

# Atomic-resolution STEM at low primary energies

Ondrej L. Krivanek<sup>1</sup>, Matthew F. Chisholm<sup>2</sup>, Niklas Dellby<sup>3</sup> and Matthew F. Murfitt<sup>4</sup>

<sup>1</sup> Nion Co., krivanek@nion.com

<sup>2</sup> Oak Ridge National Laboratory, chisholmmf@ornl.gov

<sup>3</sup> Nion Co., dellby@nion.com

<sup>4</sup> Nion Co., murfitt@nion.com

**Abstract.** Aberration correction of the scanning transmission electron microscope (STEM) has made it possible to reach probe sizes close to 1 Å at 60 keV, an operating energy that avoids direct knock-on damage in materials consisting of light atoms such as B, C, N and O. The improved resolution is allowing individual atoms to be imaged in various novel materials including graphene, monolayer boron nitride and carbon nanotubes. Some radiation damage remains even at the lower energies, and this limits the maximum usable electron dose. Elemental identification by electron energy loss spectroscopy (EELS) is then usefully supplemented by annular dark field (ADF) imaging, for which the signal is much greater. Because of its strong Z dependence, ADF allows the chemical identification of individual atoms, both heavy and light. We review the instrumental requirements for atomic resolution imaging at 60 keV and lower energies, and we illustrate the kinds of observations that have now become possible by ADF images of graphene, monolayer BN and single wall carbon nanotubes, and by ADF images and EELS spectra containing nanopods filled with single atoms of Er. We then discuss likely future developments.

## 1 Introduction

The scanning transmission electron microscope (STEM) is now able to reach a resolution close to 1 Å at 60 keV, a primary energy that is low enough to avoid direct knock-on damage in materials made of lightweight atoms such as carbon. There are no regular inter-atomic distances smaller than 1.2 Å not involving hydrogen, and 1 Å resolution therefore means that near-neighbor atoms can be resolved in such materials without destroying the observed structures by rapid knock-on damage. This is an important advance, and the main subject of this chapter.

Many researchers have contributed to the advance. The full history of the STEM is reviewed in depth elsewhere in this volume (Pennycook 2010); here we only note the major mileposts encountered along the way. Crewe's cold field-emission STEM (Crewe, Wall, and Welter 1968, Crewe 2009) was the key development, because it showed the wealth of results that can be obtained when a small probe of electrons is focused on a thin sample and several signals, such as those due to elastically and inelastically scattered electrons, are collected simultaneously.

Once their new STEM was working well, Crewe's group progressed quickly onto imaging of single heavy atoms – a significant first in electron microscopy (Crewe, Wall and Langmore 1970) – and also to a demonstration of the power of Electron

Energy Loss Spectroscopy (EELS) carried out in the STEM (Isaacson and Johnson 1975). Their work led to the two Cornell workshops on Analytical Electron Microscopy (Fraser et al. 1976, Fejes 1978), which spearheaded the adoption of the STEM technique by a large number of researchers, and largely defined the agenda for the development of the STEM over the next two decades.

It is useful to remember that Crewe's work was preceded by the development of the STEM concept by von Ardenne (1938, 1940, 1986), and a demonstration by Zworykin, Hillier and Snyder (1942) that a cold field emitter can be a useful source of electrons. But these developments were "ahead of their time": the detectors, the electronics, the ultra-high vacuum technology and the computers necessary for a practical STEM did not yet exist.

The resolution reached by the Chicago STEM was 2.4 Å at 30 kV, limited by the spherical aberration coefficient  $C_s$ , even though the coefficient had been reduced to the low value of  $C_s = 0.3$  mm. To improve the resolution further, either the spherical aberration had to be eliminated, with the help of an aberration corrector, or the operating voltage had to be raised significantly, so as to decrease the electron wavelength  $\lambda$ . Crewe's lab embarked on both approaches, but did not succeed. In the aberration corrector field, they were probably themselves "ahead of their time", because computers were not up to the task of analyzing and correcting aberrations automatically in the 1970s. Raising the operating energy significantly may have needed more funding than they were able to secure.

The field had grown much wider from the late 1970s on. Others began to make significant contributions (Hawkes 2009), e.g. by producing a STEM capable of 1.2 Å resolution at 300 keV. But to approach 1 Å-level resolution at operating energies <200 keV, a working aberration corrector was needed.

The first proof-of-principle corrector for an electron probe instrument was built by Deltrap (1964a, 1964b), based on concepts introduced by Archard (1955). Deltrap showed that three octupoles acting on an electron beam in conjunction with four quadrupoles can make spherical aberration in a probe-forming instrument zero or negative. However, he did not demonstrate an improved resolution. We now know that abandoning cylindrical symmetry together with the imperfect nature of quadrupoles (which are much harder to construct with sub-micron precision than round lenses) produces parasitic aberrations of many different orders, and that if these aberrations are not corrected along with the spherical aberration, the resolution is likely to become worse than it was before the correction attempt. Deltrap's corrector did not have enough flexibility to null the parasitic aberrations, and the task would not have been an easy one without modern-style computer control even if it did.

The correctors built in Crewe's lab were not as successful as Deltrap's advance, and no correction of aberrations was demonstrated at all. This was probably at least partly due to the fact that in the first corrector built by Beck and Crewe (1976) whose electron-optical design was similar to a  $C_s$  corrector designed by Thomson (1968), the magnetic circuit was built from permendur, whose high remanence and hysteresis made it totally unsuitable for producing the low-magnitude fields that needed to be controlled with high precision in a corrector. But Crewe's group contributed greatly to the theoretical understanding of corrector optics: the main concepts of a sextupole

corrector and of a 5-th order-optimized corrector were introduced by them (Beck 1979, Crewe and Kopf 1980, Shao 1988).

The next attempt at correcting spherical aberration in a STEM was by Krivanek, Dellby and coworkers, with a quadrupole-octupole corrector (Krivanek et al. 1997). Their work demonstrated that with a sufficient number of computer-controlled auxiliary optical elements and newly developed aberration-diagnosing software, parasitic aberration could be mastered and the resolution of the STEM the corrector was built into could be improved relative to its pre-corrector value. This group then moved onto a second-generation corrector design (Krivanek, Dellby and Lupini 1999, Dellby et al. 2001), which became the first commercially available corrector, and which allowed a more optimized STEM to reach directly interpretable sub-Å resolution, for the first time in electron microscopy (Batson, Dellby and Krivanek 2002).

At about the same time and on a parallel track, a sextupole corrector design due to Rose (1990) was developed into a practical aberration corrector for the TEM by Rose's students and collaborators (Haider et al. 1998). The same group also developed a practical aberration corrector for an SEM (Zach and Haider 1995), using a Rose (1971) design whose principles were similar to a  $C_s/C_c$  probe corrector designed and built by Hardy (1967).

With aberration correction thus attained in all the three principal types of electron microscopes, further progress in aberration-corrected STEM has been rapid. Resolution records have been bested repeatedly (Nellist et al. 2004, Sawada et al. 2007, 2009, Erni et al. 2009), the improvements coming principally from the use of higher primary energies (and thus smaller electron wavelength  $\lambda$ ). Sextupole+round lens STEM correctors were shown to be capable of similar probe-forming performance as quadrupole+octupole ones (Müller et al. 2006), the resolution of both types being typically limited by uncorrected chromatic aberration  $C_c$  that comes primarily from the objective lens.

The new instruments made possible many new types of materials studies. Elemental mapping by EELS at atomic resolution (Bosman et al. 2007, Muller et al. 2008) has become a standard tool, single atoms with large EELS cross-sections are now routinely detected (Varela et al. 2004, Suenaga et al. 2009), and the 2-D variation in the electronic bonding can be studied with atomic resolution (Muller et al. 2008, Fitting-Kourkoutis 2010). The availability of atomic resolution at 60-80 kV primary energy, i.e. low enough to avoid significant displacement damage in light-Z materials such as graphene and BN (Zobelli et al. 2007) has produced a wealth of results in these materials (Girit et al. 2009, Meyer et al. 2009, Jin et al. 2009, Alem et al. 2009, Krivanek et al. 2010a, 2010b). Several meetings and conference sessions have been devoted to aberration correction and its results, and several compendia of aberration-corrected work that include a large proportion of STEM results have been published (e.g. Hawkes 2008).

In this chapter, we focus on STEM at primary energies lower than 100 keV. We review the factors that determine the resolution and the STEM probe current, show examples of investigations that have now become possible, and discuss likely future developments.

## 2 Microscope performance at low primary energies

Predicting the attainable resolution has become more complicated in the aberration-corrected era. Spherical aberration no longer dominates, and the STEM resolution limit typically comes from one or more of the following:

- a) chromatic aberration (if not corrected)
- b) higher-order geometric or mixed geometric-chromatic aberrations
- c) finite brightness of the electron source
- d) finite size of the atoms
- e) statistical noise in the images
- f) instrumental instabilities.

Formulas for evaluating contributions a)-c) and e) numerically are given in Krivanek et al. (2008a). The principal resolution-limiting influences are discussed below.

### 2.1 Optimizing the resolution at low primary energies

Despite the large number of factors that need to be considered, the strategy for obtaining the smallest possible STEM probe is usually quite simple. When observing an experimental shadow image (Ronchigram, see for instance Krivanek et al. 2009a) of an amorphous sample, it is usually possible to tell which part of the Ronchigram will usefully contribute to the probe maximum and which parts will only contribute to the probe tail. The useful part of the Ronchigram is selected by an aperture, typically not located inside the objective lens itself, and therefore called the “virtual” objective aperture (VOA), or the “condenser aperture”. The diffraction limit due to the aperture then determines the smallest attainable probe size as:

$$d_{\text{aber}} = 0.61 \lambda / \alpha_{\text{aber}} \quad (1)$$

where  $\alpha_{\text{aber}}$  is the illumination half-angle admitted by the aperture, and  $\lambda$  is the electron wavelength. With a sufficiently precise knowledge of the relevant aberration parameters,  $\alpha_{\text{aber}}$  can of course also be calculated analytically, and detailed expression that consider geometric aberrations up to 7<sup>th</sup> order as well as first-order chromatic aberration are given in Krivanek et al. (2008).

