^2 produces density maps with the least radiation damage is consistent with estimates of the optimal dose to obtain density maps at 3-Å resolution by using cryo-EM (42). In addition to reducing the effects of beam-induced motion and mitigating the effects of radiation damage, we show that dose fractionation can also be used to improve the accuracy of contrast transfer function (CTF) estimation, which is an important resolution-limiting factor in single-particle cryo-EM (ref. 43; Fig. S5). Finally, the images we used in our analysis were recorded by using a direct electron detector at the end of an energy filter, demonstrating that the use of the filter does not limit structure determination to resolutions of ∼3 Å, and may even be beneficial because of the increased contrast of energy-filtered images.

The advances in cryo-EM that allow determination of structures of small protein complexes and membrane proteins at near-atomic resolution mark a critical shift in structural biology. It is also encouraging that we were able to determine the structure at ∼3.2-Å resolution from as few as ∼12,000 projection molecular images, bringing it closer than ever before to the theoretical limit identified by Henderson (44) for the number of "perfect" images needed to obtain atomic resolution from cryo-EM. Structure determination via X-ray diffraction requires that proteins be coaxed into conditions that enable formation of well-ordered 3D crystals; however, in some cases, the process of crystallization can significantly alter the native solution conformation of the protein. NMR spectroscopy provides an alternative route to determine solution structures, but in practice, it is not feasible for molecules larger than ∼50 kDa. With cryo-EM, structures of protein complexes as large as ∼90 MDa and as small as ∼300 kDa have now been determined at near-atomic resolution. There is every reason to hope that the structures of much smaller proteins will also be determined by cryo-EM methods. Because specimens for cryo-EM are plunge frozen rapidly from the aqueous phase, the cryo-EM density reflects the averaged, most probable protein conformation in solution and, in combination with image classification, can provide structures of a range of conformations that are sampled by the protein. Even when a number of crystal structures are already available, the determination of an atomic resolution model that captures the conformation and active site geometry in solution can be important for detailed quantitative analysis of the reaction mechanism. The prospect that the determination of protein structures to atomic resolution will no longer be limited by size or by the need for crystallization represents a significant and exciting horizon in structural biology. The stage is now set for the application of these methods to analyze structures of a variety of biologically and medically relevant multiprotein complexes and membrane protein assemblies, which have historically represented the most challenging frontier in structural biology.

Materials and Methods

Sample Preparation. β-galactosidase (Enzyme Commission 3.2.1.23; catalog no. G5635; Sigma-Aldrich) was subjected to gel filtration on a Superdex-200 size-exclusion chromatography column connected to an ÄKTA FPLC apparatus (GE Healthcare Bio-Sciences) with an elution buffer comprised of 25 mM Tris at pH 8, 50 mM NaCl, 2 mM $MgCl_2$, and 0.5 mM tris(2-carboxyethyl) phosphine. Protein suspensions at 2.3 mg/mL were deposited on 200 mesh Quantifoil R2/2 grids (Quantifoil Micro Tools) and plunge frozen by using a Leica EM GP instrument (Leica Microsystems).

Data Acquisition. The grids were imaged by using a Titan Krios transmission electron microscope (FEI Company) aligned for parallel illumination and operated at 300 kV, with the specimen maintained at liquid nitrogen temperatures. Images were recorded on a K2 Summit camera (Gatan) operated in superresolution counting mode with a physical pixel size of 1.275 Å and a superresolution pixel size of 0.6375 Å. The detector was placed at the end of a GIF Quantum energy filter (Gatan), operated in zero-energy-loss mode with a slit width of 20 eV. The dose rate used was ∼5 e^-/pixel·s (equivalent to ∼3 $e^-/Å^2$·s at the specimen plane) to ensure operation in the linear range of the detector (11). The total exposure time was 15.2 s, and intermediate frames were recorded every 0.4 s giving an accumulated dose of ∼45 e^-/A^2 and a total of 38 frames per image.

Image Processing. From a dataset of 834 micrographs, with defocus values ranging from −1.0 to −2.5 μm, a subset of 509 micrographs displaying the highest resolution Thon ring profiles, least astigmatism, and clearly visible, well-separated particles were then selected for further analysis.

