www.sciencemag.org/cgi/content/full/science.aab1576/DC1

Supplementary Materials for

2.2 Å resolution cryo-EM structure of β-galactosidase in complex with a cell-permeant inhibitor

Alberto Bartesaghi, Alan Merk, Soojay Banerjee, Doreen Matthies, Xiongwu Wu, Jacqueline L. S. Milne, Sriram Subramaniam*

*Corresponding author. E-mail: ss1@nih.gov

Published 7 May 2015 on *Science* Express
DOI: 10.1126/science.aab1576

This PDF file includes:

Other Supplementary Material for this manuscript includes the following:
(available at www.sciencemag.org/cgi/content/full/science.aab1576/DC1)

Materials and Methods

Sample Preparation

β-galactosidase (catalog #G5635; SIGMA-ALDRICH, St. Louis, MO) was subjected to gel filtration on a Superdex-200 size-exclusion chromatography column connected to an ÄKTA FPLC apparatus (GE Healthcare Bio-Sciences, Piscataway, NJ) with an elution buffer comprised of 25 mM Tris (pH 8), 50 mM NaCl, 2 mM $MgCl_2$ and 1 mM TCEP. The ligand phenylethyl β-D-thiogalactopyranoside (PETG) was purchased from Sigma-Aldrich (catalog #P1692). PETG was dissolved in 30% methanol to a final stock solution concentration of 50 mM. In order to form the protein-ligand complex, the optimal final concentration of PETG was 5 mM and that was approximately 250-fold molar excess over the β-galactosidase concentration (2.3 mg/ml) used for the experiment. The protein-ligand mixtures were deposited on 200 mesh Quantifoil R1.2/1.3 grids (Quantifoil Micro Tools GmbH, Jena, Germany), and plunge-frozen using a Leica EM GP instrument (Leica Microsystems Inc., Buffalo Grove, IL).

Data Acquisition

The grids were imaged using a Titan Krios transmission electron microscope (FEI Company, Hillsboro, OR) 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, Inc., Pleasanton, CA) operated in super-resolution counting mode with a physical pixel size of 0.637 Å. The detector was placed at the end of a GIF Quantum energy filter (Gatan, Inc.), operated in zero-energy-loss mode with a slit width of 20 eV. To minimize the effects of coincidence loss, the dose rate used was ~ 2.4 e^-/pixel/s (equivalent to ~ 5.9 $e^-/Å^2/s$ at the specimen plane) resulting in improved image contrast and maximization of amplitudes at low resolution (11), which allowed us to use lower defocus and still be able to correctly pick and align particles. The total exposure time was 7.6 s and intermediate frames were recorded every 0.2 s giving an accumulated dose of ~ 45 $e^-/Å^2$ and a total of 38 frames per image.

Image Processing

The data set used consisted of 1487 micrographs, displaying Thon ring profiles to resolutions beyond ~ 3 Å, with minimal astigmatism, with defocus values ranging from -0.6 to -2.0 μm and clearly visible particles. To compensate for drift and beam induced motion during the electron exposure, individual frames of each movie were aligned by cross-correlation using as reference the cumulative average of previously aligned frames as described previously (*18*). Image displacements used to correct for beam induced motion were restricted to be multiples of the super-resolution pixel, eliminating the need for interpolating image values at sub-pixel positions which can lead to dampening of high-resolution information present in the raw images. CTF determination for each micrograph was done using radially averaged power spectra obtained by periodogram averaging with tiles extracted from all frames of each movie as described previously (*18*).

The maximum resolution used to estimate the defocus using TOMOCTFFIND was set to 2.5 Å, and oscillations from the radially averaged power spectra for many micrographs were observed at this frequency (Fig. S1C).

