Macromolecular electron microscopy in the era of structural genomics

Wolfgang Baumeister and Alasdair C. Steven

Macromolecular machines carry out many cellular functions. Cryo-electron microscopy (cryo-EM) is emerging as a powerful method for studying the structure, assembly and dynamics of such macromolecules, and their interactions with substrates. With resolutions still improving, ‘single-particle’ analyses are already depicting secondary structure. Moreover, cryo-EM can be combined in several ways with X-ray diffraction to enhance the resolution of cryo-EM and the applicability of crystallography. Electron tomography holds promise for visualizing machines at work inside cells.

OVER THE PAST 10 years, major strides have been made towards harnessing the instrumental resolving power of the electron microscope for biological purposes. At <3 Å, this resolving power should suffice to disclose the folds of macromolecules. One key step has been the recognition that preservation of native structure is a performance-limiting factor of paramount importance, and measures have been devised to preserve proteins through the ordeal of electron imaging. Destructive dehydration effects are averted by immobilizing the specimen in a thin layer of vitrified buffer by rapid freezing, then transferring it to a cryogenic specimen holder for observation – hence ‘cryo-electron microscopy’. Damage from electron irradiation is limited by restricting exposures to the minimum needed for statistically defined images. The second pivotal innovation has been the pervasive use of computers: to suppress noise by averaging over many individual images; to synthesize 3D density maps from 2D micrographs; to correct images for optical (phase-contrast) effects; and, to an increasing extent, for data acquisition, replacing human operation of the microscope by computer control.

To date, the fruits of this progress include the derivation of atomic models for several intrinsic membrane proteins, such as bacteriorhodopsin and light-harvesting complex, at resolutions in the 3.5–4.5 Å range, and visualization of a dozen or so other membrane proteins in three dimensions at resolutions of 5–10 Å. An atomic model has also been obtained for tubulin, a major cytoskeletal protein, and several other proteins that are not integral membrane components but nevertheless pack into well-ordered 2D lattices have also been studied at somewhat lower resolution. Work on ‘single’ (i.e. noncrystalline) particles has succeeded in depicting secondary structure at resolutions in the range of 7–10 Å, so far mainly for icosahedral viruses like the hepatitis B virus capsid. In addition, many studies at resolutions of 15–30 Å have yielded a wealth of information on domain and subdomain interactions, and intersubunit interactions in various protein complexes.

As resolutions continue to improve, the question arises whether electron microscopy will emulate or even supplant X-ray crystallography as a source of high-resolution structures. Currently, this appears unlikely (but not impossible); rather, the emerging relationship is one of complementarity. Cryo-EM offers a promising approach for molecules that are refractory to X-ray diffraction analysis, for example complexes that are too large, unstable or insoluble, or are available in insufficient amounts for crystallographic studies to be feasible. Moreover, as outlined below, cryo-EM and X-ray crystallography – and to some extent, nuclear magnetic resonance (NMR) spectroscopy – can be combined to overcome the limitations of either method alone. Indeed, this hybrid approach is emerging as a viable strategy for elucidating large complexes, including the macromolecular machines whose functional pre-eminence in a cellular context is becoming increasingly evident.

The aim of this article is to review the major lines of development of 3D cryo-EM; to illustrate its potentialities with some recent or otherwise notable examples; and to assess its prospects for a major role in the structural genomics initiative.

Electron crystallography

Ordered protein arrays, often only one molecule thick, are too insubstantial to be analysed by X-ray diffraction but make ideal specimens for a mode of EM analysis that is crystallographic in spirit. The goal is to build up the 3D Fourier transform of the repeating element by recording data from arrays tilted through various angles (Box 1). Each such view provides a central plane through the 3D transform, obtained by calculating the Fourier transform of the corresponding digital image. As such, it provides phases as well as amplitudes.

For well-ordered crystals, electron diffraction yields amplitudes that are superior to those from transformed images, and consequently are incorporated into the transform. When all available data have been integrated into a composite 3D transform, the density map is obtained by inverse Fourier transformation.

