

BIROn - Birkbeck Institutional Research Online

Patwardhan, A. and Ashton, A. and Brandt, R. and Butcher, S. and Carzaniga, R. and Chiu, W. and Collinson, L. and Doux, P. and Duke, E. and Ellisman, M.H. and Franken, E. and Grünewald, K. and Heriche, J.-K. and Koster, A. and Kühlbrandt, W. and Lagerstedt, I. and Larabell, C. and Lawson, C.L. and Saibil, Helen R. and Sanz-García, E. and Subramaniam, S. and Verkade, P. and Swedlow, J.R. and Kleywegt, G.J. (2014) A 3D cellular context for the macromolecular world. *Nature Structural & Molecular Biology* 21 (10), pp. 841-845. ISSN 1545-9993.

Downloaded from: <https://eprints.bbk.ac.uk/id/eprint/10672/>

Usage Guidelines:

Please refer to usage guidelines at <https://eprints.bbk.ac.uk/policies.html>
contact lib-eprints@bbk.ac.uk.

or alternatively

A 3D cellular context for the macromolecular world

Ardan Patwardhan¹, Alun Ashton², Robert Brandt³, Sarah Butcher⁴, Raffaella Carzaniga⁵, Wah Chiu⁶, Lucy Collinson⁵, Pascal Doux³, Elizabeth Duke², Mark H Ellisman⁷, Erik Franken⁸, Kay Grünewald⁹, Jean-Karim Heriche¹⁰, Abraham Koster¹¹, Werner Kühlbrandt¹², Ingvar Lagerstedt¹, Carolyn Larabell¹³, Catherine L Lawson¹⁴, Helen R Saibil¹⁵, Eduardo Sanz-García¹, Sriram Subramaniam¹⁶, Paul Verkade¹⁷, Jason R Swedlow¹⁸ & Gerard J Kleywegt¹

We report the outcomes of the discussion initiated at the workshop entitled A 3D Cellular Context for the Macromolecular World and propose how data from emerging three-dimensional (3D) cellular imaging techniques—such as electron tomography, 3D scanning electron microscopy and soft X-ray tomography—should be archived, curated, validated and disseminated, to enable their interpretation and reuse by the biomedical community.

Structural biology is currently transitioning from a largely molecular perspective to a much wider range of length scales. This is made possible by the rapid development of 3D cellular imaging methods, such as electron tomography (ET), 3D scanning electron microscopy (3D-SEM) and soft X-ray tomography (SXT). Combining these new methods with established molecular structure determination techniques paves the way for understanding of biology in 3D, on scales from molecules to cells. However, for large-scale integration, interpretation, analysis and reuse of structural information at a range of scales, coordinated global efforts are required to archive, annotate, validate and disseminate data produced by methods for which no archives exist today.

To discuss the needs and opportunities in this rapidly developing area, a workshop entitled A 3D Cellular Context for the Macromolecular World was held in Cambridge in December 2012. The proposals and recommendations that emerged from this meeting and subsequent discussions are presented here, with two main aims: (i) to inform the cellular and molecular biology communities of ongoing developments in 3D cellular imaging and of the potential of synergistic use of structural information across a wide range of scales and (ii) to inform the structural biology

community (including funding agencies and journals) of the critical importance of addressing the archiving needs so that data produced by emerging experimental techniques will be captured, adequately curated and integrated with other data resources.

Emerging 3D cellular imaging techniques

Of the 3D imaging techniques providing information about the cellular context of biomacromolecules and complexes, ET is the most mature one. Its main advantage over single-particle EM is its ability to provide 3D reconstructions of individual objects (as opposed to ensemble averages), thus enabling studies of biologically relevant pleomorphic systems, i.e., systems varying in shape and size, such as the coxsackievirus¹. On the other hand, ET data typically have lower resolution and more noise than EM data. The interpretation of tomograms relies on prior knowledge of the morphology, distribution and localization of cellular features, compartments and components, and it can be facilitated by correlative imaging experiments, for instance using fluorescence microscopy and antibody labeling. The interpretation is usually conveyed through segmentation, i.e., decomposition of a reconstruction into regions of biological interest.