In practice, in a well-tuned,  $C_s$ -corrected STEM with minimized or corrected fifth-order aberrations operating at 200 kV and below, the smallest attainable probe size is usually determined by the chromatic aberration. To reach the smallest probe, the illumination aperture is maximized by setting it to a half-angle given by (Krivanek et al. 2008a):

$$\alpha_{\text{aber}} = \alpha_{\text{chrom}} = 1.2 (\lambda / (C_c \delta E / E_0))^{1/2} \quad (2)$$

where  $C_c$  is the coefficient of chromatic aberration,  $\delta E$  the half-width of the energy spread of the primary beam, and  $E_o$  the primary energy. The minimum attainable probe size is then given by:

$$d_{\text{aber}} = 0.5 (\lambda C_c \delta E / E_o)^{0.5}. \quad (3)$$

Equations (1-3) do not consider the influence of the finite size of the electron source that must be projected onto the sample if the probe is to contain a non-zero current. For a useful probe current, i.e. a source demagnification value that is not infinite, the probe size is broadened to:

$$d_{\text{probe}} = (d_{\text{aber}}^2 + d_{\text{source}}^2)^{0.5} \quad (4)$$

where  $d_{\text{source}}$  is the diameter of the source projected onto the sample (assuming perfect (Gaussian) optics).  $d_{\text{source}}$  depends on the selected probe current  $I_p$  and the source brightness as:

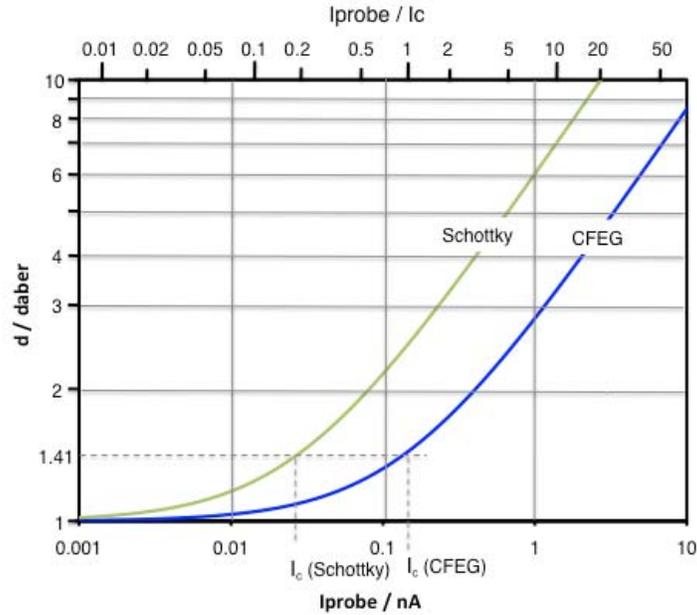
$$d_{\text{source}} = 2 (I_p / B_n V_o^*)^{0.5} / (\pi \alpha_{\text{aber}}), \quad (5)$$

where  $B_n$  is the “normalized” (or “reduced”) brightness  $B_n = B / V_o^*$ , and  $V_o^*$  is the relativistically corrected accelerating voltage.  $\alpha_{\text{aber}}$  appears in this expression because the probe current is proportional to the area of the Ronchigram selected by the illumination aperture. Combining equations (4) and (5) leads to an expression for the probe size with a non-zero probe current, which is applicable in the presence of any type of aberration, and at any primary voltage (Krivanek et al. 2008a):

$$d_{\text{probe}} = (1 + 7.3 \times 10^8 I_p / B_n)^{0.5} d_{\text{aber}} \quad (6)$$

in which the probe current is specified in nA and the normalized brightness in  $A/(m^2 \text{ sr V})$ . Fig. 1 shows the dependence of  $d_{\text{probe}}$  on  $I_p$  for two different normalized beam brightnesses:  $B_n = 1 \times 10^8 A/(m^2 \text{ sr V})$ , typical of cold field emission guns, and  $B_n = 2 \times 10^7 A/(m^2 \text{ sr V})$ , typical of field-assisted thermionic (=Schottky) guns. This graph is universally applicable to all STEM columns, corrected or not, operating at all primary energies.

The universality arises because the fractional deterioration of the probe size at larger beam currents compared to  $d_{\text{aber}}$  is independent of the optical performance of the microscope. When aiming to produce the smallest possible electron probe, one simply uses the largest illumination aperture for which the electron wavefront converging on the sample can be made not to deviate by more than  $\lambda/4$  from the ideal spherical wavefront. Keeping the wavefront deviations under this limit guarantees that the probe will not be broadened by the aberrations that are present. This means that the aberrations simply set the allowed size of the illuminating aperture, and that the rest of the imaging process is the same no matter what aberrations there are.



**Fig. 1.** Probe size as a function of the probe current for the typical CFEG and Schottky electron guns, calculated for  $B_n(\text{CFEG}) = 1 \times 10^8 \text{ A}/(\text{m}^2 \text{ sr V})$  and  $B_n(\text{Schottky}) = 2 \times 10^7 \text{ A}/(\text{m}^2 \text{ sr V})$ .

Another way of explaining Fig. 1 is to remember that the probe size is determined by the diffraction limit due to the selected aperture, provided that the aperture is coherently illuminated. Complete coherence would require that the electron source be infinitely small. Since a zero-size source cannot produce any current at all if its brightness is finite, a non-zero-size source is used, but typically small enough so that the source projected onto the sample is smaller than the diffraction limit due to the illuminating aperture.

When the projected source size equals the diffraction limit due to the illuminating aperture, the width of the small patch on the source that contributes electrons to the probe is equal to the spatial coherence width of the source. In other words, when the probe current is such that the probe is widened by  $\sqrt{2}$  times relative to the zero-current limit, the electrons in the probe come from a source patch whose width is the source's transverse coherence length. Such a probe current can therefore be called **the coherent current  $I_c$** . It is given by (in nA):

$$I_c = 1.4 \times 10^{-9} B_n \quad (7)$$

The above discussion makes it clear that the coherent current depends only on the value of the reduced brightness, and not on the optical performance of the column.

Typical coherent current values are  $I_c = 0.13 \text{ nA}$  for CFEG and  $I_c = 0.025 \text{ nA}$  for Schottky. Since using the full coherent probe current does cause a 41% broadening

of the probe compared to the diffraction limit, those aiming for the highest resolution typically only use probe currents of  $0.1 - 0.3 I_c$ . On the other hand, experimentalists needing to optimize the signal-to-noise ratio in noisy spectra often run with the source demagnified less strongly, e.g. at a beam current of  $3I_c$ , at which point the probe is 2x as wide as the diffraction limit, and is mostly incoherent. Increasing the beam current to  $3 I_c$  gives about 0.5 nA current for the typical CFEG and 0.1 nA for the typical Schottky gun.

The typical coherent current values for the two types of high brightness guns are also shown in Fig. 1. The graph can be made universal for all values of  $I_c$  simply by rescaling the abscissa to show the probe current value as a fraction of  $I_c$  rather than as an absolute current. This scaling is implemented at the top of Fig. 1, for the CFEG curve.

When operating in the large-current regime, the illumination aperture can usually be opened up a little to provide an even bigger beam current, without the additional aberration-caused broadening affecting the probe more than the increased source size. This adjustment is capable of giving about a 2x increase in the probe current before the additional aberration-caused broadening becomes objectionable.

Counteracting the above is the fact that because the solid angle of the beam emitted from the source that ends up contributing to the final probe on the sample grows at higher beam currents, the diameter of the used part of the beam emitted from the tip becomes larger at high beam currents. With a wider beam traversing them, the contributions of the gun and the condensers to the total aberrations of the optical system then increase in importance. At some value of the beam current, typically around 5-10  $I_c$  for the CFEG, these gun and condenser aberrations begin to dominate the total aberrations of the probe-forming system. The aberration corrector then needs to worry more about correcting the aberrations arising in the gun and the condensers than those of the final probe-forming lens. The exact point at which this happens depends on the details of the gun optics, and a lens that restricts the spread of the beam coming out of the gun (a “gun lens”) is useful for delaying the onset of the effect.

Equation (6) can be expressed, in the case of a STEM whose optical performance is limited by chromatic aberration, by a simple practical equation::

$$d_{\text{probe}} = 5.5 \times 10^5 \left( (1 + I_p / I_c) C_c \delta E \right)^{0.5} / E_o^{*0.75} \quad (8)$$

where  $d_{\text{probe}}$  is in pm,  $E_o^*$  is the relativistically corrected primary energy (in eV),  $I_p$  and  $I_c$  can use any unit as long as it's the same for both of them, but are most conveniently specified in nA,  $C_c$  is in mm, and  $\delta E$  in eV. In our experience, the equation predicts the attainable probe size in a  $C_c$ -limited STEM with about 10% accuracy in the low-current regime, and within about 40% in the high-current regime, in which there complicating factors such as sometimes being able to increase the size of the illuminating aperture slightly without greatly worsening the resolution, sometimes needing to decrease the aperture size because of extra gun and condenser aberrations, and not knowing the coherent current sufficiently accurately.

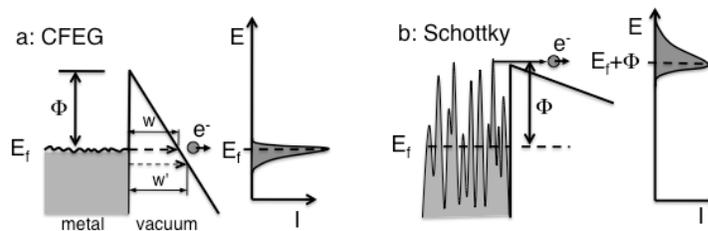
Equation (8) shows that in  $C_c$ -limited STEMs, the probe size for a given probe current  $I_p$  and given primary energy  $E_o$  depends on just three parameters: the coher-

ent current  $I_c$ , the chromatic aberration coefficient  $C_c$ , and the energy spread  $\delta E$ . This means that for most present-day aberration-corrected STEMs, two out of the three parameters that determine the attainable probe size depend on the characteristics of the electron source. Despite this, there has been much less emphasis on improving the electron source than on aberration correction, and the source performance has not progressed greatly since the first consistent use of the cold field emission source by Crewe et al. (1968) over 40 years ago. This may change in the future, and it is therefore useful to review the present status of high-brightness electron sources.

## 2.2 Cold field emission and Schottky sources

High-brightness electron sources used by present-day STEMs come in two principal varieties: cold field emission guns (CFEGs), and field-assisted thermionic emission guns, i.e. Schottky guns.

Fig. 2 shows a schematic comparison of the emission mechanism from a cold field emission gun (Crewe et al. 1968b), and the Schottky gun (Swanson and Schwind 2009). In the CFEG, the electric field at the surface of the emission tip is typically about 10x higher than in the Schottky gun. This decreases the width  $w$  of the potential barrier due to the metal's work function  $\Phi$  so much that electrons at the Fermi energy level  $E_f$  are able to tunnel through the barrier. Electrons of slightly lower energy than  $E_f$  are able to tunnel through the barrier too, but the tunneling distance  $w'$  for them is longer, which significantly reduces the tunneling probability. The tunneling of these electrons causes a "low energy tail" of the emission peak (see figure). With an emission tip at room temperature, some electrons have energies slightly greater than  $E_f$  and these are able to tunnel out too. This causes a high energy tail, whose extent is typically much smaller in the CFEG than the low energy tail, giving a characteristically asymmetric zero loss peak in energy spectra.