This approach, electron crystallography, was pioneered on bacteriorhodopsin, a protein that forms ordered arrays in the membrane of a halophilic bacterium under conditions of oxygen deprivation. The method is particularly well suited for intrinsic membrane proteins that are visualized in their natural environment; that is, embedded in a lipid bilayer. While the number of naturally crystalline membrane proteins is limited, a more widely applicable approach is to solubilize the molecule of interest in detergent, purify it in micelles, and reconstitute the protein with lipids, aiming to form 2D (or 3D) crystals. This method of structural analysis is equally well applicable to monolayer arrays of water-soluble proteins, such as those formed by tubulin

W. Baumeister is at the Dept of Molecular Structural Biology, Max-Planck-Institut für Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, Germany; and A.C. Steven is at the Laboratory of Structural Biology, National Institute of Arthritis, Musculoskeletal and Skin Diseases, National Institutes of Health, Bethesda, MD 20892-2717, USA. Emails: baumeist@biochem.mpg.de; steven@calvin.niams.nih.gov
in the presence of Zn$^{2+}$ ions$^4$, or proteins arrayed on planar lipid monolayers$^8$, as with annexin$^9$. Rational strategies have been devised to promote the association of proteins carrying high-affinity tags (e.g. oligo-His) with monolayers of suitably derivatized lipids, thus using membranes as crystallization templates$^{10}$. 

With large well-ordered arrays, electron crystallography can yield atomic models, although the resolution tends to be considerably lower in the dimension perpendicular to the plane. This anisotropy occurs because it is not practicable to tilt through a full 90°; thus, that part of the 3D transform – the 'missing cone' – remains undefined (see Box 1). Consequently, features organized in the plane, such as the packing of transmembrane α helices, tend to be well represented, whereas features arranged perpendicular to it, such as interhelical linkers, are less well conveyed. Advances in atomic force microscopy now make it possible to delineate surface topography in submolecular detail and compensate at least in part for this missing information$^{11}$. 

Alternatively, with tubular membranes containing helically ordered arrays of proteins, as have been obtained with the acetylcholine receptor$^{12}$, and the Ca$^{2+}$-ATPase$^{13}$, the missing cone problem is circumvented because (almost) all views of the repeating subunit are presented in a single tube.

Electron crystallographic studies of bacteriorhodopsin are now targeting distinct states of the photocycle$^{14}$. An atomic model was obtained for a light-harvesting complex, revealing the spatial distribution of its prosthetic groups, in addition to the protein fold$^3$. Recent work on aquaporins$^{15-17}$, an extensive family of water channels, has discovered a novel hairpin motif, with two copies per subunit, penetrating the membrane from either side (Fig. I). A landmark result was the structure determination of tubulin arrayed in planar sheets$^1$. Another protein of great interest that long resisted attempts to elicit structural information below -20Å resolution is connexin, the structural subunit of gap junctions. By exploiting well-ordered arrays obtained by expressing the protein in HeLa cells, Unger et al. succeeded in visualizing the configuration of four transmembrane α helices per subunit in the hexameric ring$^{18}$. 

Box 1 Combining data to calculate 3D density maps in electron crystallography, single-particle analysis and electron tomography.

The 3D Fourier transform of the density distribution of the molecule of interest serves as a melting pot for combining cryo-electron microscopy (cryo-EM) data collected in each of three strategies: electron crystallography, single-particle analysis and tomography. These strategies are appropriate for different kinds of specimens. The key relationship exploited in these syntheses is that the Fourier transform of an EM image – which is a projection – gives a 2D transform that is a central plane of the full 3D transform, that is a plane that passes through its origin. For ease of illustration, the transforms are represented in Fig. I in two dimensions, so that this simplification relates to 1D projections and central lines in 2D transforms.

In any case, the digital transforms are sampled on Cartesian lattices. The goal in each structural analysis is to fill in the values of the transform at each point of this lattice as completely and accurately as possible. Then the density map can be obtained by simply calculating an inverse Fourier transform. The resolution attained depends on the radius to which reliable data extend in the transform (reciprocal space): the limits marked in Fig. I represent the current state of the art for the three approaches. However, resolution is also reduced if parts of the transform remain empty; that is, if there are no data to be inserted. For instance, the ‘missing cone’ (actually, in Fig. I, a missing wedge) is a consequence of being unable to tilt through a full ±90° and results in anisotropic resolution (see main text). Resolution is also limited by the amount of data missing in the spaces between sampled planes; for instance, in tomography it is limited by the angular increment between tilts (Δω, in radians), according to the relationship: r > (ΔωD)/2, where r is the resolution and D is the diameter of the object.