3D-SEM methods (such as focused-ion-beam² and serial block-face³ SEM) and SXT⁴ enable visualization of the ultrastructure of whole cells and tissue samples. In 3D-SEM, the

surface of the specimen is scanned, and thin layers are progressively abraded with a focused ion beam or an ultramicrotome; this yields a stack of two-dimensional (2D) images from which a 3D representation of the original specimen can be reconstructed. Resolutions better than 1 nm in the specimen plane and ~5 nm along the imaging axis can be achieved. This technique has been successfully used to study the distribution of HIV virus in the cell⁵ and to elucidate a 3D adenovirus-polymer network formed in the nucleus⁶.

SXT uses diffractive optics to image a specimen in the X-ray 'water window' (2.34–4.4 nm), in which organic material absorbs an order of magnitude more radiation than water. By rotating the specimen, a series of projection images can be collected for a 3D reconstruction with standard ET image-processing techniques to resolutions up to ~30 nm. A key advantage over ET is that fully hydrated specimens up to 15 μm thick can be imaged with SXT, and the data are quantitative, thus making it possible to segment many organelles according to their unique linear absorption coefficients. SXT has been used to study monoallelic gene expression in mouse olfactory neurons⁷ as well as the 3D ultrastructural organization of intact *Chlamydomonas reinhardtii* algae⁸. It has also been used in combination with fluorescence microscopy to investigate autophagosomes in whole mammalian cells⁹.

A full list of affiliations appears at the end of the paper.

Light microscopy (LM) techniques enable further expansion of the horizons of cellular imaging, including investigating temporal dynamics ('4D imaging') and living specimens. Moreover, super-resolution LM techniques¹⁰ such as stimulated emission-depletion (STED) microscopy, photoactivated localization microscopy (PALM) and structured illumination microscopy (SIM) achieve resolutions in the 10- to 100-nm range, making it possible to locate individual macromolecules within the cell. In addition, LM can be combined with EM (correlative light and electron microscopy; CLEM) in two different ways: LM can identify an object of interest that is subsequently studied with EM¹¹, or LM can be used directly to collect data conveying complementary structural information¹².

Importance of archiving data

The importance of the public archiving of scientific data and results is well appreciated by the molecular structural biology community. The Protein Data Bank (PDB)¹³ is the oldest surviving electronic archive of biomolecular data, and it now stores over 100,000 atomic models determined by a variety of techniques; the Biological Magnetic Resonance Data Bank (BMRB; <http://www.bmrb.wisc.edu/>) archives data related to NMR experiments; and the Electron Microscopy Data Bank (EMDB)¹⁴ archives 3D EM reconstructions, including electron tomograms. Since 2003, the PDB has been managed by the Worldwide PDB consortium (wwPDB; <http://wwpdb.org/>)¹⁵, which includes the Protein Data Bank in Europe (PDBe; <http://pdbe.org/>)¹⁶, the Research Collaboratory for Structural Bioinformatics (RCSB PDB; <http://rcsb.org/>), the Protein Data Bank Japan (PDBj; <http://pdbj.org/>) and the BMRB. Similarly, since 2007 the EMDB has been managed by the EMDatabank (<http://emdatbank.org/>) partners PDBe, RCSB PDB and the US National Center for Macromolecular Imaging (NCMI)¹⁷. These professionally and collaboratively managed archives provide mutual integration (for example, describing which PDB entries have been fitted to a particular EMDB map) and also link to other key biological resources such as UniProt¹⁸, Gene Ontology (GO)¹⁹ and Pfam²⁰ through the Structure Integration with Function, Taxonomy and Sequences (SIFTS) resource (<http://pdbe.org/sifts/>)²¹.

The EMDB was established at the EMBL European Bioinformatics Institute (EMBL-EBI) in 2002 and has become the single global archive for EM-derived 3D reconstructions of biomacromolecules and their complexes, with over 2,400 released maps (<http://pdbe.org/emstats/>; August 2014). Most EMDB entries

are based on single-particle experiments, but the EMDB also archives results from other EM techniques such as 2D electron crystallography (~1% of EMDB holdings) and ET (~14%). In part as a result of the workshop and ensuing discussions, the PDBe is developing a pilot archive for raw 2D image data underpinning EM and ET structures in the EMDB (Electron Microscopy Pilot Image Archive (EMPIAR); <http://pdbe.org/empiar/>). This archive will be invaluable for the development of new validation and data-processing methods as well as for teaching and software testing.