Structure Refinement and Analysis

93,686 particles were picked automatically as described previously (*18*) and extracted using a binning factor of 4 and a box size of 384x384 pixels. Class averages showing typical β-galactosidase projections are shown in Fig. S1E. The same *de novo* low-resolution initial model determined previously (*18*) was used as reference for 3D refinement. Eight rounds of refinement followed by 16 rounds of 3D classification into 3 distinct classes were done with the program FREALIGN (*25, 26*). Only one of the classes presented high-resolution features and yielded the best FSC curve. The 41,123 particles that were assigned to this class were then re-extracted from the original micrographs using a binning factor of 2 and a box size of 768x768, and subsequently subjected to 8 additional rounds of refinement using a high-resolution frequency limit of 6 Å. At this point, the contribution of each frame to the 3D reconstruction was assessed by performing single-frame reconstructions and computing their FSCs against the map obtained using all the frames (Fig. S2A). Based on this curve, the subset of frames 4 to 20 (equivalent to an accumulated dose of ~ 20 $e^-/Å^2$) was identified as having the least amount of beam-induced motion and radiation damage and particles were re-extracted from the original micrographs using only this interval of frames. 8 additional rounds of refinement were done using this fraction of the exposure and the final 3D reconstruction was then re-calculated using the most recent set of Euler angles and displacements for each particle, only including information contained in frames 4-13 (interval containing the highest resolution information according to Fig. S2A), equivalent to an accumulated dose of ~ 12 e^-/A^2. The later portion of the exposure essentially improves the signal-to-noise ratio of the low-resolution terms of the structure at the expense of dampening the high-resolution components. For this reason, high-resolution features are better resolved in maps computed from the early fraction of the exposure as compared to maps computed from the full exposure. For all FSC calculations, maps were multiplied by a soft mask that followed the outer edges of the protein density (Fig. S2B, inset). The final map was corrected by a B-factor of -75 $Å^2$ for purposes of map visualization.

Atomic Model Refinement

Structure of the apo β-galactosidase tetramer (PDB 3J7H) was initially refined using the program COOT (*27*). Water molecules and ions were subsequently added to the refined atomic model followed by an additional real space refinement using the program PHENIX (*28*). At this point, the conformation of the inhibitor PETG was adjusted using the PyMOL Molecular Graphics System, added to the refined coordinates and subjected to additional cycles of COOT/PHENIX refinement. Figures were produced using UCSF Chimera (*29*), COOT (*27*), and Cinema4D (MAXON).

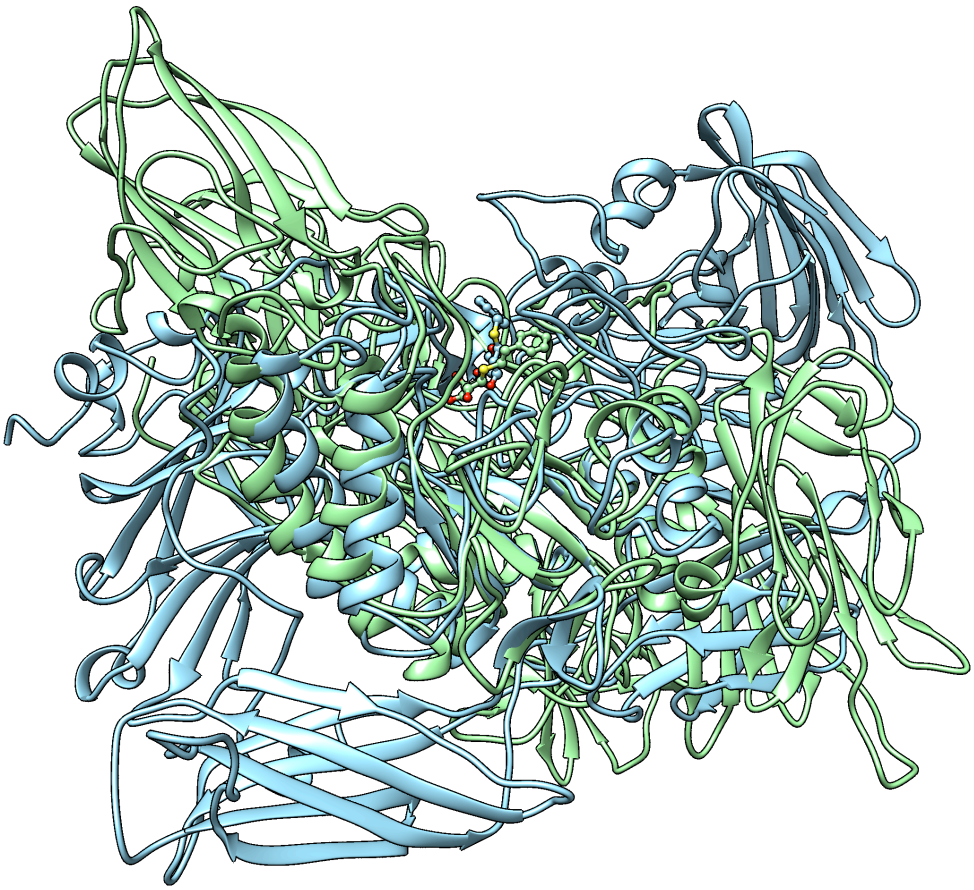


Fig. S1.