Because micrographs are noisy, their Fourier transforms are also noisy and must be averaged to suppress the noise (not possible in tomography, where only one image is obtained at each tilt angle). A rich variety of averaging strategies have been developed to synthesize the available data. In electron crystallography, the transform is not sampled continuously, but only at discrete points on each plane – the reciprocal lattice (shown in Fig. I as bold points on each line). At these points, the signal-to-noise ratio is highly favorable compared to corresponding points in the transforms of individual particle images, on account of implicit averaging. Each term in the Fourier transform is a complex number. With large, well-ordered 2D crystals, electron diffraction provides superior amplitudes that can then be combined with phases taken from image transforms (an Argand diagram for one such Fourier component is shown in Fig. I, top right). In practice, many procedures in current use do not literally fill up the 3D transform and calculate the inverse transform: for instance, the data can be used to calculate the coefficients of a functional expansion in Fourier space, or the density map can be calculated entirely by real space operations. However, these algorithms are, in effect, carrying out the same operations, and the scenario outlined here allows a systematic comparison of the three approaches.

'Single-particle' analysis

Free-standing macromolecules offer certain practical advantages for cryo-EM. Because there is no requirement for crystallinity, virtually any particle is eligible providing that it exceeds a certain size – currently at the 250–500-kDa level, depending on structure. For particles that assume random orientations in vitrified films – which is usually the case – the ‘missing cone’ problem does not occur, so that resolution should be isotropic (Box 1). Calculation of 3D structures from noncrystalline specimens
is widely called ‘single-particle’ analysis. Commendably succinct, this term is arguably misleading because the final 3D density map is derived not from a single particle but from thousands of particles. However, it should be recalled that each of these particle images has to be classified on the basis of its individual information content.

The price to be paid for these advantages is that the viewing geometry of each particle must be determined before a 3D density map can be calculated. In general, the analysis proceeds in two stages. First, an initial density map at moderate resolution, say 30 Å, is somehow acquired. A variety of methods can be followed to obtain it, including reconstruction of data acquired by tilting the specimen, as in the ‘random conical tilt’ method or electron tomography (see below); or without tilting, as in ‘angular reconstitution’; or by matching it against a grid of reference images. The second stage involves cyclic model-based refinement. At each stage, more accurate values for the viewing angles of each particle are obtained by matching it against a grid of reprojections of the current model, each of which represents a particular view. Translational alignment is also fine-tuned. Then, a refined map is calculated and this procedure is iterated exhaustively.

The resolution attained depends on several factors, including the number of particles in the data set, the precision (and correctness) of the orientation and alignment parameters, and, of course, the quality of the original data. Even under the most favorable conditions, macromolecules yield noisy, low-contrast images, and for particle orientations to be determined successfully, they must generate sufficient signal to allow reliable discrimination among the various projections. In practice, this requirement places a lower size limit on macromolecules that can feasibly be analyzed – currently 250–500 kDa. It might be possible to extend this limit downwards by binding heavy-metal clusters as fiducial markers, or by incorporating the molecule of interest into a larger (thus more tractable) complex, or even a related complex, reconstructing a tilt series of negatively stained data, or simply computer-generating a model based on a priori information. The second stage involves cyclic model-based refinement. At each stage, more accurate values for the viewing angles of each particle are obtained by matching it against a grid of reprojections of the current model, each of which represents a particular view. Translational alignment is also fine-tuned. Then, a refined map is calculated and this procedure is iterated exhaustively.

The most general method for obtaining 3D information by EM is tomography; this is applicable not only to isolated particles but also to pleomorphic structures like mitochondria and other organelles, whole cells, or tissue sections. The basic concept is not new, but formidable technical obstacles prevent realization of its potential until recently. The key problem lies in reconciling two conflicting requirements: to obtain a detailed and undistorted reconstruction, a tilt series must be recorded that covers as wide an angular range as possible in as many increments as possible (Box 1); at the same time, the electron dose must be kept subcritical if radiation damage is not to erase important details. In principle, one could partition a fixed total dose over as many projections as are required; however, the resulting projections would be correspondingly noisier, making alignment of them for reconstruction unreliable. This problem is aggravated by the limited mechanical accuracy of specimen holders. With the advent of computer-controlled microscopes and large-area digital cameras, it became possible to implement automated procedures for the acquisition of

**Figure 1**

(a) Three-dimensional density map of human aquaporin 1 (AQP1) water channel, determined at 4 Å resolution by electron crystallography of ice-embedded 2D crystals. The view is normal to the bilayer and the membrane and shows density at 1.5 and 2 Å resolution, respectively. Scale bar, 10 Å.