In cellular structural biology, huge volumes of imaging data are routinely collected, and the metadata and results are very complex. This has driven the development of databases for imaging data, and there are now several open-source projects providing image-database and repository functionality. The Open Microscopy Environment (OME; <http://openmicroscopy.org/>) releases OMERO, an open-source image data-management platform²². OMERO is used at thousands of sites worldwide; it is the basis for several large public repositories (<http://jcb-dataviewer.rupress.org/>)²³; <http://www.cellimagelibrary.org/>)²⁴, and its tomogram-slice viewer has been integrated into the EMDB entry pages (e.g., <http://pdbe.org/emd-1906/slice/>)²⁵. Bio-Image Semantic Query User Environment (BISQUE; <http://www.bioimage.ucsb.edu/bisque/>) is an open-source data-management system, developed at the University of California, Santa Barbara, that is used by a major plant bioimaging project (<http://bovary.iplantcollaborative.org/>) and several image-processing groups in the United States. Thus, the technology to build and operate multidimensional image-data repositories exists, and the first proofs of concept have been demonstrated.

The value of data integration

Integration of EMDB and PDB data with each other and with other major biological data resources makes structural data more easily discoverable and accessible to the wider biomedical community and allows nonspecialists to benefit from the rich biological insights afforded by the archived structural data. The cellular imaging techniques described above all have the potential to contribute unique 3D perspectives on the cellular context of biomacromolecules and their complexes. However, for this potential to be fully realized, the relevant data have to be publicly archived, annotated and integrated with other data. At present, there are no dedicated archival resources for most of these imaging techniques. In addition, appreciation of the importance of mandatory deposition (and validation) of experimental data upon

publication is not yet as pervasive in the cellular-imaging community as it is in the molecular structural biology community.

It is in this context that PDBe and OME organized discussions on the needs, challenges and opportunities for the archiving, integration and dissemination of data from emerging cellular imaging techniques. The participants included experts on 3D imaging, on scales from molecules to cells, by a range of techniques including X-ray crystallography, macromolecular EM, ET, 3D-SEM, SXT and LM; experts on archiving 3D structural data, including representatives from the EMDatabank, the wwPDB, The Cell: an Image Library (<http://www.cellimagelibrary.org/>), the JCB DataViewer (<http://jcb-dataviewer.rupress.org/>) and the Cell Centered Database (<http://ccdb.ucsd.edu/>); and experts on visualization and segmentation of 3D imaging data.

The principle of public archiving of cellular imaging data is broadly supported, but there are major obstacles that have to be addressed, including the sheer complexity of representing metadata from diverse methods. There are also technical challenges involved in storing and especially in transporting large data sets. Many of the new techniques routinely produce data sets of 10 GB or more. In contrast, the largest data set currently in the EMDB is 5 GB (EMD-2489). Changing the mindset of the cellular-imaging community about the importance of public archiving of its imaging data could be a challenge. Finally, even with public archives in place, the integration of 3D cellular data with molecular structural data and the effective dissemination of structural data, information and insights to nonspecialist users are anything but trivial. The development of the tools and capabilities that can integrate these linked yet different data types should have high priority.

To illustrate the benefits of data deposition, curation and integration, we describe molecular and structural resources related to herpes simplex virus (HSV) and show how linking of different data sets could lead to new biological insights and engender hypotheses that could not be derived from the individual data sets.

HSV is endemic in the human population, with infection manifested by cold sores, genital herpes and in rare cases encephalitis. The virus can establish a lifelong latent infection in the nervous system²⁶. The structure of the HSV virion has been shown to be pleomorphic by ET²⁷. There are high-resolution X-ray crystal structures of individual HSV proteins in the PDB (e.g., 1NO7, 1DML, 2GUM and 2GJ7) and reconstructions in the EMDB of the entire virion including the envelope and tegument (layer between the envelope and capsid) and the capsid with the portal vertex (e.g., EMD-1956,

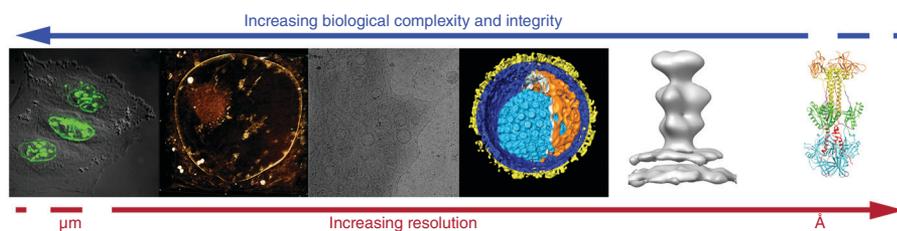


Figure 1 The HSV life cycle studied at different length scales. An integrated approach combining high-resolution structure determination methods and CLEM, SXT, 3D EM and ET makes it possible to study dynamic processes at different levels of resolution and complexity, ultimately leading to a better understanding of those processes. Figure and legend adapted from Zeev-Ben-Mordehai *et al.*²⁶ under the terms of the CC BY 3.0 license.

