Abstract: Scanning transmission electron microscopes can both measure the mass of single protein complexes and take amazingly clear images thereof, offering a wide range of applications in structural biology. The principle of mass measurement is presented and discussed and the scope of scanning transmission electron microscopy is illustrated by selected examples.

Keywords: Mass determination · Scanning transmission electron microscopy · STEM

Introduction

The invention of the scanning tunneling microscope in 1982 by Binnig and Rohrer [1] opened the door to direct investigation of the nano-world. A wide spectrum of nanotechnologies has since been developed and applied to answer biological questions [2]. In contrast the scanning transmission electron microscope (STEM) was already employed to make nano-scale measurements on biological samples in the 1970s [3][4] and retains its importance today (for reviews see [5][6]). Key was the development of the field emission gun and its incorporation in the STEM [7]. This electron source delivers a highly coherent beam of electrons that can be focussed to a spot less than 0.5 nm in diameter. In a calibrated instrument [8][9] these beam characteristics allow the mass of proteins as small as a few nanometres in size to be determined from their scattering power. The necessary requirement of a highly tuned dark-field detector system able to count single electrons allows the unstained proteins to be clearly visualised under low dose conditions. Each pixel of the digital image is a quantitative scattering experiment from which the mass of the irradiated volume can be calculated [9]. Thus, the mass of a single protein complex can be determined by integrating over all pixel values that lie within its boundary. The method is applicable to macromolecules of widely varying mass and form (reviewed in [10]). When combined with SDS-gel electrophoresis and, where available, sequence information, STEM mass measurements serve to define protein stoichiometry. Further, as electron energy loss spectroscopy [11][12] or energy dispersive X-ray spectroscopy [13] can be performed while imaging, the STEM also offers the possibility of determining the distribution of specific elements in protein complexes. Indeed the feasibility of mapping single atoms bound to protein molecules has recently been demonstrated [14].

The high contrast delivered by the dark-field detector system [15] can also be exploited in negative or positive stain microscopy [16]. Here the STEM yields clear single-shot images that are free from phase contrast fringes. These images frequently provide information only otherwise visible after averaging several hundred projections recorded with a transmission electron microscope (TEM). They thus make it possible to document the presence of distinct protein conformations, information that would be lost by averaging techniques. Indeed, STEM images of both negatively stained [17] and unstained [18] protein complexes have been used to reconstruct their 3D structure.

The few dedicated STEMs employed in biology today are invaluable. In particular, the coupling of image and mass gives them the power to answer questions not resolvable by ultracentrifugation or mass spectrometry. Here we discuss the use of the STEM both as an imaging tool and to measure mass, and examine the uncertainties to be expected in STEM mass data sets.

The Importance and Versatility of STEM Mass Measurements

Worldwide, only four laboratories routinely use dedicated STEMs to measure the mass of biological samples. Their facilities are open for external collaborations and the many papers in the literature containing STEM mass measurements document the importance of this rare technique to the scientific community. Although the precision of STEM cannot match that of the mass spectrometer, its capability to measure an enormously wide mass range, its independence from shape assumptions and the presence of an image make the STEM an invaluable tool.

The methodology of mass spectrometry has developed to a high level of sophistication over the last decade [19][20], allowing the proteins present in hetero-oligomeric complexes to be identified. In combination with electron microscopy, in particular with STEM as an initial step, this opens an important new avenue that could result in the full structural definition of protein complexes. STEM measurements provide the first link in this chain, as the total mass of individual protein complexes is determined. In com-
bination STEM measurements would both reveal whether the proteins identified by the bulk mass spectrometry technique were part of a single complex or a number of complexes that collectively have the determined composition, and allow their stoichiometry to be calculated.

The mass range over which STEM can be applied is enormous. Proteins as small as 100 kDa and larger than 100 MDa [21][22] have been successfully examined and this independent of their shape. STEM is also the only technique that not only allows the total mass of a large protein complex to be directly determined, but also the mass-per-length of filamentous structures and the mass-per-area of planar structures such as surface layers [23][24] or membrane assemblies. The mass-per-length is a useful if not essential parameter defining the structure of filaments [25], indicating the number of strands present [26–29] and in the case of helices restricting the number of possible assembly rules [30–32]. Similarly, the mass-per-area defines the unit cell stoichiometry in 2D crystals or the number of layers in a sheet-like structure [33–36], information that is otherwise only accessible by chance at their edges, or by atomic force microscopy [36][37]. Further, in conjunction with a lipid assay by analytical ultracentrifugation mass-per-area measurements can be used to confirm the predicted packing of membrane proteins in two-dimensional crystalline arrays [35]. Finally, the presence of an image allows the mass of specific, well defined regions of intact complexes of any shape to be determined and a mass map to be generated [22][23][38–45].

The Principle

In the STEM, an electron beam is focused to an atomic scale probe and raster scanned over a thin sample to assess the structure of biological matter. Theoretical data on elastic scattering are sufficiently precise to permit the calculation of the absolute mass scale for proteins [8], making a fully calibrated instrument [9] independent of mass standards. The error of the measurement is given by the statistics of single electrons counted by the annular detector and by the cleanliness of the carbon film.