(b) Stereo view of the threading of AQP1 polypeptide chain through the lipid bilayer membrane. The six tilted transmembrane α helices (in solid color) in a monomer form a barrel enclosing the water-selective pathway. This pathway is also lined by two interhelix loops (in solid white) that contain short α helices and house a pair of functionally important, highly conserved NPA tripeptide motifs. Courtesy of A. Mitra. Reproduced, with permission, from Ref. 17.

Electron tomography

The most general method for obtaining 3D information by EM is tomography; this is applicable not only to isolated particles but also to pleomorphic structures like mitochondria and other organelles, whole cells, or tissue sections. The basic concept is not new, but formidable technical obstacles prevent realization of its potential until recently. The key problem lies in reconciling two conflicting requirements: to obtain a detailed and undistorted reconstruction, a tilt series must be recorded that covers as wide an angular range as possible in as many increments as possible (Box 1); at the same time, the electron dose must be kept subcritical if radiation damage is not to erase important details. In principle, one could partition a fixed total dose over as many projections as are required; however, the resulting projections would be correspondingly noisier, making alignment of them for reconstruction unreliable. This problem is aggravated by the limited mechanical accuracy of specimen holders. With the advent of computer-controlled microscopes and large-area digital cameras, it became possible to implement automated procedures for the acquisition of
fact, with less experimental effort. In general, however, the latter approach needs an unbiased and sufficiently detailed starting model, and this is where the two approaches can be combined most effectively. Electron tomography can provide reliable starting models, which can then be refined and extended by single-particle analysis.

Hybrid crystallography: the marriage of X-ray diffraction and cryo-EM

Visualizing interactions.

Cryo-EM, while not yet yielding individual structures at resolutions as high as those obtained with X-ray crystallography, is readily applicable to macromolecular complexes. If the structures of the individual components are known to high resolution, a cryo-EM density map of the entire complex, in the intermediate resolution range of 15–30 Å, will usually allow their relative positions and orientations to be specified to within a few angstroms, thus identifying the interaction surfaces. This crucial information about interactions cannot, in general, be inferred from the structures of the individual components, even at very high resolution, and crystals of functional complexes tend to be a rather rare commodity.

The first example of such work described the docking of a monoclonal Fab onto rhinovirus, pinpointing its epitope on the capsid surface, subsequently confirmed by crystallography. Shortly thereafter, the myosin S1 ATPase was fitted into a cryo-EM map of S1-decorated actin filaments, depicting a rigor cross-bridge. A third success came from determining the placement of adenovirus hexons – trimers of 109-kDa subunits – on the virion surface. Variations on this theme have since proliferated in studies of the cytoskeleton, among other topics.

In most such analyses, the docking has been done by hand, using molecular graphics. The EM density map is represented as a transparent surface, and a stick model of the molecule to be docked is manipulated to optimize the fit; that is, to minimize excursions of the molecule beyond the surface and unoccupied lacunae inside it. However, the need for quantitative, objective fitting procedures has been recognized. In particular, it will be valuable to impose calibrated figures of merit for the fits, and to develop criteria for whether discrepancies reflect shortcomings of the cryo-EM map or the subunit model, e.g. in the crystal structure vis-à-vis the solution structure, or are indicative of conformational changes induced in forming the complex. To confirm the orientations of components in complexes, heavy-metal clusters can serve as valuable site-specific markers.

Phasing X-ray crystallographic data sets with EM-derived molecular envelopes. Diffraction-quality crystals of large complexes call for different phasing methods than are applicable to crystals of typically sized globular proteins. For instance, heavy atoms might lack tomographic data sets. Rather complex protocols now allow correction for the imperfections of the experimental setup, while minimizing the cumulative electron dose. Data sets of 100–200 projections can be recorded at doses low enough not to visibly damage ice-embedded specimens.

As applied to single particles, the method has produced a 3D structure of the archaeal thermosome, a hexadecameric chaperone, that was good enough to allow docking of its domains of known structure (Fig. 3). Tomographic density maps of numerous individual particles were calculated, screened to remove outliers, and averaged. In this particular case the same result could, no doubt, have been obtained by single-particle methods – in particular, it will be valuable to impose calibrated figures of merit for the fits, and to develop criteria for whether discrepancies reflect shortcomings of the cryo-EM map or the subunit model, e.g. in the crystal structure vis-à-vis the solution structure, or are indicative of conformational changes induced in forming the complex. To confirm the orientations of components in complexes, heavy-metal clusters can serve as valuable site-specific markers.