Superposition of the structures of the PETG-bound *E. coli* β-galactosidase protomer (cyan) and PETG-bound monomeric *T. reesei* β-galactosidase (green) determined by cryo-EM and X-ray crystallography, respectively.

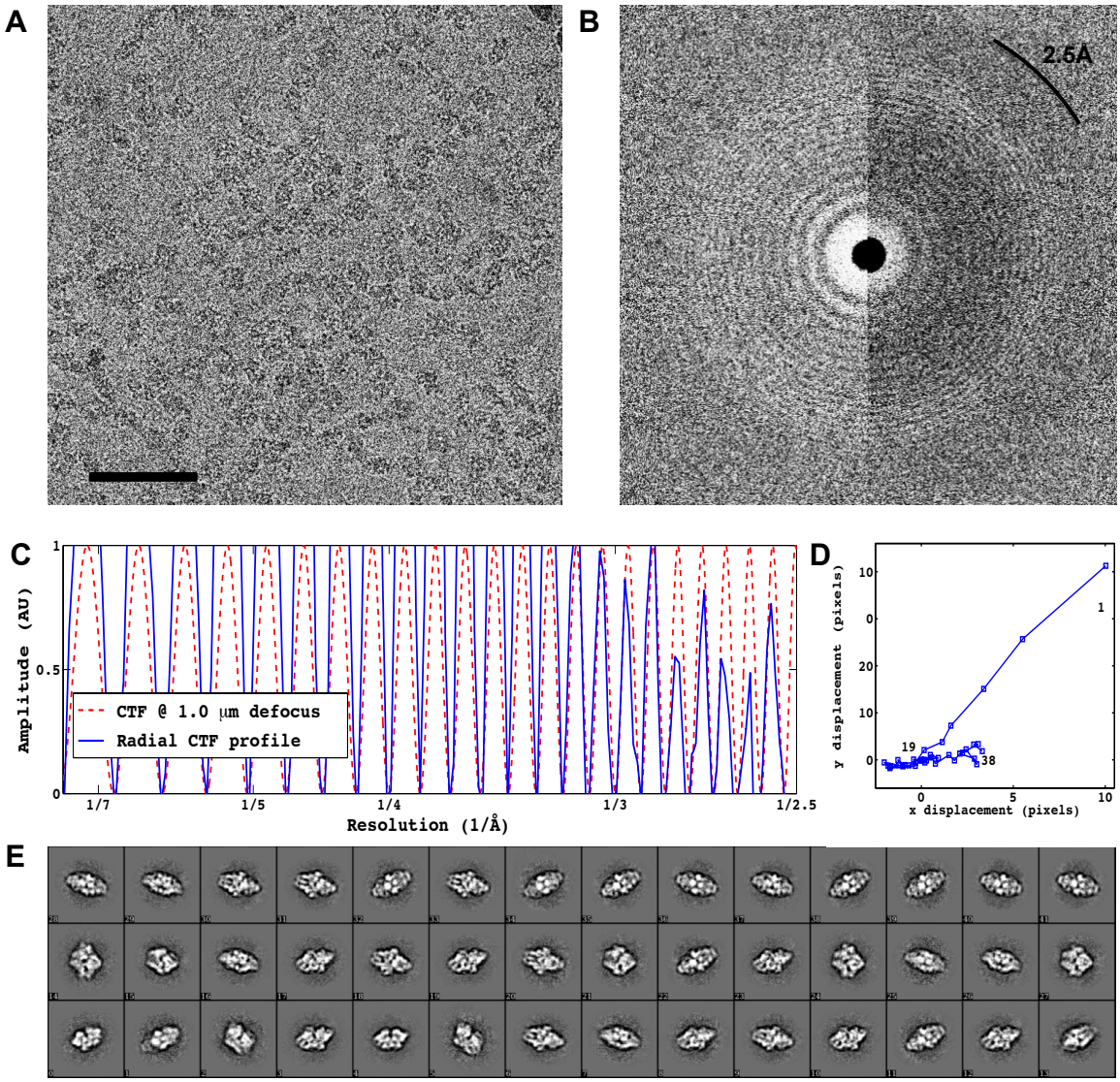


Fig. S2

Cryo-EM data collection from vitrified specimens of the PETG-bound *E. coli* β-galactosidase complex. **(A)** Average of 38 aligned movie frames acquired over a 7.6 s exposure window (45 $e^-/Å^2$ total accumulated dose). Images were recorded at 300 kV and -1 µm defocus in super-resolution mode using a physical pixel size of 0.637 Å. Scale bar is 50 nm. **(B)** FFTs of image in (A) showing the extent of Thon rings that are visible to resolutions beyond ~ 3 Å, obtained using periodogram averaging from the average of aligned frames (left panel) and from the raw individual frames (right panel). **(C)** 1D power spectrum profile obtained by radially averaging the Fourier Transform obtained from the individual frames (right panel in B) with peaks visible at 2.5 Å resolution. **(D)** Trajectory of movement for each of the 38 frames during the 7.6 s exposure used to correct the drift in image shown in (A). Displacements are measured in multiples of the super-resolution pixel size. (**E**) 2D reference-free class averages showing characteristic β-galactosidase projection views (shown only for visualization purposes, 2D classification was not used during refinement).