EMD-2379, EMD-5260 and EMD-5453). There have been numerous ET studies of different stages of the HSV life cycle, for example at the point of cell entry, and 3D snapshots of HSV during axonal transport from primary neurons during latent infection. HSV capsid assembly and egress from the nucleus have been studied with correlative SXT, ET and fluorescence LM²⁶. Given the importance of understanding the HSV life cycle for human health and the unique insights afforded by these 3D-imaging techniques, the amount of experimental 3D structural information covering all the steps of the HSV life cycle is likely to increase considerably in the next few years. In an ideal (but as yet hypothetical) future, all these data would be publicly archived, integrated with other relevant sequence, structure and function resources and made accessible through a web-based viewer to provide an integrated view of the data, including the ability to navigate between different scales and different types of information (Fig. 1). A biologist could then quickly gain an overview of the entire HSV life cycle on the basis of the experimentally determined 3D structures. A user interested in axonal transport could bring up related ET reconstructions, possibly on different kinds of neurons, offering alternative views on the stages of transport. Drilling down into the molecular world, users could retrieve high-resolution structures of the capsid, identify related proteins on the basis of sequence similarity and locate publications that provide additional information and insights.

A structural biologist studying the nuclear pore complex (NPC) would find not only 3D subtomogram-averaged structures of the entire NPC but also 3D reconstructions of the cellular environment showing the NPC distribution and its colocalization with other important assemblies and complexes, e.g., HSV. These reconstructions would be linked to related publications on HSV-NPC interaction and to higher-resolution structures showing the capsid portal through which viral DNA is injected into the cell nucleus. Furthermore, the structures of

different viruses interacting with the NPC could be compared to provide new insight, transcending what could be learned from the published literature. Annotations of the segmented objects would lead to other resources such as Pfam, UniProt and GO and enable further exploration of proteins that are related in sequence, function or cellular localization.

Looking ahead

We present here a number of key recommendations, some of which are challenges to existing and emerging communities and methods developers. We hope that these recommendations will be discussed further in the various communities and with key stakeholders (such as funding agencies and journals) to ensure that the archival needs will be addressed in a timely and globally coordinated fashion so that the data generated in studies of molecular and cellular structural biology can be properly retained, annotated, integrated, disseminated and analyzed to reach their full potential.

A pilot archive should be established for SXT and 3D-SEM data. Addressing the archiving needs for all the cellular imaging techniques in a single archive is unrealistic and impractical. It would require substantial funding to set up and maintain such an archive, and it would also require a wide range of disparate imaging communities to contribute to, support and promote this effort and to agree on standards, policies, formats, etc. However, establishing a pilot archive focusing on a few targeted techniques would allow for experience to be gained in dealing with the technical challenges and in working with the stakeholder communities to promote data deposition. Depending on the success of such a pilot archive and the future funding landscape, it might be opportune to incorporate the archive into the EMDB and gradually expand it to accommodate other imaging techniques.

Two techniques recommended as the initial focus for such a pilot archive are SXT and 3D-SEM. Each still has a relatively small community that includes many scientists with a

background in X-ray crystallography or EM who are aware of the importance of public data archiving. The EMDB data model could be fairly easily adapted to cover SXT and 3D-SEM. Nevertheless, both techniques offer sufficient challenges such as establishing community consensus on formats and deposition requirements, linking with existing molecular archives and handling 3D-SEM data sets, which are typically large (commonly 0.1–1 TB) compared to those from ET.