To compensate for drift and beam-induced motion during the electron exposure, individual frames of each movie were aligned by cross-correlation using the cumulative average of previously aligned frames as a reference to align the remaining frames (for more details, see *SI Materials and Methods*). The increased signal-to-noise ratio of the cumulative average of frames, which was used as a reference to align the raw frames, results in significant improvements in the accuracy of motion correction as evidenced by better defined Thon rings in the drift-corrected images (Fig. S3). In terms of the cross-correlation between power spectra extracted from the corrected averages and the theoretical CTF profiles, our results compare favorably with those produced by the technique presented in ref. 11, which uses redundant measurements between raw frames to improve the consistency of motion correction (Fig. S4).

Accuracy of CTF determination is one of the main factors that impact resolution in single-particle cryo-EM specially in the near-atomic resolution regime. Taking advantage of the capability of direct electron detectors to produce dose fractionated movies, we show that in addition to providing opportunities for motion correction and mitigation of radiation damage effects, movies also provide a way to improve the accuracy of CTF determination and, therefore, the resolution of single-particle reconstructions. The way we accomplish this improved accuracy is by estimating the parameters of the contrast transfer function of each micrograph by using radially averaged power spectra obtained by periodogram averaging with tiles extracted from all frames of each movie. Compared with the standard approach of using tiles extracted from the average of frames (after drift correction), our method does not require frames to be prealigned and results in better defined 1D CTF profiles, allowing more accurate determination of defocus (Fig. S5 and for more details, see *SI Materials and Methods*). The improvement in the resolvability of zero crossings obtained by using this approach allowed us to set the upper resolution limit for CTF estimation at 3 Å while still providing reliable CTF fits.

Model Refinement and Analysis. Particles (24,750) were picked automatically from the best 509 micrographs by detecting the local maxima of

correlation of each image with a Gaussian disk of 100 Å in radius and extracted by using a binning factor of 2 and a box size of 384 × 384 pixels. Particles were then subjected to reference-free 2D classification using EMAN2 (22). From 250 classes, 160 were used for de novo initial model determination by using e2initialmodel.py imposing D2 symmetry. Class averages showing typical β-galactosidase projections and class averages with corresponding model reprojections are shown in Fig. S6. A subset of 23,452 particles (corresponding to the 160 classes used for the initial model building) was then used for 3D refinement by using a gold-standard approach. Two half-stacks (each containing 11,726 particles) were independently subjected to eight rounds of iterative refinement in FREALIGN (23) by using a high-resolution frequency limit of 8 Å and using the best 50% of particles from each stack according to phase residual values (equivalent to 5,863 particles) to calculate the reconstructions at each iteration. At this point, particles were reextracted from the original unbinned micrographs by using a box size of 768 × 768 pixels and further refined in FREALIGN starting from the most recent set of alignments obtained with the binned data. The final map was obtained by averaging the two half reconstructions in real space and corrected by a B-factor of −85 Å^2 for the purpose of visualization.

Side-chain densities were visualized clearly throughout the map, allowing us to build and refine the entire atomic model in the environment of the program COOT (45). For details, see *SI Materials and Methods*.

To analyze the variations in side-chain densities, values at atom positions of each residue were calculated in the environment of UCSF Chimera (46) and averaged for each residue type. To distinguish between exposed and buried residues, the surface area of each residue was calculated by using SURFV (47, 48). The maximum surface area for each residue type was calculated by using a GLY-XXX-GLY motif, analyzed for each residue (XXX) one at a time for each of the 20 aa.

To study the effects of radiation damage on the final map, a subset of particles that was used in the reconstruction was reextracted from the averages of the first 8, 16, and 24 frames, and the reconstruction using each set was obtained by using identical alignment parameters (i.e., Euler angles and displacements for each particle were unchanged). Figures were produced by using UCSF Chimera (46) or PyMOL (The PyMOL Molecular Graphics System, Version 1.3., Schrödinger, LLC). The final map and refined atomic model have been deposited with the Electron Microscopy Data Bank (EMDB ID code EMD-5995) and Protein Data Bank (PDB ID code 3J7H), respectively.

ACKNOWLEDGMENTS. We thank Drs. Lesley A. Earl and Mario J. Borgnia for helpful comments on the manuscript, Dr. Erin Tran for assistance with specimen preparation, and Robert Mueller for technical assistance with electron microscopy. The complete set of electron micrographs used to obtain the density map presented here is available through the Electron Microscopy Pilot Image ARchive (EMPIAR). This study used the high-performance computational capabilities of the Biowulf Linux cluster at the National Institutes of Health (NIH). This research was supported by funds from the Center for Cancer Research, National Cancer Institute, NIH.

