Macromolecular structures without crystals

Robert M. Glaeser*
Lawrence Berkeley National Laboratory, University of California, Berkeley, CA 94720

Electron microscopy has many attractive capabilities as a tool for experimentally visualizing the structures of biological cells and macromolecules. Particles can be imaged while freely suspended in biochemically "native" buffers, and electrons can be focused to resolutions exceeding anything that a biochemist's heart might desire. To be sure, biological electron microscopy also is subject to certain physical limitations, perhaps the most problematic of which arise from the fact that short-wavelength electrons also are a form of ionizing radiation. However, among the goals for which technology rather than physical principles currently appears to be the limitation, many regard the most desirable to be the production of 3D density maps that show the positions and rotamer conformations of every amino acid by using specimens in which the proteins are in the form of well dispersed, single particles. The structures of such proteins would not only be free of crystal-packing constraints, they would also be free of the influences of the associated crystallization buffers, which often have unphysiological pH, ionic strength, ionic composition, or water activity. With the publication of the structure of the icosahedral "inner capsid particle" of a human rotavirus in this issue of PNAS, Zhang *et al.* (1) have come remarkably close to that goal. The technological frontier now moves to accomplishing the same results with macromolecular complexes that are significantly smaller than viruses and with structures for which a high degree of symmetry cannot be exploited to ease the task.

Atomic Models from Electron Micrographs of Crystals and Helices

The first protein structure to be solved by electron microscopy at a high enough resolution to build an atomic model of the polypeptide chain was that of a naturally crystalline membrane protein, bacteriorhodopsin (2). In retrospect, the chief reason for using crystals rather than single, dispersed molecules was to make it easy to average the images of ≈10,000 identical copies of the protein for each independent view of the structure. This large amount of averaging overcomes the extremely poor signal-to-noise ratio that exists in images recorded with electron exposures low enough to not destroy the sample. Structures of additional 2D crystals then followed, including tubulin (3), an important cytoskeletal protein, and aquaporin-0 (4), whose electron diffraction patterns were used to refine the structure to a resolution of 0.19 nm. However, electron microscopy of 2D crystals is a technically tedious method of structure determination because of the frequently poor quality of the data that are obtained when specimens are tilted to high angles. Electron crystallography (5) nevertheless remains an important option for determining macromolecular structures, especially when the structures within 2D crystals are uniquely relevant to the protein's biochemical function, as may often be the case for membrane proteins (6).

The next milestone in the development of this technology was to obtain high-resolution density maps of helical protein structures. This was a greater challenge because the signal-to-noise ratio for helices is considerably lower than it is for 2D crystals. The weaker signal is due in part to the fact that there are fewer copies of the repeating structural motif in a "1D" helix than in a "2D" sheet. The weaker signal also is due in part to the fact that the Fourier transform of a helix is distributed along "layer lines" rather than being concentrated into discrete diffraction spots. Even so, once the appropriate computational methodology was developed, electron microscopy was able to deliver chain-trace models of the nicotinic acetylcholine receptor (7) and of the bacterial flagellar filament (8).

The signal-to-noise ratio for helices is considerably lower than it is for 2D crystals.

Icosahedral Particles Were the Next Logical Step

Obtaining a high-resolution density map for the protein shell of an icosahedral virus was the next logical challenge. Although the basic 60-fold symmetry of an icosahedron makes this task simpler than that of working with particles that have no symmetry at all, it nevertheless requires every virus particle in the dataset to be independently aligned to every other particle, and, in addition, the orientation of each particle must be determined relative to that of every other particle. Bottcher *et al.* (9) and Conway *et al.* (10) were the first to do this at a resolution below 1 nm, sufficient to visualize separate densities for individual α-helices within a four-helix bundle of the capsid protein of the hepatitis B virus. Three-dimensional maps with a resolution better than 0.8 nm are becoming fairly routine for icosahedral structures, some of which even allow tracing the position of the polypeptide backbone within β-sheets and building a ribbon-diagram interpretation of the structure (11). At least a few of these structures are sure to soon advance to a resolution sufficient to build an atomic model into the density. To make such significant advances over the past decade has required the use of improved microscopes, the development of improved algorithms for aligning single particles (and assigning their angular orientations in space), and the fairly recent access to high-capacity computing that is provided by laboratory-scale "clusters."