Correlative LM data should be archived with the companion EM or ET data. In principle, CLEM data could be archived in extant public archives dedicated to LM, with cross-references provided to the accompanying EMDB entries. However, currently, there is no single LM archive that has a stature comparable to that of the EMDB in the EM community. Moreover, users would need to deposit data and provide metadata to two different archives for correlative experiments. It is therefore recommended that the EMDataBank investigate the feasibility of archiving correlative LM data in the EMDB with the provision of metadata descriptors to allow the LM and EM data to be linked. The pilot archive for SXT and 3D-SEM proposed above should be designed to accommodate correlative LM data.

Segmentations with biological annotation are essential for integration of cellular and molecular (structure) data resources. Biologically annotated segmentations need to be archived alongside the 3D reconstructions to facilitate the integration of cellular imaging data with other biological data. Most segmentation packages allow for free-text annotation of segmented objects, but this is not sufficient for integration purposes, which require database identifiers (for example, UniProt identifiers) and terms from established ontologies (for example, GO). After an earlier workshop (Data Management Challenges in 3D Electron Microscopy)²⁸, the PDBe has been working with members of the EM community to develop a data model that supports biological annotation of segmented objects. The data model is generic and software and method agnostic and could be used to handle segmentations from SXT and 3D-SEM as well. Tools are needed to make it easy to produce, annotate and deposit segmentations, including format-conversion tools and a tool for interactively annotating segmentations. Currently, EuroBioImaging is developing a common region-of-interest (ROI) data model (<http://scijava.org/roi-model/>) that could be useful for representing segmentations.

Integrated multiscale viewers are needed to allow effective dissemination, analysis and knowledge discovery of cellular and molecular

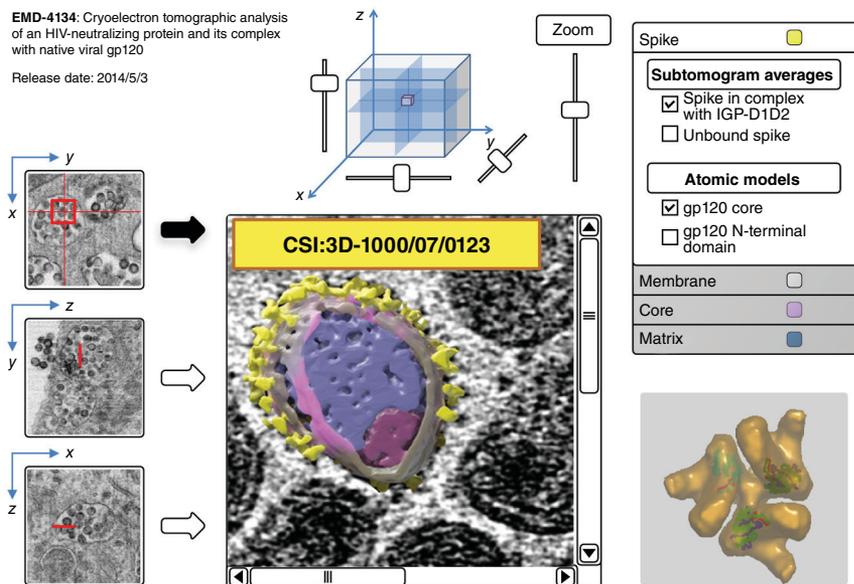


Figure 2 Mock-up of what a volume browser for 3D cellular imaging data could look like, using HIV and simian immunodeficiency virus (SIV) as an example, by integrating data from different imaging modalities. The three leftmost images are derived from a 3D-SEM reconstruction and show HIV virion reservoirs in infected macrophages (images adapted from ref. 29 under the terms of the CC BY 2.5 license). These three orthogonal cross-sections and the cube help users orient themselves in the data. The central panel shows a slice from an ET reconstruction in which the features of individual SIV virus particles can be identified³⁰. (Image adapted from ref. 30. This research was originally published in *J. Biol. Chem.* Bennett, A. *et al.* Cryoelectron tomographic analysis of an HIV-neutralizing protein and its complex with native viral gp120. *J. Biol. Chem.* 2007; 282:27754-27759. © the American Society for Biochemistry and Molecular Biology.) Here, biologically meaningful segmentations have been colored and overlaid on the tomogram, and the corresponding annotations are shown at top right. Bottom right shows a 3D rendering of data from the EMDB and PDB (image adapted from ref. 31, Nature Publishing Group). Figure and legend adapted from ref. 32 under the terms of the CC BY 2.0 UK license.