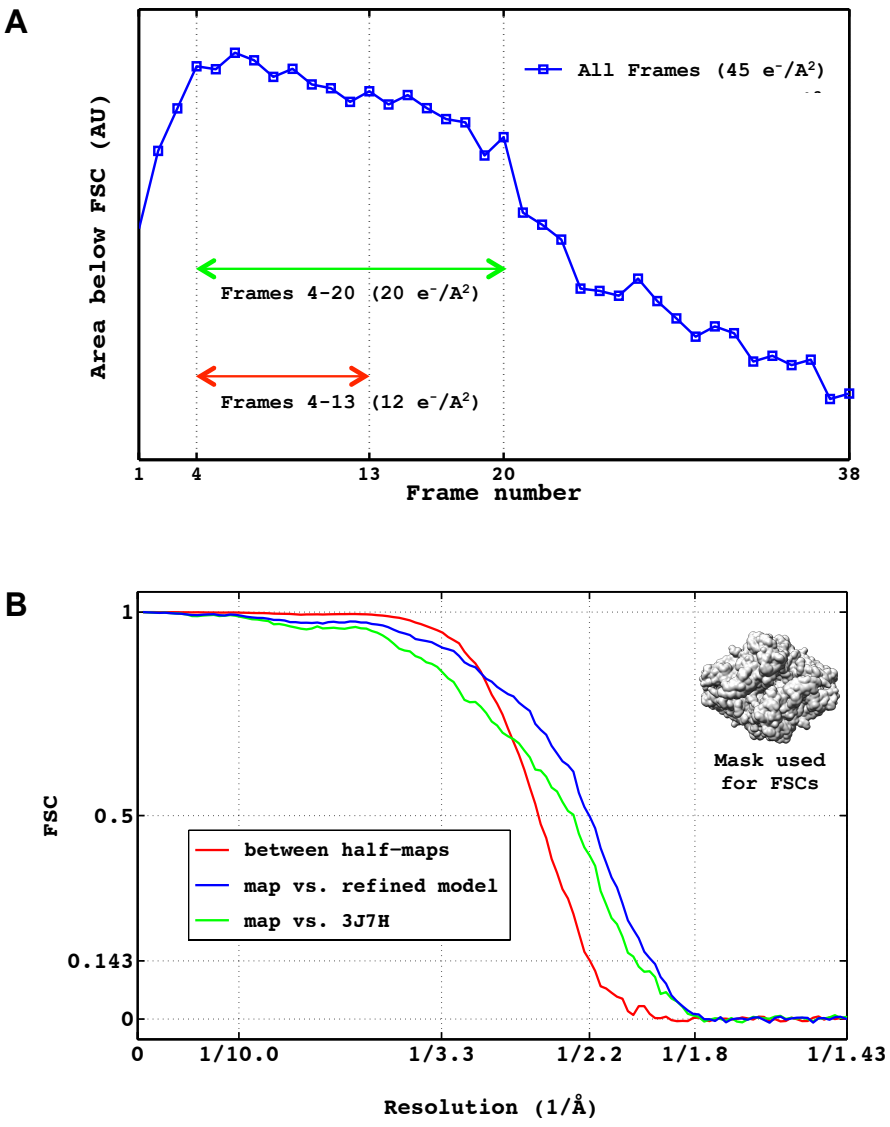


Fig. S3

Assessment of resolution and contribution of individual frames to the cryo-EM density map of β-galactosidase-PETG complex at 2.2 Å resolution. **(A)** Contribution of individual frames to the density map measured by integrating the FSC curves between maps obtained from the individual frames and the map obtained using all the frames. From a total of 38 frames, the average of frames 4-20 (the least affected by beam-induced motion and radiation damage) was used for 3D refinement, and the average of frames 4-13 (containing the highest resolution information) was used to obtain the final reconstruction shown in Figs. 1-4. **(B)** Fourier Shell Correlation (FSC) curve between two semi-independently refined halves of the data (red curve) and FSC plot calculated between the map shown in Figs. 1-4 and the map computed from the cryo-EM-derived atomic model (blue curve) both showing a resolution value of ~ 2.2 Å at the 0.143 and 0.5 cutoffs, respectively. For comparison, the corresponding plot derived using our earlier structure of apo β-galactosidase at 3.2 Å resolution (PDB 3J7H) is also shown (green curve). All FSC calculations were done using the soft shape mask illustrated in the inset.