**Fig. 2.** Comparison of the cold field electron emission mechanism with the field-assisted thermionic (Schottky) emission mechanism.

In the Schottky gun, the electric field at the tip is normally too weak for the electrons to be able to tunnel out. They therefore go over the top of the barrier. This is made possible by the extra energy supplied to them by the elevated temperature of the tip. A picturesque way of viewing this is that the Fermi sea is rough, and some of its "waves" are able to "splash" over the workfunction barrier. The energy of the

Schottky emission peak is therefore displaced by about  $\Phi$  relative to the CFEG peak. There is a finite probability that an electron will be excited to an energy some distance above the workfunction barrier, and this causes a high-energy tail of the Schottky emission peak.

Lowering the work function  $\Phi$  lowers the energy spread of the energy distribution for both types of guns, but the mechanisms are different. In the CFEG, lower value of  $\Phi$  means that the same tunneling width  $w$  is reached at a lower applied electric field. The shallower slope of the potential line means that the width of the tunneling barrier increases faster for energies lower than  $E_F$  (see Fig. 2), and this causes a faster decay of the tunneling tail. In the Schottky gun, the temperature required to excite electrons over a lower workfunction barrier is lower, and this decreases the width of the tail of the energy distribution extending above the barrier. The workfunction is typically lowered in the Schottky gun case, by building up a layer of ZrO or a similar oxide covering a (100) W tip (Swanson and Schwind 2009). Attempts to do the same for the CFEG (e.g. Batson 1987) have so far not resulted in a practical electron source.

The maximum attainable brightness of both the CFEG and the Schottky gun is chiefly determined by electron-electron Coulomb interactions, which modify the electron trajectories and thereby increase the size of the virtual electron source the electrons appear to come from. Without the interactions, one could simply increase the extraction voltage (in the CFEG case) or the tip heating (in the Schottky case) and thereby extract more electrons from the tip and thus operate with increased brightness. In reality, the broadening of the source due to the interactions causes an increase in the virtual source size and hence a drop-off in the brightness (current per unit solid angle per unit area) at total emission currents greater than about  $5 \mu\text{A}$  in the W (310) CFEG (Bacon 2008), and at larger total currents for blunter Schottky sources (which typically emit in the forward direction from a larger total area than CFEG sources). The Coulomb interaction is much reduced once the electrons have been accelerated, and the faster acceleration in the stronger extraction field gives the CFEG an edge over Schottky in the ultimate brightness it can attain.

For the cold field emission source, there is another important aspect beyond the brightness and energy spread: how stable is the emitted current? With the W (310) and (111) tips used in normal CFEG guns, contamination of the tip typically causes an increase in the work function and thus a drop in the emitted current. Further, the adsorbed contaminants tend to be mobile, and their rearrangement causes short-term changes in the tip geometry and work function. The current emitted from a “dirty” tip therefore fluctuates more than the current emitted from a “clean” tip.

When a tungsten cold field emission tip becomes dirty, it is readily restored to the clean condition by brief heating (flashing) at a high temperature, typically by passing a current through it for about a second. The clean tip then once again becomes gradually covered by adsorbates, and the emission decays. The time over which the emission current decays to one half of the starting value is usually denoted by  $t_{1/2}$ . In actual electron guns,  $t_{1/2}$  varies widely, from less than a minute to several hours or even several days (Martin et al. 1960).

When  $t_{1/2}$  is greater than about 30 minutes, the usual practice is to use the gun in its “clean” state, and to flash the tip as often as is required to keep it clean. This way

of operating a CFEG appears to have been introduced by Vacuum Generators (VG). It is also used with the Nion CFEG (Bacon et al. 2010). We call this type of gun clean cold field emission gun (CCFEG). Its advantages are increased brightness and reduced energy spread relative to a gun with a higher workfunction, and greatly reduced emission noise. Being able to operate a CFEG in this way requires that the vacuum near the tip be  $<10^{-10}$  torr even when emitting, and ideally in the high  $10^{-12}$  torr to the low  $10^{-11}$  torr regimes.

In CFEG guns with vacuum level worse than about  $1 \times 10^{-10}$  torr, the emission current typically decays much too fast for the gun to be usable in the freshly flashed “clean” state. The gun is then operated with the emission occurring with the adsorbate layers present, rather than from a clean tungsten tip. The extraction voltage is raised as needed to get a usable emission current, sometimes repeatedly to compensate for the current drop-off due to additional adsorbates. We call this type a “dirty” field emission gun (DFEG). Its advantage is that vacuum requirements are reduced. Its disadvantages are that the emission current fluctuates much more than with clean tips, and that the gun brightness is typically reduced compared to the “clean” state.

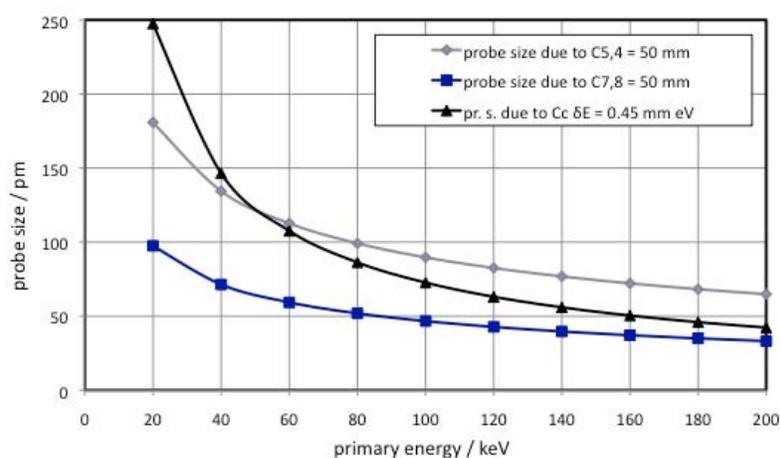
The Schottky gun can operate in a poorer vacuum of around  $10^{-9}$  torr, because arriving contaminants are continuously desorbed from the hot tip, which is in a state of dynamic equilibrium. Because of the high temperature, the mobility of the adsorbants is high and the changes in the work function are therefore much faster than they would be in a CFEG. Furthermore, the tip radius is typically larger in the Schottky gun than in a CFEG, and the emitting area that contributes to the final probe is therefore also larger. The high adsorbant mobility and the larger emitting area lead to a more stable emission current than in the DFEG case.

The electric field at the tip in the Schottky gun is weaker than the field that would be needed to obtain field emission by tunneling. If tunneling were present, electrons would go both *over the top* of the workfunction barrier as well as *through* the barrier, and the energy width of the emitted beam would grow to several eV (Swanson and Martin 1975, Swanson and Schwind 2009). This needs to be avoided in a practical source of electrons, and a practical Schottky gun is therefore never run at an applied field high enough to allow electrons of the Fermi energy to tunnel out of the tip. The standard Schottky gun therefore behaves very differently from a CFEG: the Schottky emission current goes to zero when the heating is turned off. The advantages of the Schottky source are that it can operate in an even poorer vacuum than DFEG, and that its emission current remains stable over long periods of time. Its disadvantages are a wider energy spread than achievable with CFEG, typically of the order of 0.5 to 1.0 eV, lower brightness, and shorter longevity: the Schottky emitter typically runs out of the pool of Zr needed to replenish its ZrO coating after about a year of continuous operation, whereas a well run clean CFEG emitter can typically last more than 3 years.

The Schottky gun is nowadays often referred to as a field emission gun (FEG), almost certainly following a practice coined by a marketing department. It is ironic and regrettable that many scientists have started to employ this terminology, i.e. to use an acronym based on an emission mechanism that the Schottky gun *must avoid in order to operate as an optimized electron source*.

### 2.3 Practical probe-forming performance

Fig. 3 shows the influence of chromatic aberration of  $C_c = 1.3$  mm combined with an energy spread  $\delta E = 0.35$  eV, and of geometrical aberrations of  $C_{5,4} = 50$  mm and  $C_{7,8} = 50$  mm, on the theoretical probe size. The assumed probe current is 32 pA, which is equal to 0.25  $I_c$  for a CFEG. The theoretical curves were computed using equation (8) above and expressions given in Krivanek et al. (2008).  $C_{5,4} = 50$  mm is typical of Nion-corrected VG columns (Dellby et al. 2001),  $C_{7,8} = 50$  mm is typical of a well-tuned Nion UltraSTEM (Krivanek et al. 2008). The chromatic aberration and the energy spread of both these instruments correspond to the ones used in the graph.



**Fig. 3.** Theoretical probe sizes for the indicated parameters and probe current of 0.25  $I_c$  in an aberration-corrected CFEG STEM.

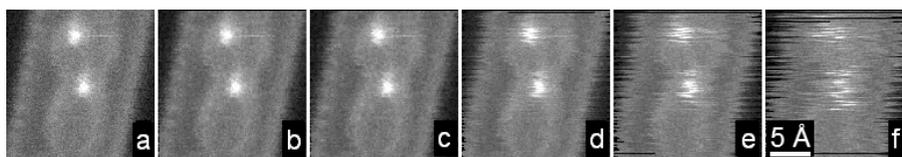
The probe size is plotted as a function of the primary energy. In the  $C_s$ -corrected VG columns, the uncorrected 5<sup>th</sup> order aberrations limit the probe size more than the chromatic aberration above about 50 keV. In the  $C_s$  and  $C_s$ -corrected UltraSTEM column, the chromatic aberration dominates at all primary energies up to 200 keV (and would dominate up to 500 keV if a microscope with similar aberration performance was operated at those energies). The strong dependence of the  $C_c$ -limited probe size on the primary energy means that it is about 5x worse at 20 keV than at 200 keV. This underscores the need for developing practical  $C_c$  correctors as the next step in aberration-corrected STEM, particularly for STEMs designed to operate at low primary energies.

The probe sizes predicted by Fig. 3 are largely borne out in practice. For instance, we have reached 1.09 Å probe size at 60 keV at 30 pA (Krivanek et al. 2010) and 0.8 Å at 100 keV. With our new 200 kV CFEG (Bacon et al. 2010), we should be able to reach a probe size of 0.5 Å. We have now reached 0.6 Å, and should be able to progress to 0.5 Å with better tuning.