A theoretical assessment of the statistical noise related to the low-dose requirement is based on the assumption of spherical protein complexes of mass \( M \) and density \( \rho \) = 0.82 Da Å\(^{-3}\), being composed of atoms with an average atomic mass of \( < M_{ad} > \) = 13.42 Da that have an average scattering cross section \( <\sigma> \) = 0.0102 Å\(^{-2}\), which holds for 80 kV electrons [8]. The radius \( r \) and cross section \( A \) of such proteins (taking the mass in Daltons and the length in Å) are:

\[
r = \sqrt[3]{3M/(4\pi\rho)} = 0.6628\sqrt[3]{M} \tag{3}
\]

\[
A = \pi r^2 = 1.38(\sqrt[3]{M})^2 \tag{4}
\]

The fraction \( \eta \) of the incident electrons scattered onto the annular detector (\( \epsilon = 0.69 \)) by \( N = M/\langle M_{ad}\rangle \) atoms of the protein and the total number of electrons \( N_{Sprot} \) incident on the annular detector when the same protein is irradiated by dose \( D \) (in electrons/Å\(^2\)) are then:

\[
\eta = M\epsilon <\sigma>/ <\sigma, Ma > 1.38(\sqrt[3]{M}) \tag{5}
\]

\[
N_{Sprot} = \eta D M\epsilon <\sigma>/ <\sigma, Ma > = 5.24 * 10^{-4} MD \tag{6}
\]

For mass measurements, samples are adsorbed to thin carbon films and must necessarily be left unstained i.e. the standard negative staining technique of electron microscopy whereby heavy ions (e.g. uranyl or phosphotungstate ions) are used to enhance contrast, cannot be employed. To minimize background, all traces of non-volatile salts are subsequently removed by a series of washes with quartz double-distilled water or, where instability of the sample demands it, with volatile buffer solution, e.g. 10–100 mM ammonium acetate. This is critical because the background can be the largest source of random errors and can thus significantly increase the scatter of the measurements. Freeze-drying generally follows. Images are recorded at low electron dose, typically 3–4 electrons/Å\(^2\), to minimize beam-induced mass-loss. When measurements are made at ambient temperature, the small mass-loss correction necessary is experimentally determined [9].

Image evaluation requires summation over all picture elements covered by the structure of interest, followed by normalization to the recording dose, and subtraction of the background (i.e. the carbon film scattering). This allows the mass of particles, the mass-per-length of filamentous structures and the mass-per-area of layers to be calculated. The statistical error for a homogeneous preparation of a spherical protein complex on a thin thick clean carbon film is given by the standard deviation, \( SD \), where:

\[
SD = \sqrt{SD_{tot}^2 + SD_{bac}^2}
\]

with \( SD_{tot} = \sqrt{SD_{prot}^2} \) and \( SD_{bac} = \sqrt{SD_{bac}} \)

\[
SD_{tot} = 0.166 * D(\sqrt[3]{M})^2 \tag{7}
\]

\[
SD_{bac} = 0.166 * D(\sqrt[3]{M})^2 \tag{8}
\]

From this, we calculate the \( SD \) to be 68 kDa for a typical protein complex of 1 MDa mass that is imaged at a dose of three electrons/Å\(^2\). As a result of additional sources of error (contaminants on the carbon film, dissociation of the protein complex, variable amounts of detergent bound per complex), the experimental \( SD \) tends to be larger. Nevertheless, under ideal conditions, such standard deviations have been achieved. Thus only a relatively small numbers of particles, in general a few hundred, need to be measured to obtain a standard error that is better than the instrument’s calibration (typically \( \pm 2\% \)).

Applications of STEM

Over the years both dark-field imaging and mass measurement by scanning transmission electron microscopy have frequently been employed to address specific questions not directly accessible by other techniques. The results discussed below have been selected to illustrate the versatility of the technique. They focus on components of the Type II (T2) and Type III (T3) secretion systems used by the Gram-negative bacteria *Klebsiella oxytoca* and *Yersi-
nia enterocolitica, respectively, to transfer certain proteins from their cytoplasm to the exterior environment.

The single, integral outer membrane proteins of T2 and T3 secretion systems belong to a large superfamily of homologous proteins, the secretins [46]. These are present in the bacterial outer membrane as homo-oligomers. STEM has been employed to define the oligomerisation state of two; the T2 secretin complex of K. oxytoca formed by the protein PulD and the corresponding T3 complex of Y. enterocolitica formed by the protein YscC [47][48]. In the more extensive study made on the K. oxytoca system [48][49], STEM mass measurements indicated the presence of twelve monomeric subunits in the intact PulD complex and in its trypsin-resistant fragment (Fig. 1a; [48]). They also showed this stoichiometry to be retained when the secretin was associated its pilotin PulS [49]. In contrast, the T3 secretin complex of Y. enterocolitica was shown to be a 13-mer (Fig. 1b; [47]). Given their mass, the two minor peaks on this histogram arise from the aggregation of two and three complexes, respectively. As documented by the insets, this interpretation was confirmed by sorting the corresponding particles according to mass and inspecting the resulting galleries. The possibility to link mass and shape is unique to STEM and allows complex mass histograms obtained from heterogeneous samples to be interpreted [50].