Phasing X-ray crystallographic data sets with EM-derived molecular envelopes. Diffraction-quality crystals of large complexes call for different phasing methods than are applicable to crystals of typically sized globular proteins. For instance, heavy atoms might lack tomographic data sets. Rather complex protocols now allow correction for the imperfections of the experimental setup, while minimizing the cumulative electron dose. Data sets of 100–200 projections can be recorded at doses low enough not to visibly damage ice-embedded specimens.

As applied to single particles, the method has produced a 3D structure of the archaeal thermosome, a hexadecameric chaperone, that was good enough to allow docking of its domains of known structure (Fig. 3). Tomographic density maps of numerous individual particles were calculated, screened to remove outliers, and averaged. In this particular case the same result could, no doubt, have been obtained by single-particle methods – in particular, it will be valuable to impose calibrated figures of merit for the fits, and to develop criteria for whether discrepancies reflect shortcomings of the cryo-EM map or the subunit model, e.g. in the crystal structure vis-à-vis the solution structure, or are indicative of conformational changes induced in forming the complex. To confirm the orientations of components in complexes, heavy-metal clusters can serve as valuable site-specific markers.
the scattering power to alter the structure factors sufficiently to allow solution by isomorphous replacement. Alternatively, such crystals might yield to a phase extension approach based on exploiting molecular envelopes defined by cryo-EM, supplemented by noncrystallographic symmetry. If the number of complexes and their positions in the unit cell are known, the envelope specifies those portions of the unit cell whose density is at a constant level. This provides a ‘solvent flattening’ constraint, which – with the additional constraint of noncrystallographic symmetry – is imposed iteratively, starting with a low-resolution model (e.g. the molecular envelope from cryo-EM). This model is progressively embellished by introducing diffraction data to higher and higher resolution, and reimposing the constraints at each cycle.

To date, this approach has worked well with icosahedral viruses with 5–3–2 symmetry, for example Norwalk virus\textsuperscript{43} and bacteriophage HK97\textsuperscript{44}, the latter study revealing a novel capsid protein fold; and for ClpP, a barrel-like protease consisting of two heptameric rings with 7–2 symmetry\textsuperscript{45}. The method appears to have much potential for the analysis of large complexes. A line of enquiry bearing on its applicability is to determine the minimum constraints needed for such phasing schemes to converge to stable, correct solutions, and the extent to which the requirement for a high order of symmetry might be relaxed or compensated by improved resolution in cryo-EM density maps. Work on the ribosome – which lacks internal symmetry – has demonstrated that EM-derived information can be usefully exploited in crystallographic studies of such particles\textsuperscript{46}.

**Visualizing ‘excited’ states of macromolecular complexes.** The reaction cycles and assembly pathways of macromolecular complexes often involve passing through short-lived, structurally distinct states, concomitant with the binding and processing of ligands or cofactors. Elucidation of such elaborate phenomena requires structural knowledge of each step in the pathway. If a complex is amenable to crystallization, it is likely to be captured in its most stable ‘ground state’, whereas the precursor or transitional states, as ‘excited states’, can be more elusive\textsuperscript{47}. It can also happen that salting out of protein complexes into crystals fixes them in some nonphysiological configuration\textsuperscript{48}. However, intermediate states of complexes in solution can be arrested by cryo-EM, using rapid-freezing\textsuperscript{49} or by adjusting buffer conditions to extend the lifetime of a given state or its prevalence in mixed populations\textsuperscript{50}. In such situations, an atomic model of the ground state may be adapted into pseudoatomic models of transitional states by assuming that the main structural elements – domains or subunits – are conserved, but undergo rigid-body movements. This assumption reduces the number of free parameters describing the intermediate state to the point that they can be stably determined from a cryo-EM reconstruction at moderate resolution. However, the possibility remains that such transitions might also involve local refolding or remodeling, and this eventuality can ultimately be addressed only by improving the resolution of the reconstruction or obtaining a crystal structure.