In experimental practice, there are of course many additional components that need to be optimized in addition to the aberrations, the gun brightness and the energy spread. These include:

- i) the high voltage, which must be stable enough so that the energy spread does not increase, thereby worsening the chromatic aberration,
- ii) the tuning of aberrations, which must be accurate enough so that mistuned aberrations (including defocus and astigmatism) do not worsen the resolution,
- iii) the power supplies for the optics, which must be stable enough so that the tuning remains unchanged over time periods long enough to find the areas of interest on the sample and record the experimental data,
- iv) the mechanical properties of the microscope column, which must be rigid and stable enough so that its elements do not drift with respect to each other, thereby causing a change of the aberrations,
- v) the sample stage, which must be free of vibrations and drift,
- vi) the shielding of the microscope column and of the electronics, which must be good enough to keep out external influences such as stray magnetic fields, cell phone transmissions and acoustic noise,
- vii) the water cooling of the lenses, which must not introduce vibrations or thermal drift,
- viii) the microscope room, which must be acoustically quiet, free of floor vibrations and stray magnetic fields, and have a stable temperature,
- ix) the post-sample detector-coupling optics, which must be able to bring the right signals to the right detectors,
- x) the detectors, which must be fast and sensitive enough to record scattering events with good detective quantum efficiency,
- xi) the vacuum of the microscope, which must be high and clean enough so that contamination and sample etching are avoided.

Not meeting these requirements results in the resolution becoming worse, the data becoming noisier, the atomic images becoming “squiggly”, or the sample being destroyed prematurely.



**Fig. 4.** Images of a small part of a nanotube filled with nanopods filled with single Er atoms, with various amounts of probe “wiggles” added artificially, recorded with the Nion UltraSTEM at 60 keV. a) no added wiggles, b) 0.05 Å r.m.s, c) 0.1 Å, d) 0.25 Å, e) 0.5 Å, f) 1.0 Å.

The stability requirements are rather high, but are now being attained. Fig. 4 illustrates this with a small portion of an experimental image shown in six versions: as recorded (a), and with intentionally added noise simulating probe “wiggles” of 0.05 Å r.m.s. (b) to 1.0 Å r.m.s. (f). Only the image with 0.05 Å added wiggles looks

substantially the same as the unprocessed image a), indicating that 0.05 Å r.m.s. was about the actual level of the microscope instabilities

Note also that the horizontal streak to the right of the top atom visible in images a-c, which is due to the atom's jump to a neighboring site for just one scan line (the nanotube image is not disturbed for this scan line, as it would be if the streak was due to an instrumental instability), becomes impossible to distinguish from streaks caused by the injected instabilities in d-f. Any one of the 11 stability requirements listed above not being fulfilled can introduce streaks that are similar to the injected instabilities, and it is a considerable achievement that overall stabilities better than 0.1 Å (10 pm) are now being reached.

## 2.4 HAADF and EELS image resolution

Experimental images are obtained by scanning the probe over the sample, and can be described as a convolution of an ideal image (which would be obtained with an infinitesimally small probe) with the actual probe. Obtaining good experimental resolution therefore requires a small probe *and* an electron-sample interaction that is sufficiently well localized.

Electrons that form high-angle annular dark field images originate from Rutherford scattering by the deep potential well surrounding the atomic nucleus. ADF images of single atoms that were free of the influence of aberrations, source size broadening and of resolutions limits due to noise would show the potential well rather than the electron orbitals, and would therefore be very small – typically less than 0.3 Å in diameter for heavier atoms (e.g. Batson 2006). Their size would be partly due to the finite dimension of the potential well and partly due to the thermal vibration of the atomic nucleus.

With probe sizes greater than 1 Å, the broadening effect due to the finite size of atoms can usually be neglected and the ADF resolution taken as equal to the probe size. With probe sizes smaller than 1 Å, however, this approximation may produce significant errors. Beck and Crewe (1975) showed already 35 years ago that a  $C_s$ -corrected 100 keV STEM using an illumination half-angle of 30 mr will produce ADF images of C atoms that are 0.9 Å wide even though the probe size will be 0.8 Å. The ADF resolution for sub-Å probe sizes therefore needs to be worked out either by summing the squares of the probe size and the atom size, or by a full calculation of the expected images.

Inelastic scattering at energy losses of tens of eV and higher originates from the interaction of the incident fast electrons with the sample's electrons (Egerton 1996), i.e. from an interaction that is much more spread out than Rutherford scattering. The resolution in EELS maps is determined by probe size plus the spatial spreading of inner shell loss scattering (and also statistical noise, discussed in Krivanek et al. 2008a). The spreading is called “delocalization”, and it is well described by a formula given by Egerton for the diameter  $d_{50}$  of the area that contains 50% of the scattering events (Egerton 2006):

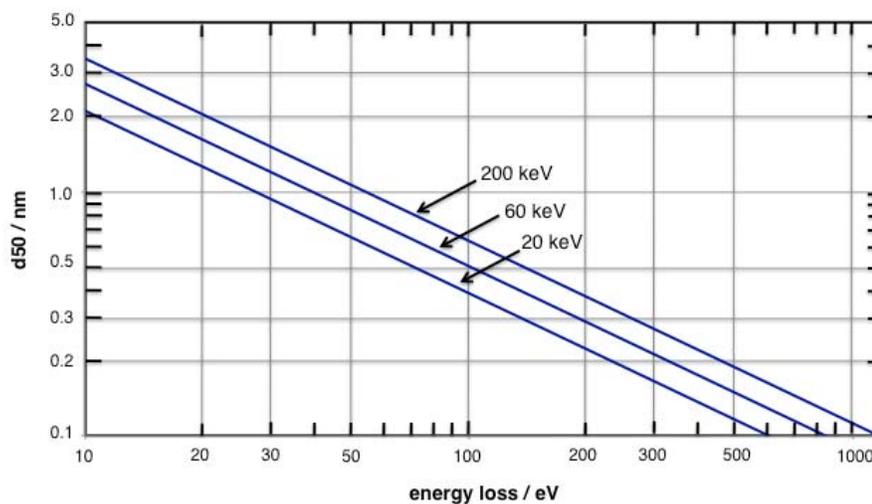
$$d_{50} = 0.5 * \lambda / (\Delta E / E_0)^{3/4}, \quad (9)$$

where  $\Delta E$  is the energy loss. The formula assumes that the EELS collection angle is large enough for the resolution not to be limited by the diffraction limit arising from a small range of scattering angles forming the inelastic image. This is often a problem in energy-filtered TEM imaging, in which the chromatic aberration of post-sample imaging optics blurs the image severely except when the admitted energy range and the range of scattering angles are both made narrow. In the STEM, the EELS collection optics is much less affected by aberrations, and the large collection angle condition is readily met by making either the illumination or the EELS collection angle large. With EELS-coupling optics capable of accepting an angular range larger than the probe cone without worsening the EELS energy resolution, both the angles are typically large: the illumination angle is made large in order to optimize the probe size and its current, and the collection angle is made large in order to optimize the collection efficiency.

The resolution expected in EELS maps is given by:

$$d_{\text{EELS}} = (d_{\text{probe}}^2 + d_{50}^2)^{0.5} \quad (10)$$

Fig. 5 shows  $d_{50}$  as a function of the energy loss for 3 primary energies: 20, 60 and 200 keV. Taking 2 Å as a convenient benchmark for “atomic” resolution, this level is achieved at energy losses greater than 460 eV at 200 keV, 340 eV at 60 keV, and 240 eV at 20 keV. Such an inverse dependence of the delocalization on the primary energy has not yet been verified experimentally, but it should be a significant additional advantage of operating at lowered primary energies.



**Fig. 5.** EELS delocalization according to Egerton’s (2006) formula as a function of the energy loss, for 3 different primary energies.

Another way to improve the spatial resolution of EELS maps may be to concentrate on large-angle scattering events, such as those that occur at energies considera-

bly above an edge threshold and give rise to the Bethe ridge (Egerton 1996) in the angular scattering distribution. This could be done for instance by using an annular EELS entrance aperture that prevents electrons scattered by low angles from entering the spectrometer. It also remains to be verified experimentally.

It is also worth noting that even though the formula for delocalization shows that atomic resolution EELS elemental mapping is not possible with very low energy losses of the order of a few tens of eV, the literature contains a number of experimental images that have been recorded at low losses and yet appear to show atomic resolution. This is readily explained by the fact that low loss, high spatial resolution EELS images can result from double scattering: low angle inelastic scattering that provides a new “primary” beam at the detected energy, plus high-angle elastic scattering that gives the fine detail in the image. The resultant images are not indicative of the sample composition. Substituting the atoms in the sample by other atoms of similar atomic weight would result in a very similar double-scattering image.

## 2.5 Image acquisition and processing

When imaging thin low  $Z$  materials at low primary energies, a slightly different strategy is called for than when imaging heavier and thicker samples at higher energies. First, the high-angle scattering from the low- $Z$  atoms not being particularly abundant, the image signal is increased about 2x when annular dark field images are formed using a lower cut-off semiangle of 50-60 mr rather than the 80-90 mr we normally use for heavier atoms at 60 keV primary energy. We call these images medium-angle annular dark field (MAADF). A slight increase in non-linearity is expected in MAADF images compared to HAADF ones, whereby 2 atoms lying on top of each other will give more than two times the signal of a single atom. We have tested for the effect by comparing the intensities of MAADF images of aligned single and double layers, which showed that it is much less than statistical noise typically present in the images.

Second, the dark field detector gain has to be increased (by increasing the PMT voltage) so that the signal from a single B or C atom is reliably detected above the background detector noise. We usually increase the gain to a level whereby the signal from about 10 graphite layers will saturate the detector, and then stay in thin sample areas where the saturation is not a concern.

Third, to get a good signal-to-noise ratio in the images of single light atoms, the exposure level (the electron dose per unit sample area) usually needs to be higher than would be needed for imaging thicker and heavier materials. Raising the probe current would worsen the resolution, and this leaves just one useful option: increasing the time the probe spends scanning over each atom. Instead of simply raising the per-pixel dwell time, we usually accomplish this by decreasing the pixel size. The resultant images are then greatly oversampled, with each atom occupying an area of 10x10 or even 20x20 pixels, but the oversampled images contain useful information on the exact atomic position and on the atomic movement, as will be shown below.

The resultant MAADF images typically show individual atoms well resolved even in graphene and monolayer boron nitride, in which the nearest neighbour spacings are 1.42 Å and 1.45 Å respectively. However, the image intensity does not go

to zero in the center of the hexagonal rings in these structures, and it decays over several Å away from the specimen edge. Both these effects are due to an extended “tail” of the electron probe, which is typically much stronger in actual imaging experiments than the tail of probe modeled by a Gaussian profile. The tail creates a “background fog” in MAADF images, and decreases their clarity. It also contributes extra intensity to the images of the nearest neighbours of each atom, and thereby makes the MAADF intensity of each individual atom depend on how many neighbours it had, and what their atomic numbers were.

It is important to note that the effects of the probe tail are well visible in MAADF images because of the quantitative nature of dark field imaging, in which the intensity in the vacuum next to the sample goes to zero, and thus provides a baseline to which the intensity in the center of the hexagons can be compared. In bright field (BF) phase contrast imaging, by comparison, the intensity in the vacuum is 1. A tail in a BF image of an atom results in a slight change in the image pattern and a reduction of the overall phase contrast. But there is no readily visible change in the DC level of any image area, and the contrast is easily boosted back up, rendering the reduction nearly invisible.