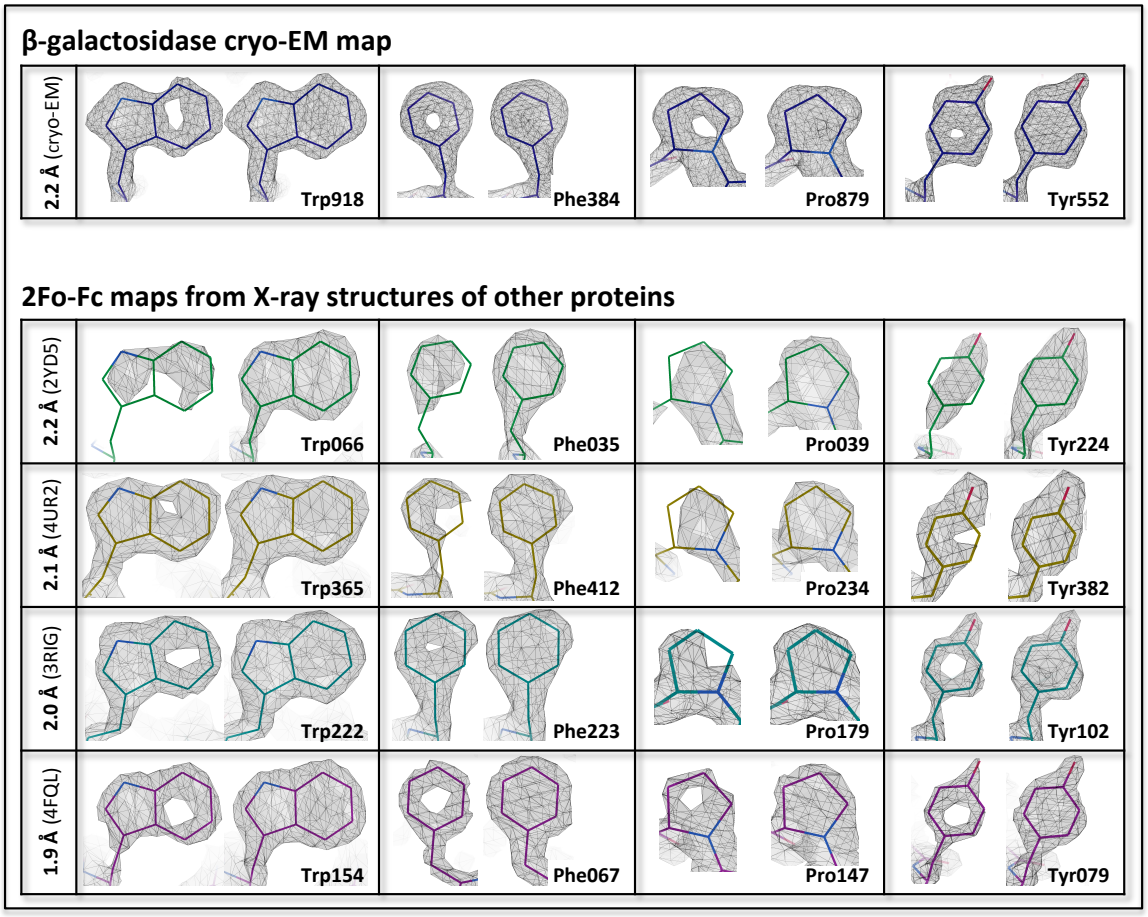
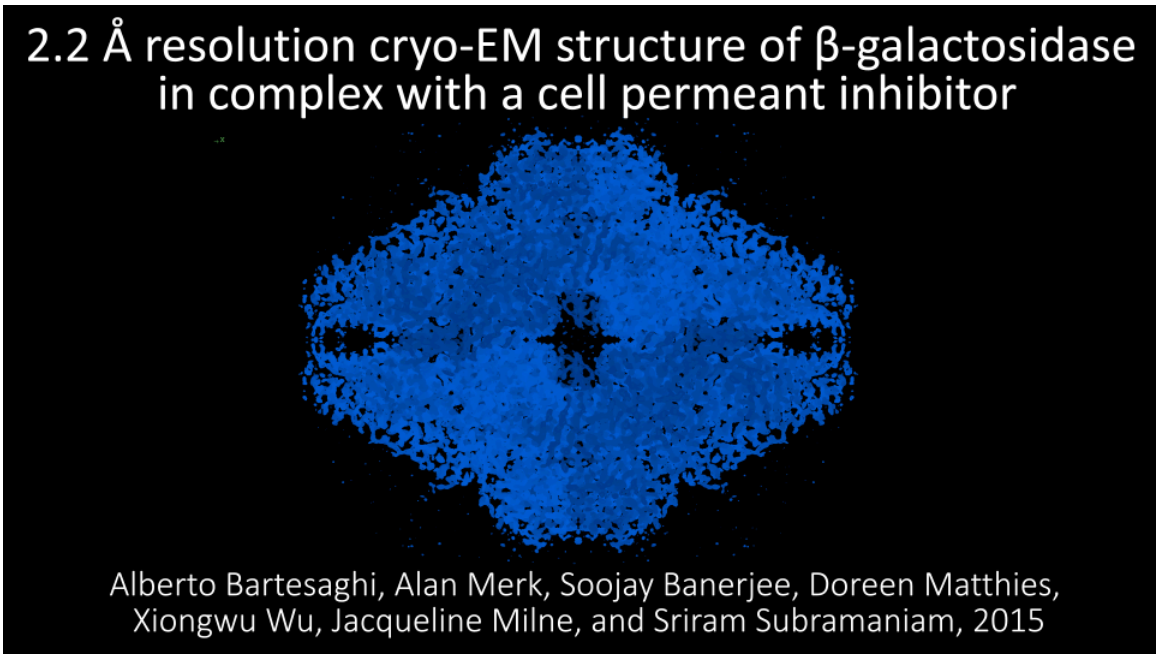