Among other applications, this approach – integrating cryo-EM and X-ray crystallography – has been used to investigate the ‘cell-entry intermediate’ states of poliovirus (Fig. 4). Compared with the native virus, these particles exhibit sizeable movements of their β-sandwich domains that create gaps between the capsid proteins through which RNA and internal peptides might exit\textsuperscript{51}. Another incisive example has been the domain movements elicited when the chaperone GroEL switches between nucleotide states\textsuperscript{52}. This combined approach has great potential in, among other topics, structural studies of ribosomes in which atomic models are emerging from crystallography (e.g. Ref. 53). Cryo-EM is well placed to exploit this information in elucidating the translation cycle by visualizing ribosomes interacting with elongation and initiation factors, nascent peptides, etc. (e.g. Refs 54,55) (Fig. 5).

**Cryo-EM and structural genomics**

The goal of structural genomics is to determine the structures of the proteins identified in genome sequences, or domains thereof; in particular, to solve representatives of each topologically distinct fold. With so many targets, emphasis has been put on automation to achieve high throughput in expression and crystallization trials. Doubtlessly, structural genomics will yield many domain structures, both because domains tend to be easier to crystallize than intact proteins and because they are of a size amenable to NMR. Although this massive repository of structural information will surely contribute to solving the protein folding problem, its implications for a mechanistic understanding of biological functions are harder to predict, as most cellular functions are carried out, not by individual domains but by large complexes\textsuperscript{56}. The latter entities do not readily lend themselves to
high-throughput tactics because methods do not yet exist for mass production by co-expression and assembly of large numbers of gene products. Thus there is a distinct need for methods that yield overall structures of large complexes that might be available only in limited supply, or detailed structures of other classes of automation-unfriendly macromolecules such as membrane proteins.

In this context, further technical progress coupled with systematic integration with bioinformatics should allow cryo-EM to emerge as a major player in structural genomics. There are two aspects to be considered: first, accelerated exploitation of state-of-the-art methods; and second, further extension of resolution. In either event, desiderata include (for single-particle analysis): automation of data collection\(^5\); streamlining and standardization of analysis procedures; and acquisition of primary data from the most powerful microscopes currently in operation. To make these resources available, there appears to be a persuasive case for establishing a number of cutting-edge installations to serve the cryo-EM community in much the same way as synchrotron beamlines meet the needs of crystallographers.

Computational advances will also be essential, and further refinement, integration and standardization of data analysis must be achieved in order to distill the information from the very large data sets that are needed, for example \(\geq 10^5-10^6\) particles. Here, parallel computing\(^5\), which lends itself to this application, should be a major factor. A goal of -4.5-Å resolution in single-particle analyses appears realistic, for at least some classes of molecules, and should suffice to resolve strands in \(\beta\) sheets and the connectivity of secondary structure elements, and consequently, the backbone folds of these proteins.

**Challenges for bioinformatics: enhancing resolution.** Exciting as the above prospects are, it appears unlikely that cryo-EM of single particles will advance in the short term to resolutions at which side-chain configurations and the binding of water molecules can be assessed (say, 2-2.5 Å). In this context, an intriguing proposition is whether it might be possible to draw empirically on the growing database of high-resolution X-ray structures to refine a structure solved at, say, 4.5-Å resolution in which the path followed by the backbone is mapped out, to one at 2.5 Å in which the stereochemistry is more fully defined. Although certainly challenging, this might turn out to be a more achievable goal than *de novo* structural prediction from amino acid sequences.

A second attractive possibility is to develop methods that scan the database of high-resolution domain structures to discriminate domains in EM-derived density maps at the currently achievable resolutions of 7-15 Å, thereby allowing the maps to be interpreted in much greater detail than these numbers would imply. In this notion, the problem of structural determination is reduced to one of pattern recognition in which the elements to be recognized are domains, not amino acids. Because domains are much larger than amino acids, and the
number of distinct domain folds is not inordinate, they should, in principle, be recognizable in density maps at much lower resolution. 

Visualizing molecules inside cells. It has long been a dream of biologists to see the molecular architecture of living cells. In addition to robust machines like the ribosome, there are undoubtedly many ‘teams’ of interacting molecules performing cellular functions, that are held together by forces too weak to survive biochemical isolation. To this end, recent developments in automated electron tomography have made it possible to obtain 3D reconstructions of whole ice-embedded prokaryotic cells or to obtain 3D reconstructions of whole cells. Electron tomography has now advanced to the point where it is a realistic goal to glimpse molecular machines operating inside cells: again, value will be added by integrating higher-resolution data obtained on isolated complexes.