A second undesirable effect in the as-recorded highly oversampled MAADF images is that by spreading the available signal over many pixels, the signal per pixel is reduced, and the statistical noise increased artificially. However, the extra statistical noise is occurring at spatial frequencies much higher than the spatial frequencies of sample details captured in the image, and can therefore be readily filtered out. Provided that the noise introduced by the detector at every pixel is negligible, which is the case for well-designed MAADF detectors, the image with the high-frequency noise filtered out will then have no extra noise compared to an image acquired at a sampling frequency corresponding to the spatial frequency of the filtering.

Both the above effects can be corrected by a simple Fourier filtering procedure described by Krivanek et al. (2010a, 2010b). The procedure is illustrated in Fig. 6, which shows a part of an experimental 1kx1k MAADF image acquired at 60 keV with about 50 pA beam current, at 64 μs per each 0.12 Å wide pixel. The figure also shows the filtering steps and the end result.

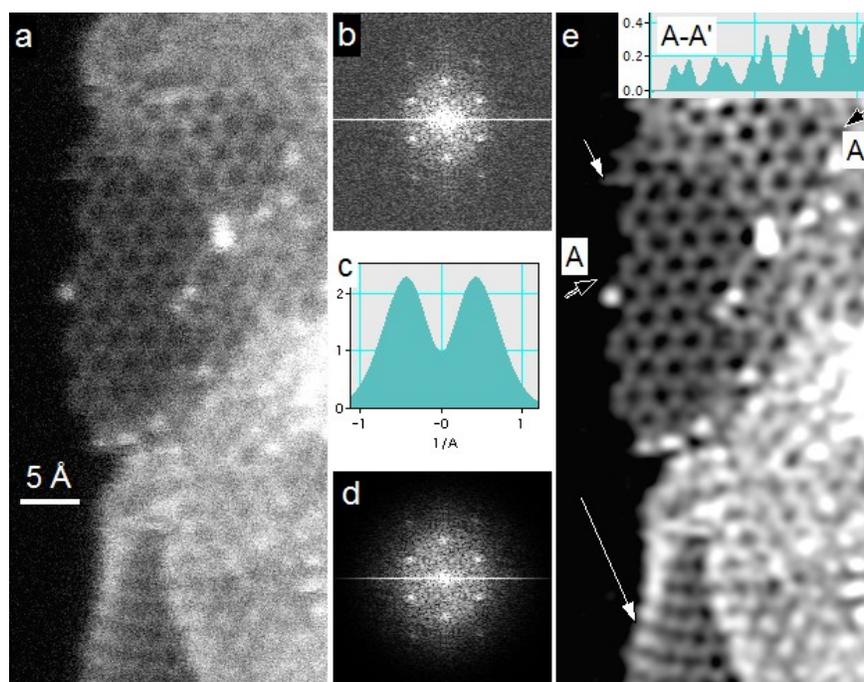
The filtering amounts to convoluting the image with:

a) a broad Gaussian, whose width corresponds to the experimental resolution, and which therefore filters out the artificial statistical noise occurring at spatial frequencies higher than the highest actual sample frequencies captured in the image, and

b) a negative Gaussian, whose width corresponds to the width of the probe tail, and whose intensity equals the intensity of the probe tail. The negative Gaussian causes the central dip of the filter, and amounts to subtracting the experimental probe tail contribution from the image, i.e. to “de-fogging” the image.

The shapes of the probe and of the probe tail are typically not known exactly, and they vary from image to image and especially from one autotuning operation to the next. This is the reason for choosing a particularly simple filtering procedure, in which the probe tail is greatly reduced compared to the unfiltered image, even though it is not subtracted exactly. Because the filtering is rotationally symmetric and has no sharp cut-offs that might cause “ringing” in the processed image, the probability of creating misleading artifacts out of random noise is small.

Fig. 6 also shows an inset profile (Fig. 6e), taken along the line A-A', which starts in vacuum, crosses a monolayer of graphene, and ends in a double-layer. The profile traverses the centers of the graphene hexagons, where it drops to about 10% of the single atom intensity. In unprocessed images, the intensity in the center of the hexagons was typically 50-70% of the single atom intensity, and this provided a reliable measure of the strength of the tail at 1.42 Å from the probe center. We avoided subtracting the probe tail completely, which would have produced negative intensities in the centers of some of the hexagons and also in some places along the sample edge.



**Fig. 6.** MAADF images of graphene illustrating a Fourier-filtering procedure designed to remove probe tails and artificially introduced statistical noise. a) as-recorded image, b) fast Fourier transform (FFT) of image, c) profile through the applied Fourier filter, d) resultant FFT of image, e) processed image obtained by an inverse FFT. Black arrows in (e) mark the direction of profile A-A' shown as an insert in the image. Sample courtesy Dr. V. Nicolosi, Oxford U. (Ultramicroscopy, by permission).

It is interesting to note that the second graphene layer was aligned over the first layer in A-A stacking in the sampled area, even though the normal stacking on graphene is A-B, in which atoms in the second layer lie over the centers of hexagons in the first layer. However, the second layer was probably pinned by amorphous carbon and hydrocarbons present around the edges of the monolayer, and was therefore not in an equilibrium configuration. Reassuringly, the intensity recorded in the dou-

ble layer for atoms aligned on top of each other is about 2x the intensity of single carbon atoms.

Two other interesting details in the image details are marked by white arrows. The short white arrow marks a location that probably had a single carbon atom dangling off the graphene edge, but which ran away while the probe was scanning over it. This can be seen in the unsmoothed image (a), in which there is an extra intensity off the graphene edge that is cut off abruptly, from one scan line to the next. The long white arrow marks a monolayer graphene sheet that curled over at the edge, thereby creating a shape resembling one quarter of a complete nanotube. Many other interesting details of sample structures are shown and discussed in the next section.

### 3 Graphene and monolayer BN imaged at 60 keV

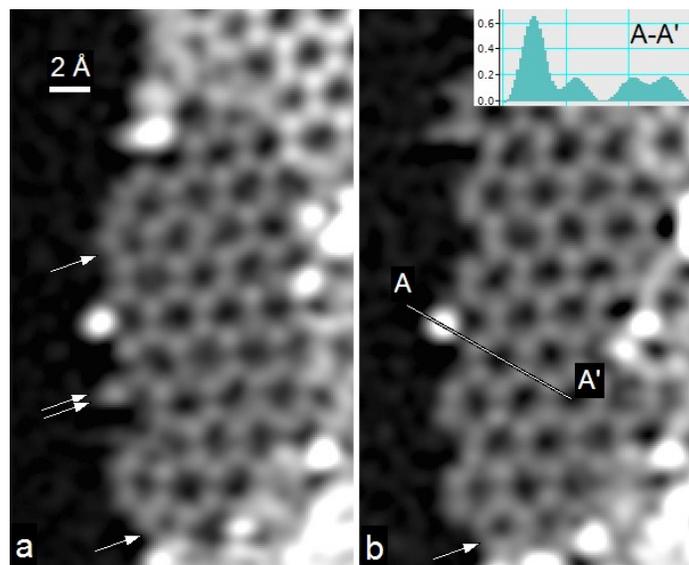
Graphene and BN samples were prepared by liquid phase exfoliation of bulk graphite and BN powders in N-Methyl-2-Pyrrolidone (NMP), which gave monolayer dispersions with a good yield (Hernandez et al. 2008). Full details of the sample preparation are given in Krivanek et al. (2010a).

The exfoliation produced graphite and BN flakes with small monolayer regions in various locations at or near the flake edges. The size of the regions varied. Smaller monolayer areas of around 10 by 10 (to 30 by 30) hexagons, surrounded by thicker regions, were typically the most stable under the beam and were therefore very suitable for observation.

#### 3.1 Graphene: lattice defects and adatoms at graphene edge

Fig. 7 shows the central portion of the image of Fig. 6 at higher magnification (Fig. 7(b)), and the same part of the sample imaged immediately before (Fig. 7(a)). Both the images were processed by the de-fogging filter and displayed slightly non-linearly, in order to make the monolayer graphene clearly visible without saturating the images of impurity adatoms. The impurity atom at the graphene edge stayed in its place, which was only about 50% probable, as could be seen by observing, with the same electron dose, the mobility of impurity edge atoms at other locations in the same sample.

The structure of the edge itself had undergone major modifications. In the left image, a variety of atomic arrangements is seen at the edge: two five-fold rings (indicated by single white arrows), a single dangling carbon atom (indicated by a double arrow), a distorted “armchair” (in which a complete carbon hexagon sits right at the sample’s edge) just above the bottom five-fold ring, and some atoms that were moving and left streaks behind. In the right image, the edge terminates in 4 regular armchairs. The rearrangement required the addition of just one carbon atom below the impurity atom and the removal of one carbon atom above the impurity atom. The armchair-terminated edge is similar to graphene edges imaged by bright field phase contrast TEM (Girit et al. 2009), but the observations of a 5-fold ring at graphene’s edge and of a single dangling carbon atom appear to be new.



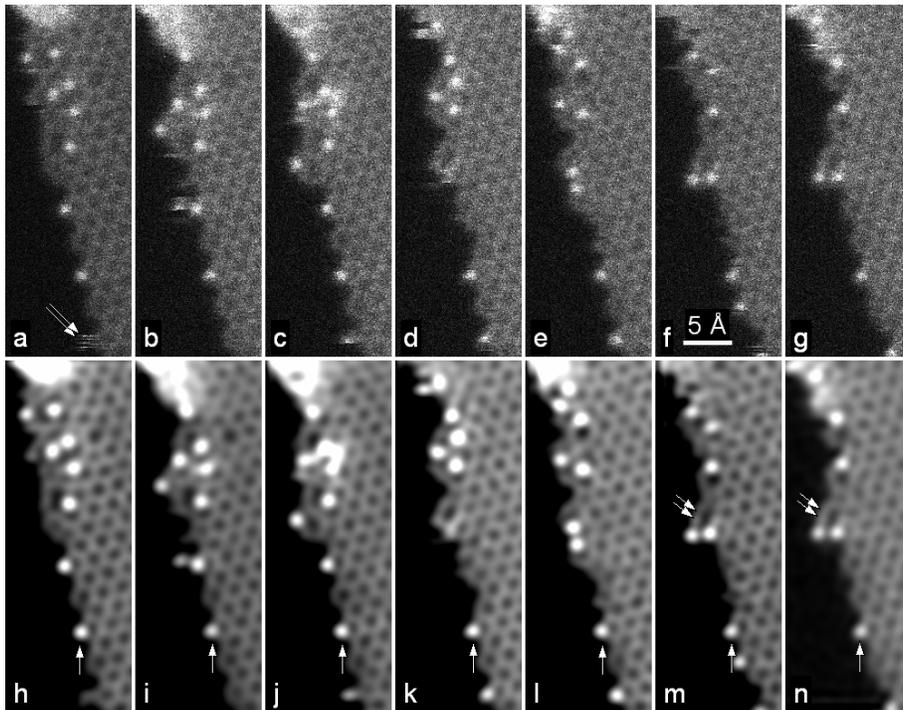
**Fig. 7.** MAADF images of monolayer graphene taken about 2 minutes apart. Image (b) is a higher magnification version of Fig. 6 (e). The single arrows in (a) and (b) point to 5-fold rings at the graphene's edge, the double arrow in (a) points to a single atom of carbon dangling off the graphene edge. A-A' profile through the impurity atom at the edge is shown as an insert in (b). Nion UltraSTEM, 60 keV, sample courtesy Dr. V. Nicolosi, Oxford U. (Ultramicroscopy, by permission).