Fig. S4

Comparison of densities visualized for four residue types in the experimentally determined cryo-EM map of β-galactosidase at an overall resolution of 2.2 Å, with 2Fo-Fc maps from previous X-ray structures at resolutions ranging from 1.9 Å to 2.2 Å. Density for residues Trp, Phe, Tyr and Pro are contoured at two different levels (columns 1 and 2). Raw cryo-EM map (first row), PDB 2YD5 at 2.2 Å resolution (*30*), PDB 4UR2 at 2.1 Å resolution (*31*), PDB 3RIG at 2.0 Å resolution (*32*), and PDB 4FQL at 1.9 Å resolution (*33*). The appearance of densities in our map is comparable to features observed in X-ray maps at resolutions ranging from 1.9 to 2.2 Å. The holes in Trp, Phe and Tyr that are visible in our cryo-EM map are visible in the X-ray structures at 2.1 Å or better, while the holes in Pro visible in our map are visible in the X-ray structures at 2.0 Å or better.



Movie S1

View through selected regions of the density map of β-galactosidase where densities for amino acid side chains and coordinated water molecules are clearly visible.

References and Notes

1. N. Grigorieff, S. C. Harrison, Near-atomic resolution reconstructions of icosahedral viruses from electron cryo-microscopy. *Curr. Opin. Struct. Biol.* **21**, 265–273 (2011). Medline doi:10.1016/j.sbi.2011.01.008

2. F. Guo, Z. Liu, P. A. Fang, Q. Zhang, E. T. Wright, W. Wu, C. Zhang, F. Vago, Y. Ren, J. Jakana, W. Chiu, P. Serwer, W. Jiang, Capsid expansion mechanism of bacteriophage T7 revealed by multistate atomic models derived from cryo-EM reconstructions. *Proc. Natl. Acad. Sci. U.S.A.* **111**, E4606–E4614 (2014). Medline doi:10.1073/pnas.1407020111

3. H. Liu, L. Jin, S. B. Koh, I. Atanasov, S. Schein, L. Wu, Z. H. Zhou, Atomic structure of human adenovirus by cryo-EM reveals interactions among protein networks. *Science* **329**, 1038–1043 (2010). Medline doi:10.1126/science.1187433

4. X. Liu, Q. Zhang, K. Murata, M. L. Baker, M. B. Sullivan, C. Fu, M. T. Dougherty, M. F. Schmid, M. S. Osburne, S. W. Chisholm, W. Chiu, Structural changes in a marine podovirus associated with release of its genome into *Prochlorococcus*. *Nat. Struct. Mol. Biol.* **17**, 830–836 (2010). Medline doi:10.1038/nsmb.1823