1. Chen JZ, et al. (2009) Molecular interactions in rotavirus assembly and uncoating seen by high-resolution cryo-EM. *Proc Natl Acad Sci USA* 106(26):10644–10648.
2. Liu X, et al. (2010) Structural changes in a marine podovirus associated with release of its genome into Prochlorococcus. *Nat Struct Mol Biol* 17(7):830–836.
3. Mills DJ, Vitt S, Strauss M, Shima S, Vonck J (2013) De novo modeling of the F420-reducing [NiFe]-hydrogenase from a methanogenic archaeon by cryo-electron microscopy. *eLife* 2:e00218.
4. Settembre EC, Chen JZ, Dormitzer PR, Grigorieff N, Harrison SC (2011) Atomic model of an infectious rotavirus particle. *EMBO J* 30(2):408–416.
5. Subramaniam S, Henderson R (2000) Molecular mechanism of vectorial proton translocation by bacteriorhodopsin. *Nature* 406(6796):653–657.
6. Walz T, et al. (1997) The three-dimensional structure of aquaporin-1. *Nature* 387(6633):624–627.
7. Yu X, Jin L, Zhou ZH (2008) 3.88 A structure of cytoplasmic polyhedrosis virus by cryo-electron microscopy. *Nature* 453(7193):415–419.
8. Zhang X, Jin L, Fang Q, Hui WH, Zhou ZH (2010) 3.3 A cryo-EM structure of a nonenveloped virus reveals a priming mechanism for cell entry. *Cell* 141(3):472–482.
9. Zhang X, et al. (2008) Near-atomic resolution using electron cryomicroscopy and single-particle reconstruction. *Proc Natl Acad Sci USA* 105(6):1867–1872.
10. Allegretti M, Mills DJ, McMullan G, Kühlbrandt W, Vonck J (2014) Atomic model of the F420-reducing [NiFe] hydrogenase by electron cryo-microscopy using a direct electron detector. *eLife* 3:e01963.
11. Li X, et al. (2013) Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM. *Nat Methods* 10(6):584–590.
12. Liao M, Cao E, Julius D, Cheng Y (2013) Structure of the TRPV1 ion channel determined by electron cryo-microscopy. *Nature* 504(7478):107–112.
13. Bai XC, Fernandez IS, McMullan G, Scheres SH (2013) Ribosome structures to near-atomic resolution from thirty thousand cryo-EM particles. *eLife* 2:e00461.
14. Fernández IS, et al. (2013) Molecular architecture of a eukaryotic translational initiation complex. *Science* 342(6160):1240585.
15. Greber BJ, et al. (2014) Architecture of the large subunit of the mammalian mitochondrial ribosome. *Nature* 505(7484):515–519.
16. Dubochet J, et al. (1988) Cryo-electron microscopy of vitrified specimens. *Q Rev Biophys* 21(2):129–228.
17. Juers DH, Matthews BW, Huber RE (2012) LacZ beta-galactosidase: Structure and function of an enzyme of historical and molecular biological importance. *Protein Sci* 21(12):1792–1807.
18. Cohn M (1957) Contributions of studies on the beta-galactosidase of Escherichia coli to our understanding of enzyme synthesis. *Bacteriol Rev* 21(3):140–168.
19. Fowler AV, Zabin I (1970) The amino acid sequence of beta galactosidase. I. Isolation and composition of tryptic peptides. *J Biol Chem* 245(19):5032–5041.
20. Kalnins A, Otto K, Rüther U, Müller-Hill B (1983) Sequence of the lacZ gene of Escherichia coli. *EMBO J* 2(4):593–597.
21. Juers DH, et al. (2000) High resolution refinement of beta-galactosidase in a new crystal form reveals multiple metal-binding sites and provides a structural basis for alpha-complementation. *Protein Sci* 9(9):1685–1699.
22. Tang G, et al. (2007) EMAN2: An extensible image processing suite for electron microscopy. *J Struct Biol* 157(1):38–46.
23. Grigorieff N (2007) FREALIGN: High-resolution refinement of single particle structures. *J Struct Biol* 157(1):117–125.
24. Scheres SH (2012) RELION: Implementation of a Bayesian approach to cryo-EM structure determination. *J Struct Biol* 180(3):519–530.
25. Jacobson RH, Zhang XJ, DuBose RF, Matthews BW (1994) Three-dimensional structure of beta-galactosidase from E. coli. *Nature* 369(6483):761–766.