Many carbon hexagons are seen to be somewhat distorted, and the distortion of the same hexagon is typically different in the two images. There were three principal causes for the distortions: a) statistical noise, which randomly enhanced different parts of the spread-out atomic images, and thus caused the smoothed images of individual atoms to shift randomly from frame to frame, b) sample movement, which translated into the displacement of some parts of the image but not others, and c) real distortions present in the carbon sheet, plus apparent distortions caused by the fact that the sheet was not aligned perpendicular to the beam and was probably also slightly buckled. The best way to separate the random distortions from the real ones is to image the same area in a sequence of images. The two images shown here indicate that most of the distortions in the present case were of the random kind. In stable samples the random distortions grow smaller at larger electron doses, and our practical experience (Krivanek et al. 2010a) indicates that they can be kept as small as about  $0.1 \text{ \AA}$  if the dose is increased about 4x relative to the one used here.

There were several impurity adatoms, which gave much stronger contrast than the carbon atoms. Adatoms on the right side of the images were located on top of the graphene sheet and were moving frequently, and this made their analysis difficult. The single adatom at the graphene's edge was stationary, and formed the apex of a 5-fold ring, with larger separation from its neighbours than the apex atom in the carbon-only 5-fold ring seen just above the adatom in Fig. 7(a). A profile through the adatom (insert in (b)) shows its intensity to be 3.6x larger than that of the C images.

Using the  $I = a Z^{1.64}$  dependence of the atomic intensity  $I$  on the atomic number  $Z$  that we have measured experimentally (Krivanek et al. 2010a) on images of B, C, N and O atoms obtained under essentially the same conditions as here, gave  $Z_{\text{impurity}} = 6 \times 3.6^{1/1.64} = 13.1$ , and we therefore tentatively identified the atom as aluminum. However, the extrapolation to  $Z = 13$  based on experimental data obtained for  $Z = 5$  to 8 is a stretch, and it is therefore possible that the impurity atom was Mg or Si, or even Na or P. EELS was tried on the atom and similar intensity impurity atoms in the vicinity, but it was not conclusive: the atoms were not strongly attached and tended to run away under the beam.

Fig. 8 shows a time sequence of MAADF images of the edge of a graphene monolayer that was decorated by several adatoms, recorded as a sequence of images each one of which took 8 seconds to record. The adatom intensities were similar to the adatom whose profile is shown in Fig. 7 (b), and they were therefore probably also Al. Once more, trying to identify the adatoms by EELS resulted in them running away, without providing useful EELS data.

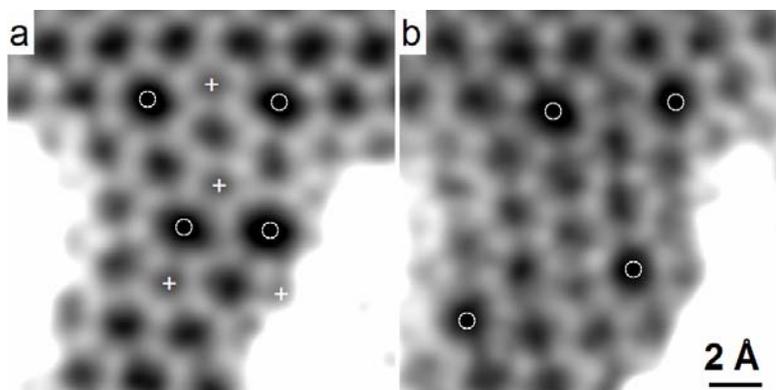


**Fig. 8.** Time sequence of MAADF images of a graphene edge decorated by several adatoms, most of which were rather mobile. Top row: unprocessed images; bottom row: smoothed and tail-removed images. Nion UltraSTEM, 60 keV, sample courtesy Dr. V. Nicolosi, Oxford U.

Arrows mark various interesting features in Fig. 8. The double arrow in (a) marks an adatom that came and went while the beam was scanning over the area,

resulting in short streaks. The single arrows in (h)-(n) mark an adatom that remained stationary throughout the sequence. The double arrows in (m) and (n) mark a single chain of C atoms, about 3 Å long, terminating in a single adatom. The bottom half of the portion of the graphene edge shown in the image was relatively stable, with the armchair termination dominant. The top half was much more mobile, and had 5-fold and 7-fold rings of carbon that came and went. The whole sequence illustrates the detailed nature of the studies the dynamics of low Z materials that have now become possible.

Fig. 9 shows a pair of MAADF images of monolayer graphene recorded about 1 minute apart, some distance away from the sample edge. Both show four 7-fold carbon rings (marked by white circles) and 5-fold rings (marked in by white crosses in (a)). The atomic arrangement for the top two 7-fold rings in (a) is close to a Stone-Wales defect (Saito et al. 1998), but with a 6-fold ring in place of one the 5-fold Stone-Wales ones. In (b), two of the 7-fold rings have moved to different places, and the atomic arrangement has grown more complicated. Overlayers at bottom left and bottom right are only an additional layer thick, but they appear as saturated white in the present images, whose contrast has been adjusted to show the monolayer clearly.



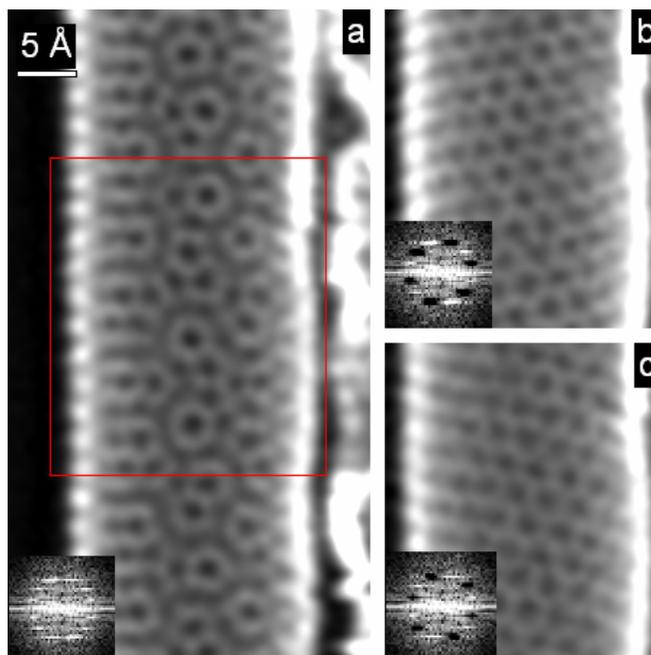
**Fig. 9.** Defects in monolayer graphene monolayer imaged 1 minute apart. MAADF, 60 keV. White circles mark 7-fold carbon rings, white crosses mark 5-fold rings. Nion UltraSTEM, sample courtesy Dr. V. Nicolosi, Oxford U.

### 3.2 Single wall nanotubes imaged with atomic resolution

Nanotubes are essentially a graphene sheet rolled up into a tube. The orientation of the rolled up sheet determines the nanotube chirality and “helical pitch”, which in turn determines the conducting properties of the nanotube (Saito et al., 1998). Fig. 10 (a) shows an MAADF image of a single wall nanotube obtained at 60 keV and processed using the noise and tail-removing deconvolution procedure.

The nanotube displays an interesting periodic structure (with a longitudinal periodicity of 31 Å), but the image is not clearly indicative of the nanotube structure. However, a Fourier transform of the nanotube gives two sets of mirror-related reflec-

tions (insert in (a)). Masking one set followed by an inverse FFT produces the image (b), which is simply either the front or the back half of the nanotube. Masking the other set produces an image of the complementary half of the nanotube. Determining the pitch of the nanotube helix becomes very easy with the two halves of the nanotube separated in this way. Determining the nanotube's chirality should also be possible, for instance by tilting the illuminating beam by 10 mrad or more and observing the resultant shift between the top and bottom halves. It is also interesting to note that the nanotube is slightly deformed, with a shape that conforms to the shape of an irregular nanotube pressing against it from the right side, but remaining about 3.6 Å away.

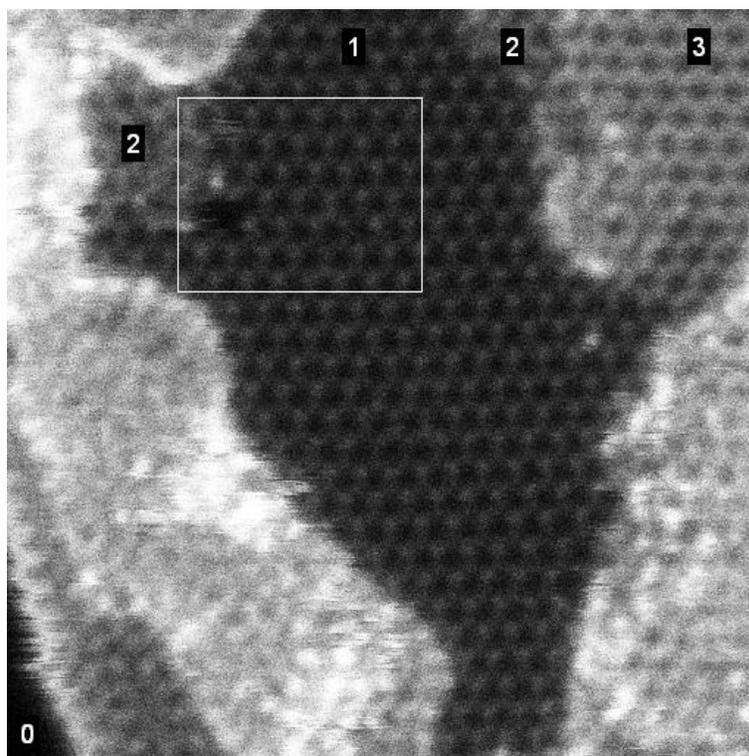


**Fig. 10.** a) MAADF image of a single wall carbon nanotube obtained at 60 keV, with the diffractogram shown in an insert. b) one half of the nanotube from the area marked by the red rectangle in (a), obtained by Fourier filtering that masked one half of the nanotube reflections. c) the other half of the nanotube. Nion UltraSTEM, sample courtesy Dr. David Geohegan, ORNL.