structure data. The potential synergy of cellular and molecular structural-biology methods can be realized fully only if data from different archives can be combined and integrated. This requires appropriate cross-references and common annotation between entries in the various structural and image archives as well as publicly available tools for searching and visualization of the combined holdings of these archives. Early efforts include a web-based interactive 3D viewer for EMDB entries and fitted PDB models, and an OMERO-based slice viewer for interactive visualization of 2D sections from 3D volume data (for example, a tomogram reconstruction)²⁵. To create an integrated multiscale viewer, such tools could be extended and combined into a web-based volume browser and explorer. Important extensions that would be required include an option to overlay segmentations in the presented views and to examine the annotations and links to other biological resources. **Figure 2** shows a mock-up of what the user interface and functionality for such a viewer could look like.

A different mechanism for dissemination of structural insights to a broader audience is through animations, which are an excellent means to visualize dynamics and functionally

important changes in molecular structures. Journals do not refer in any consistent way to animations, nor are animations linked to the EMDB and PDB entries on which they are based. A dedicated, sustainable public archive for animations based on experimental structures that could be referred to by journals would be very useful. To facilitate citations, animations could have accession codes and metadata linking them to the parent PDB and EMDB entries.

It is clear that the field of cellular structural biology is poised for exciting developments in the coming years that will yield new insights on the 3D structure and dynamics of organelles, cells and samples as well as the cellular context of molecular complexes and machines. To fully realize the potential of the emerging 3D cellular imaging techniques, it is imperative to address the issues of data archiving (including deposition, curation, validation, analysis, integration and dissemination) as soon as possible while the fields are still relatively young and their communities small. The experience gained by archival resources in molecular structural biology (PDB and EMDB) could help jumpstart similar activities in 3D cellular imaging and ensure that existing and new archives are

adequately linked. Besides the need for archival resources themselves, there is also a need to develop new tools so that structural results can be properly integrated and annotated with biological classification systems, ontologies and other resources and so that the data can be reused by expert and nonexpert users alike. The sources of 3D structural data that we ought to start archiving in the near future are heterogeneous, and very few people are experts in more than one or a few techniques. Thus, there will be a huge potential payoff for knowledge discovery through the integrated analysis of such data.

ACKNOWLEDGMENTS

We thank P. Haslam for help with the manuscript and logistical organization of the workshop. The workshop was supported by grants from the Wellcome Trust to the PDBe (088944) and OME (095931). Work on the EMDB at the PDBe is supported by the European Commission Framework 7 Programme (284209), the US National Institutes of Health National Institute of General Medical Sciences (R01 GM079429) and the UK Medical Research Council (MR/L007835), with further support to the PDBe by EMBL-EBI and the Wellcome Trust (088944). Work on the OME at the University of Dundee is supported by the Wellcome Trust (095931) and the Biotechnology and Biological Sciences Research Council (BB/G022585).

COMPETING FINANCIAL INTERESTS

The authors declare competing financial interests: details are available in the online version of the paper ([doi:10.1038/nsmb.2897](https://doi.org/10.1038/nsmb.2897)).



This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 3.0

Unported License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-sa/3.0/>.

- Shakeel, S. *et al.* *J. Virol.* **87**, 3943–3951 (2013).
- Heymann, J.A. *et al.* *J. Struct. Biol.* **166**, 1–7 (2009).
- Weston, A.E., Armer, H.E. & Collinson, L.M. *J. Chem. Biol.* **3**, 101–112 (2009).
- Weiss, D. *et al.* *Ultramicroscopy* **84**, 185–197 (2000).
- Narayan, K. *et al.* *J. Struct. Biol.* **185**, 278–284 (2014).
- Ou, H.D. *et al.* *Cell* **151**, 304–319 (2012).
- Clowney, E.J. *et al.* *Cell* **151**, 724–737 (2012).
- Hummel, E. *et al.* *PLoS ONE* **7**, e53293 (2012).
- Duke, E.M. *et al.* *Ultramicroscopy* **143**, 77–87 (2014).
- Schermelleh, L., Heintzmann, R. & Leonhardt, H. *J. Cell Biol.* **190**, 165–175 (2010).
- Kopeck, B.G., Shtengel, G., Grimm, J.B., Clayton, D.A. & Hess, H.F. *PLoS ONE* **8**, e77209 (2013).
- Szyzborska, A. *et al.* *Science* **341**, 655–658 (2013).
- Bernstein, F.C. *et al.* *J. Mol. Biol.* **112**, 535–542 (1977).
- Tagari, M., Newman, R., Chagoyan, M., Carazo, J.M. & Henrick, K. *Trends Biochem. Sci.* **27**, 589 (2002).
- Berman, H., Henrick, K. & Nakamura, H. *Nat. Struct. Biol.* **10**, 980 (2003).
- Velankar, S. *et al.* *Nucleic Acids Res.* **40**, D445–D452 (2012).
- Lawson, C.L. *et al.* *Nucleic Acids Res.* **39**, D456–D464 (2011).