Single Macromolecules Now Seem Achievable

What then remains to be done to routinely obtain atomic models of large, asymmetric macromolecules by electron microscopy? To begin with, automation of data collection (12) and data processing should make it not too onerous to include between half a million and 5 million particles in any one dataset. This order of magnitude corresponds to the number of copies of protein molecules for which data have been merged when processing images of 2D crystals, and it is the number of protein molecules that had to be included by Zhang *et al.* (1). However, the 60-fold symmetry of an icosahedron meant that 504,000 copies of the asymmetric unit were contained within images of only 8,400 virus particles, and >6.5 million copies of the capsid proteins contributed to the map when advantage was taken of the addi-

Author contributions: R.M.G. wrote the paper.
The authors declare no conflict of interest.
See companion article on page 1867.
*E-mail: rmglaeser@lbl.gov.
© 2008 by The National Academy of Sciences of the USA

tional 13-fold (nonicosahedral) symmetry that is present in the architecture of this particle. Although building a credible chain-trace proved to be possible after enforcing just the icosahedral symmetry, using the additional T = 13 symmetry of this particular structure resulted in a density map whose quality is comparable with (or perhaps even better than) that of an x-ray crystal structure of the same particle that was obtained at a resolution of 0.38 nm. More than just a brute-force scale-up of the numbers of (asymmetric) particles may be needed, however. The size of this virus particle, large in comparison with most molecular machines of interest, provided a much stronger signal—needed during alignment and assignment of particle orientations—than will be available in more general cases. Scaling up the computational aspects of aligning even half a million particles, let alone 5 million, is another technological challenge that remains to be addressed when one attempts to extend the current accomplishment to asymmetric structures.

It has not been necessary to average data from so many asymmetric units in all cases, however, nor has it always been necessary to use such large particles to achieve accurate alignment. The atomic model of flagellin was built into a map obtained by averaging data from only ≈40,000 copies of the protein (8). In addition, the positions of individual protein trimers of bacteriorhodopsin, with a combined molecular weight of only ≈75 kDa, could occasionally be identified when an essentially noise-free template was used to compute a cross-correlation map (13). These achievements are well within what is estimated to be physically possible, provided that the signal in electron microscope images is nearly as perfect as it is in the electron wave function that is transmitted through a specimen or, equivalently, as it is in the scattered wave function (14, 15). The technological challenge that is arguably the most important of all to be addressed at this point is to find ways to reduce the beam-induced movement that in most images causes the signal to be only 10% or less of what it is in the scattered wave, i.e., ways to make things go really right most of the time rather than just so rarely as to be almost anecdotal rather than scientific. If the signal level would be consistently as high as one-third rather than 1/10 of what physics would allow it to be in a perfect image, for example, the number of asymmetric particles needed in a high-resolution dataset would drop from ≈1 million to only ≈100,000, and the particle-size for which the signal would still be high enough to support accurate alignment might drop to 100 kDa or less. When such technology for controlling beam-induced movement is combined with newly developing technology for producing in-focus (Zernike-type) phase contrast with cryoelectron microscopy specimens (16), one might even realistically expect that chain-trace structures of the type achieved by Zhang *et al.* (1) could be produced for single protein molecules as small as the hemoglobin tetramer. To paraphrase the poet Robert Frost (17), electron microscopy as a tool for molecular structure analysis still has promises to keep and miles to go before it can sleep.

1. Zhang Z, Settembre E, Xu C, Dormitzer PR, Bellamy R, Harrison SC, Grigorieff N (2008) *Proc Natl Acad Sci USA* 105:1867–1872.
2. Henderson R, Baldwin JM, Ceska TA, Zemlin F, Beckmann E, Downing KH (1990) *J Mol Biol* 213:899–929.
3. Nogales E, Wolf SG, Downing KH (1998) *Nature* 391:199–203.
4. Gonen T, Cheng YF, Sitz P, Hiroaki Y, Fujiyoshi Y, Harrison SC, Walz T (2005) *Nature* 438:633–638.
5. Glaeser RM, Downing K, DeRosier D, Chiu W, Frank J (2007) *Electron Crystallography of Biological Macromolecules* (Oxford Univ Press, New York).
6. Hankamer B, Glaeser R, Stahlberg H (2007) *J Struct Biol* 160:263–264.
7. Miyazawa A, Fujiyoshi Y, Unwin N (2003) *Nature* 423:949–955.
8. Yonekura K, Maki-Yonekura S, Namba K (2003) *Nature* 424:643–650.
9. Bottcher B, Wynne SA, Crowther RA (1997) *Nature* 386:88–91.
10. Conway JF, Cheng N, Zlotnick A, Wingfield PT, Stahl SJ, Steven AC (1997) *Nature* 386:91–94.
11. Jiang W, Baker ML, Jakana J, Weigele PR, King J, Chiu W (2008) *Nature*, in press.
12. Stagg SM, Lander GC, Cheng AC, Quispe JD, Mallick SP, Avila RM, Carragher B, Potter CS (2006) *J Struct Biol* 155:470–481.
13. Walz T, Grigorieff N (1998) *J Struct Biol* 121:142–161.
14. Henderson R (1995) *Q Rev Biophys* 28:171–193.
15. Glaeser RM (1999) *J Struct Biol* 128:3–14.
16. Danev R, Nagayama K (2008) *J Struct Biol* 161:211–216.
17. Frost R (1923) in *New Hampshire: A Poem with Notes and Grace Notes* (Holt, New York), p 87.