5. E. C. Settembre, J. Z. Chen, P. R. Dormitzer, N. Grigorieff, S. C. Harrison, Atomic model of an infectious rotavirus particle. *EMBO J.* **30**, 408–416 (2011). Medline doi:10.1038/emboj.2010.322

6. Z. Wang, C. F. Hryc, B. Bammes, P. V. Afonine, J. Jakana, D. H. Chen, X. Liu, M. L. Baker, C. Kao, S. J. Ludtke, M. F. Schmid, P. D. Adams, W. Chiu, An atomic model of brome mosaic virus using direct electron detection and real-space optimization. *Nat. Commun.* **5**, 4808 (2014). Medline doi:10.1038/ncomms5808

7. M. Wolf, R. L. Garcea, N. Grigorieff, S. C. Harrison, Subunit interactions in bovine papillomavirus. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 6298–6303 (2010). Medline doi:10.1073/pnas.0914604107

8. X. Yu, P. Ge, J. Jiang, I. Atanasov, Z. H. Zhou, Atomic model of CPV reveals the mechanism used by this single-shelled virus to economically carry out functions conserved in multishelled reoviruses. *Structure* **19**, 652–661 (2011). Medline doi:10.1016/j.str.2011.03.003

9. X. Zhang, P. Ge, X. Yu, J. M. Brannan, G. Bi, Q. Zhang, S. Schein, Z. H. Zhou, Cryo-EM structure of the mature dengue virus at 3.5-Å resolution. *Nat. Struct. Mol. Biol.* **20**, 105–110 (2013). Medline doi:10.1038/nsmb.2463

10. X. Zhang, L. Jin, Q. Fang, W. H. Hui, Z. H. Zhou, 3.3 Å cryo-EM structure of a nonenveloped virus reveals a priming mechanism for cell entry. *Cell* **141**, 472–482 (2010). Medline doi:10.1016/j.cell.2010.03.041

11. X. C. 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). Medline doi:10.7554/eLife.00461

12. A. Amunts, A. Brown, X. C. Bai, J. L. Llácer, T. Hussain, P. Emsley, F. Long, G. Murshudov, S. H. Scheres, V. Ramakrishnan, Structure of the yeast mitochondrial large

ribosomal subunit. *Science* **343**, 1485–1489 (2014). Medline doi:10.1126/science.1249410

13. B. J. Greber, D. Boehringer, M. Leibundgut, P. Bieri, A. Leitner, N. Schmitz, R. Aebersold, N. Ban, The complete structure of the large subunit of the mammalian mitochondrial ribosome. *Nature* **515**, 283–286 (2014). Medline

14. P. Lu, X. C. Bai, D. Ma, T. Xie, C. Yan, L. Sun, G. Yang, Y. Zhao, R. Zhou, S. H. Scheres, Y. Shi, Three-dimensional structure of human γ-secretase. *Nature* **512**, 166–170 (2014). Medline doi:10.1038/nature13567

15. N. Fischer, P. Neumann, A. L. Konevega, L. V. Bock, R. Ficner, M. V. Rodnina, H. Stark, Structure of the *E. coli* ribosome-EF-Tu complex at <3 Å resolution by Cs-corrected cryo-EM. *Nature* **520**, 567–570 (2015). Medline doi:10.1038/nature14275

16. X. Li, P. Mooney, S. Zheng, C. R. Booth, M. B. Braunfeld, S. Gubbens, D. A. Agard, Y. Cheng, Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM. *Nat. Methods* **10**, 584–590 (2013). Medline doi:10.1038/nmeth.2472

17. M. G. Campbell, D. Veesler, A. Cheng, C. S. Potter, B. Carragher, 2.8 Å resolution reconstruction of the *Thermoplasma acidophilum* 20S proteasome using cryo-electron microscopy. *eLife* **4**, e06380 (2015). Medline doi:10.7554/eLife.06380

18. A. Bartesaghi, D. Matthies, S. Banerjee, A. Merk, S. Subramaniam, Structure of β-galactosidase at 3.2-Å resolution obtained by cryo-electron microscopy. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 11709–11714 (2014). Medline doi:10.1073/pnas.1402809111

19. J. Jiang, B. L. Pentelute, R. J. Collier, Z. H. Zhou, Atomic structure of anthrax protective antigen pore elucidates toxin translocation. *Nature* 10.1038/nature14247 (2015). Medline doi:10.1038/nature14247