18. UniProt Consortium. *Nucleic Acids Res.* **41**, D43–D47 (2013).
19. Dimmer, E.C. *et al.* *Nucleic Acids Res.* **40**, D565–D570 (2012).
20. Punta, M. *et al.* *Nucleic Acids Res.* **40**, D290–D301 (2012).
21. Velankar, S. *et al.* *Nucleic Acids Res.* **41**, D483–D489 (2013).
22. Allan, C. *et al.* *Nat. Methods* **9**, 245–253 (2012).
23. Hill, E. *J. Cell Biol.* **183**, 969–970 (2008).
24. Orloff, D.N., Iwasa, J.H., Martone, M.E., Ellisman, M.H. & Kane, C.M. *Nucleic Acids Res.* **41**, D1241–D1250 (2013).
25. Lagerstedt, I. *et al.* *J. Struct. Biol.* **184**, 173–181 (2013).
26. Zeev-Ben-Mordehai, T., Hagen, C. & Grunewald, K. *Curr. Opin. Virol.* **5**, 42–49 (2014).
27. Grunewald, K. *et al.* *Science* **302**, 1396–1398 (2003).
28. Patwardhan, A. *et al.* *Nat. Struct. Mol. Biol.* **19**, 1203–1207 (2012).
29. Bennett, A.E. *et al.* *PLoS Pathog.* **5**, e1000591 (2009).
30. Bennett, A. *et al.* *J. Biol. Chem.* **282**, 27754–27759 (2007).
31. Liu, J., Bartesaghi, A., Borgnia, M.J., Sapiro, G. & Subramaniam, S. *Nature* **455**, 109–113 (2008).
32. Gutmanas, A. *et al.* *Acta Crystallogr. D Biol. Crystallogr.* **69**, 710–721 (2013).

¹Protein Data Bank in Europe, European Molecular Biology Laboratory, European Bioinformatics Institute, Hinxton, UK. ²Diamond Light Source, Didcot, UK. ³FEI Visualization Sciences Group, Mérégnac, France. ⁴Institute of Biotechnology, University of Helsinki, Helsinki, Finland. ⁵Electron Microscopy Unit, Cancer Research UK London Research Institute, London, UK. ⁶National Center for Macromolecular Imaging, Verna and Marrs McLean Department of Biochemistry and Molecular Biology, Baylor College of Medicine, Houston, Texas. ⁷Center for Research in Biological Systems, National Center for Microscopy and Imaging Research (NCMIR), University of California, San Diego, San Diego, California, USA. ⁸FEI Electron Optics B.V., Eindhoven, the Netherlands. ⁹Division of Structural Biology, Wellcome Trust Centre for Human Genetics, Oxford, UK. ¹⁰Cell Biology and Biophysics Unit, European Molecular Biology Laboratory, Heidelberg, Germany. ¹¹Department of Molecular Cell Biology, Leiden University Medical Center, Leiden, the Netherlands. ¹²Department of Structural Biology, Max Planck Institute for Biophysics, Frankfurt, Germany. ¹³Department of Anatomy, University of California, San Francisco, San Francisco, California, USA. ¹⁴Research Collaboratory for Structural Bioinformatics, Rutgers University, Piscataway, New Jersey, USA. ¹⁵Institute of Structural and Molecular Biology, Department of Crystallography, Birkbeck College, London, UK. ¹⁶Center for Cancer Research, National Cancer Institute, Bethesda, Maryland, USA. ¹⁷Wolfson Bioimaging Facility, School of Biochemistry, University of Bristol, Bristol, UK. ¹⁸Centre for Gene Regulation and Expression, University of Dundee, Dundee, UK. Correspondence should be addressed to G.J.K. (gerard@ebi.ac.uk) or J.R.S. (j.swedlow@dundee.ac.uk).