The front-back separation for a nanotube has been done before using a bright field image (Suenaga et al. 2007). The work of Suenaga et al. and our work, which was done at a lower operating energy and with slightly better resolution, show that nanotubes can now be imaged atom-by-atom, and defects and impurities present in them identified clearly. They are also very suitable as containers for holding molecules of unknown structures, since their contribution to the observed image of the molecule can be subtracted away quite precisely.

### 3.3 Monolayer BN: distinguishing B from N and identifying impurities

Fig. 11 shows an unprocessed 1kx1k MAADF image of a BN monolayer, surrounded by double, triple and thicker layers of BN as well as overlayers that included hydrocarbons which were revealed by a strong C K-edge in EEL spectra of the thicker regions.

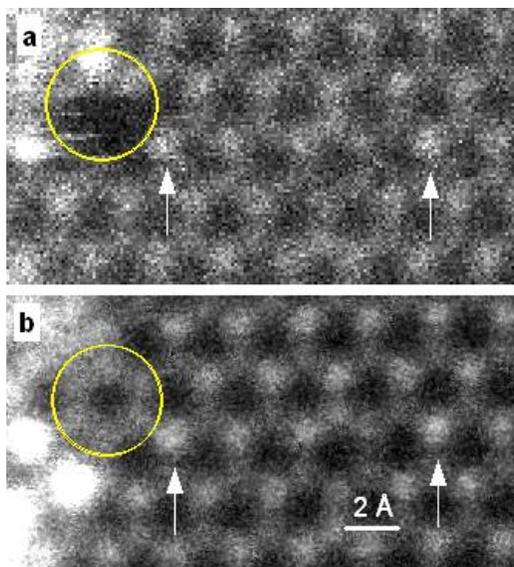


**Fig. 11.** 123 x 123 Å area of BN. Different numbers in the image mark the number of BN layers in that image area. “0” indicates the vacuum beyond the sample edge. The per-pixel dwell time was 64 μs, the pixel size 0.12 Å. The white rectangle shows an image area studied in greater detail. Nion UltraSTEM, 60 keV, sample courtesy Dr. V. Nicolosi, Oxford U. (Nature supplementary materials, by permission).

Fig. 12 compares an unprocessed image of the sample region containing the hole to an unprocessed image of the same region taken about 2 minutes later. The magnification was twice as high for the second image, in which each pixel was only 5.9 pm (0.059 Å) wide. The pixel dwell time was the same for both the images. This meant that there were 4x as many electrons per Å<sup>2</sup> in the higher magnification image, and that its statistical image noise was therefore greatly reduced.

A triangular pattern of brighter spots is clearly visible in both the images, with a darker spot in the center of each bright spot triangle. This is exactly what is expected

in ADF images of monolayer BN. The brighter spots correspond to the heavier nitrogen, and the darker ones to the lighter boron.



**Fig. 12.** Comparison of the BN area containing the hole (a) with the same area imaged about 2 minutes later, at a higher magnification. (Nature supplementary materials, by permission).

The images of Fig. 12 show several departures from the regular bright spot - dark spot pattern. Two spots that are brighter than the spots corresponding to nitrogen are indicated by white arrows. They are only just brighter than the nitrogen spots, and their location did not change from one image to the next. This indicates that they are due to heavier substitutional atoms, probably oxygen, incorporated into the BN lattice. They provide fiducials that allow individual atoms in the BN lattice to be followed from one image to the next in the sequence of several images we recorded from this area.

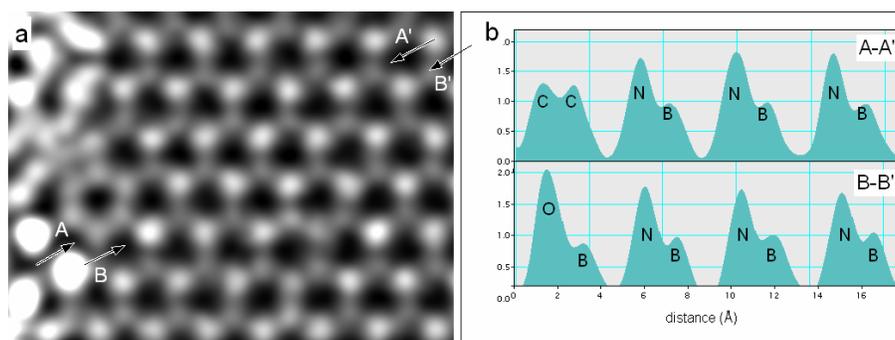
Several image spots, located on the left side of the images are considerably brighter than the spots due to substitutional atoms. These spots mostly occur in different locations in the two images. They are almost certainly due to mobile impurity adatoms on the BN surface.

The hole seen in Fig. 12 a) is marked by a small circle. There were atoms moving around in the hole, and the motion produced the short white streaks visible inside the hole. The streak lying at about 7 o'clock within the circle is two scan lines wide, meaning that the atom arrived at this location while the beam was scanning nearby, and left 66-198 ms later, 66 ms being the line scan interval (i.e., it stayed for one whole scan line interval, plus two unknown portions of scan line intervals). The streak at 9 o'clock is only one scan line wide, meaning that the atom departed 1-132 ms after its arrival.

The same circle is also shown in Fig. 12 b). The hole had filled up, but the atoms within it deviate from the bright-dark pattern of spots in the rest of the image: their intensity is roughly the same. It seems likely that carbon atoms available in the hydrocarbon deposits next to the hole on the left side filled the hole up, and that the brighter spot marked by the left white arrow was an oxygen atom that lodged itself in the BN as a part of the same hole-filling process.

Another small hole, roughly where the right oxygen atom is, was seen briefly in an earlier image and then filled up. This suggests that all the substitutional impurities seen in this image were incorporated in the BN sheet following hole creation by the electron beam, and their subsequent filling by mobile adatoms traveling over the BN sheet.

Fig. 13 shows the image of Fig. 12 (b) after smoothing and de-fogging using the double Gaussian filter, plus a small correction (of about 0.4 Å amplitude) of scan distortions present in the as-acquired image. The strength of the negative Gaussian component was adjusted so that the intensity at the center of the BN hexagons (which was 50% in the unprocessed image), became close to zero. This guaranteed that the intensity contribution of individual atoms to their near neighbour sites, which are the same distance away from the atoms as the centers of the hexagons, was also reduced to zero. In other words, the spurious contributions that the tails of the images of the nearest neighbours would have made to each atomic image have been nulled by the procedure.



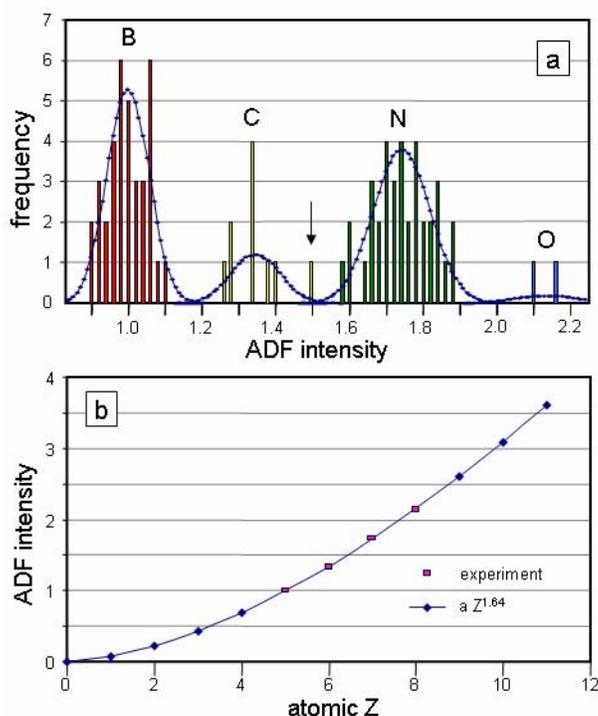
**Fig. 13.** a) filtered version of the MAADF image of the BN monolayer area containing atomic substitutions. b) profiles through marked locations in (a). (Nature, by permission)

Profiles A-A' and B-B' shown in Fig. 13 (b) therefore portray the correct intensities, rather than intensities altered by a probe tail. They show a clear and consistent pattern of peaks of alternating intensity, with the higher peaks corresponding to nitrogen atoms and the lower ones to boron. There are 3 significant deviations from the pattern: two peaks whose intensity is about half-way between the N and B peaks in profile A-A', and a single peak in profile B-B', whose intensity is significantly higher than the N peaks.

The most plausible explanation is the one already given: the intermediate peaks are due to C atoms, and the high one due to an O atom. Without a quantitative statis-

tical analysis, however, atomic assignments such as these are subject to an unquantifiable statistical uncertainty. The appropriate way to quantify the assignments is to compute a histogram showing the distribution of the atom intensities (Isaacson et al. 1979, Voyles et al. 2002) for all the atoms in a given area, and to use the histogram to determine the probability that the atomic assignments were made correctly.

Fig. 14 (a) shows a histogram of all the image peaks within the monolayer area of the corrected image of Fig. 13 (a). The histogram separates into four distinct peaks, showing that we selected the illumination dose ( $6 \times 10^6$  electrons per  $\text{\AA}^2$ ) just right: 2x fewer electrons would have resulted in enough additional statistical noise to cause the peaks to overlap, 2x more would have produced a better peak separation, but may have caused extra damage to the sample. The B and N distributions are modeled by Gaussians whose widths were derived from the data scatter; the C and O peak were modeled by Gaussians whose widths were extrapolated from the B and N distribution widths. The intensity has been normalized so that the center of the boron peak is at 1.0.



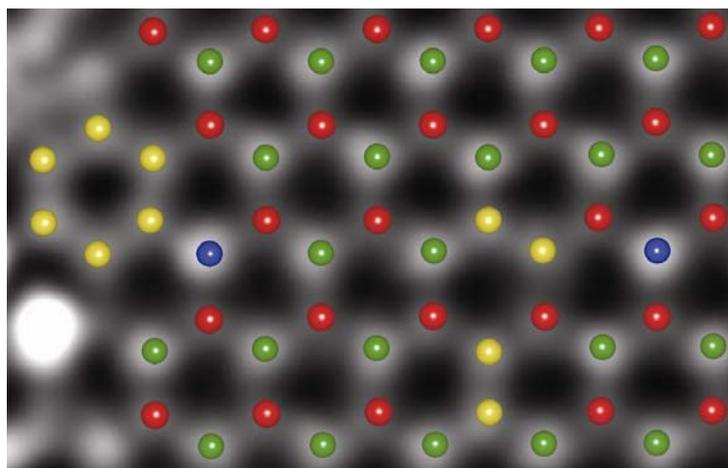
**Fig. 14.** a) histogram of the peak intensities in the monolayer area of Fig. 13 (a), b) plot of the histogram peak positions as a function of atomic number Z, together with the best fit of  $I = a Z^{1.64}$ . The uncertainty of the experimental points in (b) is indicated by the height of the small rectangles. (Nature, by permission)

Fig. 14 (b) shows the dependence of the average atomic peak intensity, i.e. the centers of the histogram peaks, on the assigned atomic number Z, plus a theoretical fit using an  $I = a Z^{1.64}$  model. The fit is excellent, passing through the experimental

data rectangles, whose heights correspond to  $2\sigma$  for the experimental points. An exponent of 1.64 is about what is expected for MAADF imaging on theoretical grounds (Hartel et al. 1996).