20. M. Liao, E. Cao, D. Julius, Y. Cheng, Structure of the TRPV1 ion channel determined by electron cryo-microscopy. *Nature* **504**, 107–112 (2013). Medline doi:10.1038/nature12822

21. T. L. Blundell, H. Jhoti, C. Abell, High-throughput crystallography for lead discovery in drug design. *Nat. Rev. Drug Discov.* **1**, 45–54 (2002). Medline doi:10.1038/nrd706

22. D. Agard, Y. Cheng, R. M. Glaeser, S. Subramaniam, in *Advances in Imaging and Electron Physics*, P. W. Hawkes, Ed. (Elsevier, Amsterdam, 2014), chap. 2, pp. 113–137.

23. R. Henderson, The potential and limitations of neutrons, electrons and x-rays for atomic resolution microscopy of unstained biological molecules. *Q. Rev. Biophys.* **28**, 171–193 (1995). Medline doi:10.1017/S003358350000305X

24. M. Maksimainen, N. Hakulinen, J. M. Kallio, T. Timoharju, O. Turunen, J. Rouvinen, Crystal structures of *Trichoderma reesei* β-galactosidase reveal conformational changes in the active site. *J. Struct. Biol.* **174**, 156–163 (2011). Medline doi:10.1016/j.jsb.2010.11.024

25. X. Li, S. Q. Zheng, K. Egami, D. A. Agard, Y. Cheng, Influence of electron dose rate on electron counting images recorded with the K2 camera. *J. Struct. Biol.* **184**, 251–260 (2013).

26. N. Grigorieff, FREALIGN: High-resolution refinement of single particle structures. *J. Struct. Biol.* **157**, 117–125 (2007). Medline doi:10.1016/j.jsb.2006.05.004

27. D. Lyumkis, A. F. Brilot, D. L. Theobald, N. Grigorieff, Likelihood-based classification of cryo-EM images using FREALIGN. *J. Struct. Biol.* **183**, 377–388 (2013). Medline doi:10.1016/j.jsb.2013.07.005

28. P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, Features and development of *Coot*. *Acta Crystallogr. D* **66**, 486–501 (2010). Medline

29. P. H. Zwart, P. V. Afonine, R. W. Grosse-Kunstleve, L. W. Hung, T. R. Ioerger, A. J. McCoy, E. McKee, N. W. Moriarty, R. J. Read, J. C. Sacchettini, N. K. Sauter, L. C. Storoni, T. C. Terwilliger, P. D. Adams, Automated structure solution with the PHENIX suite. *Methods Mol. Biol.* **426**, 419–435 (2008). Medline doi:10.1007/978-1-60327-058-8_28

30. E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, T. E. Ferrin, UCSF Chimera—A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004). Medline doi:10.1002/jcc.20084

31. C. H. Coles, Y. Shen, A. P. Tenney, C. Siebold, G. C. Sutton, W. Lu, J. T. Gallagher, E. Y. Jones, J. G. Flanagan, A. R. Aricescu, Proteoglycan-specific molecular switch for RPTPσ clustering and neuronal extension. *Science* **332**, 484–488 (2011). Medline doi:10.1126/science.1200840

32. M. Bommer, C. Kunze, J. Fesseler, T. Schubert, G. Diekert, H. Dobbek, Structural basis for organohalide respiration. *Science* **346**, 455–458 (2014). Medline doi:10.1126/science.1258118

33. J. Du, Y. Zhou, X. Su, J. J. Yu, S. Khan, H. Jiang, J. Kim, J. Woo, J. H. Kim, B. H. Choi, B. He, W. Chen, S. Zhang, R. A. Cerione, J. Auwerx, Q. Hao, H. Lin, Sirt5 is a NAD-dependent protein lysine demalonylase and desuccinylase. *Science* **334**, 806–809 (2011). Medline doi:10.1126/science.1207861

34. C. Dreyfus, N. S. Laursen, T. Kwaks, D. Zuijdgeest, R. Khayat, D. C. Ekiert, J. H. Lee, Z. Metlagel, M. V. Bujny, M. Jongeneelen, R. van der Vlugt, M. Lamrani, H. J. Korse, E. Geelen, Ö. Sahin, M. Sieuwerts, J. P. Brakenhoff, R. Vogels, O. T. Li, L. L. Poon, M. Peiris, W. Koudstaal, A. B. Ward, I. A. Wilson, J. Goudsmit, R. H. Friesen, Highly conserved protective epitopes on influenza B viruses. *Science* **337**, 1343–1348 (2012). Medline