The secretin samples were also examined by negative stain microscopy to obtain structural information. As illustrated by Fig 1c and d, the single-shot STEM images revealed details that only became visible on averaging cryo-electron microscopy (cryo-EM) or conventional negative stain transmission electron microscopy (TEM) projections. The top-views showed the PulD and the YscC secretins to form large ring-shaped complexes with a central channel (Fig. 1c, d) and revealed peripheral densities around the PulD ring that were lost on trypsin treatment. A few single complexes of the secretin PulD also adsorbed side-on giving a view of the central channel. In contrast, the trypsin-resistant PulD fragment was only seen from the side when two complexes had associated via their more ‘conical’ ends (Fig. 1c). Similarly YscC secretin complexes only adsorbed side-on when aggregated. In this case dimers formed by the interaction of the flatter ends, and single-shot STEM images revealed an inherent flexibility of the ‘conical’ region not detected by averaging procedures (Fig. 1d).

The secretin complex is itself thought to form the translocation pore used in T2 secretion, but how the translocation occurs is not known. In K. oxytoca translocation is possibly facilitated by a short and as yet undetected, periplasmic pseudopilus produced on over-expression of PulG(His)_{6} [32]. Indeed, images recorded from negatively stained samples directly revealed their periodic helical structure (Fig. 2a). As expected, straightened filament stretches yielded typical x-shaped diffraction patterns, allowing the pitch of the helix to be determined (Fig. 2b). The mass-per-length (Fig. 2c) measured for corresponding unstained samples by STEM, was key to the following image analysis, restricting the number of helical rules applicable. Accordingly, there were on average 17 PulG subunits in four turns of the left-handed helix [32]. This information allowed a 3D model of the pseudopilus to be generated based on the X-ray structure determined for a truncated form of PulG (Fig. 2d; [32]).

In contrast to T2 secretion systems, the T3 secretin of Y. enterocolitica is penetrated by a needle complex that extends into the exterior milieu. In a recent study carried out to investigate the structure of the needle, STEM images clearly revealed a complex at
its tip (Fig. 3a; [51]). Combined with further negative stain STEM, mutagenesis and immuno-labeling experiments showed this to be formed by the protein LcrV (also known as V-Antigen). First, the distinctive tip-complex was absent from the needles of an lcrV mutant strain (Fig. 3b); instead the needles had rather pointed ends. Second, polyclonal antibodies against LcrV directly visualized by STEM were seen to interact exclusively with the tip-complexes often linking two needles together via their variable domains (Fig. 3c). Particularly the latter experiment illustrates the power of STEM, which in contrast to TEM is able to clearly visualize and localize single antibodies without the use of gold tags. The presence of LcrV at the very tip of the Yersinia needle explains why this protein is essential for pore assembly in the host cell membrane and infection. Averaging a series of high signal-to-noise images of the tip-complex clearly revealed features only faintly, if at all, visible on the unprocessed STEM images, showing the needle to have a central channel that continues into the tip-complex (Fig. 3a, inset).

**Conclusion**

As illustrated by the above examples, quantitative biological STEM is a well-developed and invaluable technology for nano-analytics. As it is possible to associate mass with the shape of individual protein complexes (Fig. 1b) it provides a complementary extension of mass spectrometry to be exploited in the future. The ability of STEM to distinguish between different proteins by means of their mass promises to make this technique of vital importance to many systems biology projects.

**Abbreviations**

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ε</td>
<td>collection efficiency of the detector</td>
</tr>
<tr>
<td>η</td>
<td>fraction of electrons scattered onto the annular detector</td>
</tr>
<tr>
<td>A</td>
<td>cross section of a spherical protein complex</td>
</tr>
<tr>
<td>r</td>
<td>radius of a spherical protein complex</td>
</tr>
<tr>
<td>D</td>
<td>incident electron dose (electrons per unit area)</td>
</tr>
<tr>
<td>M</td>
<td>mass</td>
</tr>
<tr>
<td>&lt;M&gt;</td>
<td>average atomic mass of biological matter</td>
</tr>
<tr>
<td>N</td>
<td>number of atoms</td>
</tr>
<tr>
<td>SAD</td>
<td>annular detector signal</td>
</tr>
<tr>
<td>Sbac</td>
<td>standard deviation of Sbac</td>
</tr>
<tr>
<td>Stot</td>
<td>total signal, Sprot + Sbac</td>
</tr>
<tr>
<td>SD</td>
<td>standard deviation</td>
</tr>
<tr>
<td>SDtot</td>
<td>standard deviation of Stot</td>
</tr>
</tbody>
</table>

**Acknowledgments**

This work was supported by the Maurice E. Müller Foundation of Switzerland, by the Swiss National Foundation (SNF) within the framework of the National Centre of Competence in Research for Structural Biology, SNF grant 3100-059415 to AE, and the NoE 3D-EM, EU project (LSHG-CT-2004-502828).

Received: September 1, 2006


Single Biological Molecules’, University of Chicago, IL, 1971.