26. Mitsuoka K, et al. (1999) The structure of bacteriorhodopsin at 3.0 A resolution based on electron crystallography: Implication of the charge distribution. *J Mol Biol* 286(3):861–882.
27. Glaeser RM (1971) Limitations to significant information in biological electron microscopy as a result of radiation damage. *J Ultrastruct Res* 36(3):466–482.
28. Glaeser RM, Taylor KA (1978) Radiation damage relative to transmission electron microscopy of biological specimens at low temperature: A review. *J Microsc* 112(1):127–138.
29. Henderson R (1992) Image contrast in high-resolution electron microscopy of biological macromolecules: TMV in ice. *Ultramicroscopy* 46(1-4):1–18.
30. Schmid MF, Jakana J, Matsudaira P, Chiu W (1992) Effects of radiation damage with 400-kV electrons on frozen, hydrated actin bundles. *J Struct Biol* 108(1):62–68.
31. Garman EF (2010) Radiation damage in macromolecular crystallography: What is it and why should we care? *Acta Crystallogr D Biol Crystallogr* 66(Pt 4):339–351.
32. Owen RL, Rudiño-Piñera E, Garman EF (2006) Experimental determination of the radiation dose limit for cryocooled protein crystals. *Proc Natl Acad Sci USA* 103(13):4912–4917.
33. Weik M, et al. (2000) Specific chemical and structural damage to proteins produced by synchrotron radiation. *Proc Natl Acad Sci USA* 97(2):623–628.
34. Wu S, et al. (2012) Fabs enable single particle cryoEM studies of small proteins. *Structure* 20(4):582–592.
35. Bartesaghi A, Merk A, Borgnia MJ, Milne JL, Subramaniam S (2013) Prefusion structure of trimeric HIV-1 envelope glycoprotein determined by cryo-electron microscopy. *Nat Struct Mol Biol* 20(12):1352–1357.
36. Scheres SH, et al. (2007) Disentangling conformational states of macromolecules in 3D-EM through likelihood optimization. *Nat Methods* 4(1):27–29.
37. Zhang W, Kimmel M, Spahn CM, Penczek PA (2008) Heterogeneity of large macromolecular complexes revealed by 3D cryo-EM variance analysis. *Structure* 16(12):1770–1776.
38. Hohn M, et al. (2007) SPARX, a new environment for Cryo-EM image processing. *J Struct Biol* 157(1):47–55.
39. Ciccarelli L, et al. (2013) Structure and conformational variability of the mycobacterium tuberculosis fatty acid synthase multienzyme complex. *Structure* 21(7):1251–1257.
40. Brilot AF, et al. (2012) Beam-induced motion of vitrified specimen on holey carbon film. *J Struct Biol* 177(3):630–637.
41. Campbell MG, et al. (2012) Movies of ice-embedded particles enhance resolution in electron cryo-microscopy. *Structure* 20(11):1823–1828.
42. Baker LA, Smith EA, Bueler SA, Rubinstein JL (2010) The resolution dependence of optimal exposures in liquid nitrogen temperature electron cryomicroscopy of catalase crystals. *J Struct Biol* 169(3):431–437.
43. Agard DA, Cheng Y, Glaeser RM, Subramaniam S (2014) Single-particle cryo-electron microscopy (Cryo-EM): Progress, challenges and perspectives for further improvement. *Adv Imaging Electron Phys* 185:113–137.
44. Henderson R (1995) The potential and limitations of neutrons, electrons and X-rays for atomic resolution microscopy of unstained biological molecules. *Q Rev Biophys* 28(2):171–193.
45. Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 66(Pt 4):486–501.
46. Pettersen EF, et al. (2004) UCSF Chimera—a visualization system for exploratory research and analysis. *J Comput Chem* 25(13):1605–1612.
47. Nicholls A, Sharp KA, Honig B (1991) Protein folding and association: Insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins* 11(4):281–296.
48. Sridharan S, Nicholls A, Honig B (1992) A new vertex algorithm to calculate solvent accessible surface-areas. *FASEB J* 6(1):A174.