Conclusions

There is growing awareness that many key cellular functions are performed by ensembles of macromolecules – molecular machines. Whereas some machines are abundant and robust, like the ‘monolithic’ ribosome, others are present in only a few copies per cell or are too delicate to survive isolation procedures, ruling out X-ray crystallography or NMR spectroscopy as means to determine their structure. Also, during their reaction cycles, machines pass through short-lived altered states that will be, at best, difficult to capture by crystallography. In this context, there is immense opportunity for cryo-EM, especially as boosted by merging crystallographic structures of individual subunits into moderate-resolution cryo-EM density maps of whole complexes. Electron tomography has now advanced to the point where it is a realistic goal to glimpse molecular machines operating inside cells: again, value will be added by integrating higher-resolution data obtained on isolated complexes.

Acknowledgements

We thank J. Bohm for help with graphics, and D. Rees and P. Bjorkman for helpful comments.

References

Observing proteins in their natural habitat: the living cell

Philippe I.H. Bastiaens and Rainer Pepperkok

Fluorescence microscopy has played a tremendous role in uncovering the morphological features of cells and the expression pattern of proteins by immunofluorescence. Since the discovery of green-fluorescent proteins (GFPs), this technique has undergone a revival in the life sciences as the spatial distribution of ectopically expressed fusion proteins inside living cells can now be followed more easily. By further exploiting the photophysical properties of the emitted fluorescence with microspectroscopic methods, spatial information on the biochemical parameters of intracellular processes and reactions can be obtained. This possibility will not only play an important role in the understanding of biochemical reactions in signal processing and fidelity but also help to uncover the molecular mechanisms of organelle and cell morphogenesis.

THE CELL CAN be conceived as a biochemical information-processing device whose response to the environment depends on the state of spatially organized networks of protein activities. Localized protein interactions, covalent modifications, proteolytic processing and conformational changes are the weighted information transfer functions by which the state of these complex molecular networks is determined. The interconnectivity and spatial organization of protein systems that sustain basic cellular functions is only maintained in the context of the whole intact molecular architecture of the cell. It is clear that understanding cell function by integrating molecular activities with spatial organization is an important challenge for modern biology.

Numerous biochemical assays revealing protein modifications, interactions or transport have already helped to bring us closer to this goal. However, none of the methods that are based on reconstituted systems in the test-tube can fully take into account the compartmentalized and interconnected nature of these reactions in cells. It is therefore imperative to develop methods that can measure the dynamics of these biochemical reactions in the intact cell and thereby extract the spatial organization in vivo. Only recently, and most probably because of the availability of genetically encoded fluorescent proteins, has it become possible to image not only cellular processes such as protein translocation or transport but also basic reactions such as protein interactions, proteolysis and phosphorylation in intact cells.

Fluorescent proteins for live cell imaging

Fluorescence imaging of biological reactions in living cells requires techniques to fluorescently label the macromolecules involved. These fluorescent probes need to be specific, interfere as little as possible with the reactions to be visualized, must not perturb the physiological conditions of the cells, and must have favourable spectroscopic properties for efficient detection. For many years the only way to study a fluorescent protein in living cells was to purify, chemically modify and then (re)introduce it into living cells by, for example, capillary microinjection. Unfortunately, the widespread use of this approach has been inhibited by technical limitations. For example, it is often difficult to label larger proteins fluorescently because they are more difficult to express and purify in sufficient amounts by recombinant methods. Furthermore, transmembrane or luminal proteins of the secretory pathway are difficult or impossible to label because the injected, fluorescent proteins do not translocate into the lumen of the endoplasmic reticulum or the endogenous protein is not accessible to fluorescent antibodies. Therefore, one of the most exciting advances in fluorescent in vivo labelling techniques has been the discovery of the green-fluorescent protein (GFP) and its spectral variants. Because GFP is genetically encoded and emits visible fluorescence without the need for cofactors, it has been used extensively as a fluorescent tag fused to cloned proteins. In principle, any cDNA of interest can be fluorescently labelled with GFP. Therefore, the number of applications of GFP in combination with fluorescence time-lapse microscopy has been enormous. This has helped towards a better understanding of the dynamics of the organization and architecture of living cells or organisms.

Although GFP tagging of proteins has clear advantages, the GFP technology...