There was one exception to the clear separation: the arrowed bar in the valley between the C and the N peaks. The intensity of the corresponding atom's peak was 3 standard deviations from the center of the C peak, and 5.6 standard deviations from the center of the N peak. This means that the atom was likely to be carbon at 94% confidence level. For all the other atoms, the probability of having made the correct assignment was >99%.

Fig. 15 shows the distribution of the substitutional impurities in the marked area, superimposed on the experimental image. The oxygen atoms substituted for nitrogen atoms, singly, whereas the carbon atoms substituted for boron and nitrogen in pairs. The paired substitution avoided an energy penalty due to the unbalanced charge distribution associated with a single substitutional carbon atom in BN, and is therefore not surprising.



**Fig. 15.** MAADF image of BN shown in Fig. 13(a), with a DFT-relaxed atomic model corresponding to the atomic types derived from the observed intensities shown on top of the image. Boron = red, carbon = yellow, nitrogen = green, oxygen = blue. DFT model courtesy T.J. Pennycook. (Nature, by permission)

The substitutional atoms created small in-plane distortions in the BN lattice next to them. In particular, the O atoms pushed their nearest neighbours away by about 0.1 Å. This is most readily seen for the O atom next to the C hexagon: the C atom nearest to the oxygen is pushed into the carbon ring.

The stability of the substitutions was verified by density-functional theory (DFT) calculations (Krivanek et al. 2010a). The calculations also confirmed the lattice distortions caused by the substitutional atoms, although the amplitude of the distortions predicted by DFT was about 50% smaller than the distortions observed experimentally.

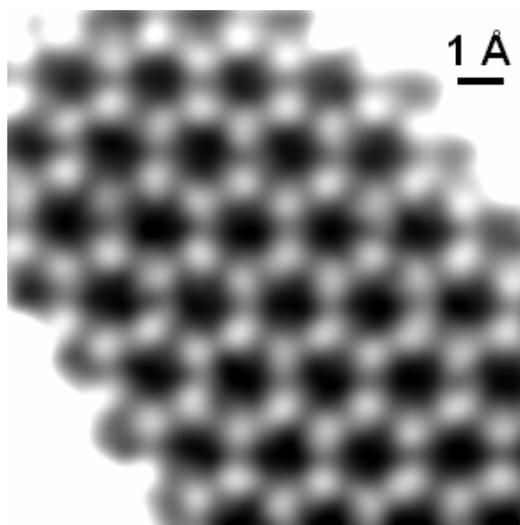
Going beyond the single layer BN, the image area on the lower left side of Fig. 13 contains 3 bright spots, whose intensity is a good match for sodium atoms sitting over N atoms in the BN layer. The image area to the left and above the carbon hexagon in Fig. 13 (a) shows a disordered second layer lying mostly over a continuation of the BN layer. It provides tantalizing glimpses of a disordered 3D structure, some of whose atoms we are able to place. However, we have not been able to model the entire structure. This kind of investigations may well become more fruitful when the low-energy STEM resolution improves further, as discussed in section 4.

It is useful to note that there have been several previous attempts to distinguish boron atoms from nitrogen atoms in monolayer BN using bright field phase contrast imaging (Meyer et al. 2009, Jin et al. 2009, Alem et al., 2009), but that none of them has succeeded in being able to identify a particular atom in a single image as being either B or N. The reason is that the bright field scattering strength of the two types of atoms is very similar (Meyer et al., 2009), which makes it difficult to distinguish them without extensive averaging to improve the signal-to-noise ratio, either over several images, or many atomic sites, or both.

The “traditional” wisdom in electron microscopy used to be that annular dark field imaging in the STEM was excellent for imaging heavier atoms, but that it was not a suitable technique for imaging single light atoms because of their small cross sections. Imaging light atoms by ADF STEM is indeed more difficult than imaging heavy atoms. But the results shown here demonstrate that aberration correction has made incoherent ADF imaging of light atoms readily possible, and that this technique enjoys the standard advantages of incoherent dark field techniques: better resolution than BF imaging in the same instrument, quantitative results, and simple interpretation.

An instructive example illustrating the quantitative nature of incoherent ADF imaging is shown in Fig. 16, which looks very much like an image of monolayer graphene: all the atomic maxima have about the same intensity. It was one of the first atomic resolution MAADF images we recorded from BN at 60 kV, and initially we were mostly producing images just like this one. This was rather confusing: we expected images in which the boron and the nitrogen differ substantially, i.e images similar to those shown in Figs. 11-13. The explanation was not long in coming: we were looking at a double BN layer. The initial sample area we examined in fact had many more double than single layer areas, a property that we have seen repeatedly in BN samples, and all our images were from double and thicker areas. The BN stacking is A-A' – boron atoms in the second layer lie over nitrogen atoms in the first layer and vice versa, and graphene-like contrast is therefore expected for the double layer.

The simple measurement which demonstrated that this was happening was a line profile spanning from the area shown in Fig. 16 to a thicker area recorded in the same micrograph. Instead of the 1:2 ratio of average intensities that we expected, we obtained a ratio of 2:3. This showed very clearly that the thinner area was in fact thicker than a monolayer. Had we been looking at our sample using bright-field phase-contrast imaging, no such quantitative tool allowing us to determine how many layers we were looking at would have been available.



**Fig. 16.** A smoothed, de-tailed and de-warped MAADF image of double-layer BN. 60 keV, 64  $\mu\text{s}$  per each 0.05  $\text{\AA}$  wide pixel. The image looks very similar to an image of monolayer graphene, except in one crucial aspect: its average intensity is double that of monolayer graphene. Nion UltraSTEM, sample courtesy Dr. V. Nicolosi, Oxford U.

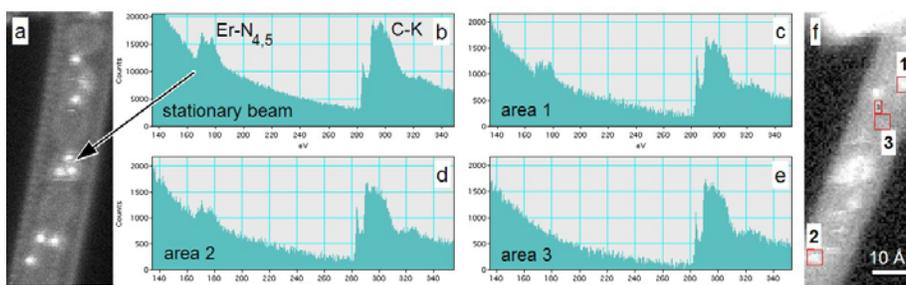
### 3.3 EELS of single heavy atoms

Our attempts to record EEL spectra from light impurity adatoms ( $Z \sim 13$ ) resting on graphene and monolayer BN have been largely unsuccessful so far. The chief reason is that the EELS cross-sections are typically 100x-1000x weaker than the MAADF ones, and obtaining an EEL spectrum with a good signal-to-noise ratio therefore requires that the electron probe spends much longer over each atom. Adatoms are not strongly bound to graphene and BN, and nearly always run away while an EEL spectrum or a spectrum-image is being recorded.

The situation is much more favorable for EELS if the impurity atom is confined, for instance when it's inside a nanopod, which can itself be inside a nanotube. Such samples are now being produced, sometimes with small molecules being confined in this way instead of single atoms (e.g., Suenaga et al. 2000, Koshino et al. 2007, Liu et al. 2007).

Fig. 17 shows an MAADF image of Er atoms inside  $\text{C}_{82}$  nanopods stuffed inside a single wall nanotube, plus Er EEL spectra and an Er EELS map extracted from a spectrum-image (Krivanek et al. 2010b). The single Er atoms are very readily visible in the MAADF image, even though the beam current was about the same (50 pA) while the per-pixel time was shorter (10  $\mu\text{s}$ ) and the pixel size larger (0.12  $\text{\AA}$ ) than the settings we typically use for optimised imaging of graphene and BN. The spectra show very good signal-to-noise ratios, more than adequate for identifying the single Er atoms.

The resolution predicted by equations (9) and (10) for a 1.4 Å probe and a 170 eV energy loss (the energy of the Er  $N_{4,5}$  edge threshold) at 60 keV primary energy is 3.6 Å. This is in good agreement with the data shown in Fig. 17, and it was confirmed more accurately experimentally (Krivanek et al. 2010b). It is much larger than the probe size. Even with the carbon K-edge energy of 285 eV, the EELS resolution at 60 keV is still a relatively poor 2.6 Å.



**Fig. 17.** a) An unprocessed MAADF image of a single-wall carbon nanotube filled with  $C_{82}$  nanopods, which originally contained one Er atom each. b) EEL spectrum recorded with the STEM probe placed between arrowed 3 Er atoms, at an acquisition time of 1 s. c)-e) EEL spectra extracted from the marked areas from in spectrum-image (f), f) post-Er  $N_{4,5}$  energy slice through a spectrum-image recorded with 9 ms per each  $0.5 \times 0.5$  Å pixel. c) and d) originate from single Er atoms, e) originates from the carbon nanotube only. Nion UltraSTEM, 60 keV, sample courtesy Dr. K. Suenaga, AIST. (Ultramicroscopy, by permission)

## 4 Current state-of-the-art and future directions

### 4.1 Present status

The observations shown here document several important points:

1) The signal-to-noise ratio (SNR) in ADF images of individual atoms as light as boron is now high enough to allow individual atoms to be identified by their ADF intensity. This technique is at its best when the atoms are non-overlapping, as in the BN example shown here, but it is also applicable to heavy atoms lying on lighter thin substrates.

To see how well the atomic identification by ADF intensity is likely to work for light and heavy atoms, we start by remembering that our measurements showed that the MAADF image intensity  $I$  increases as  $I = a Z^{1.64}$ , where  $a$  is a constant. The separation of the histogram peaks for adjacent elements thus increases as  $\Delta I/\Delta Z = 1.64 a Z^{0.64}$ . The width of the histogram peaks in an image whose SNR is limited by the finite statistics due to the limited electron dose increases as  $a^{0.5} Z^{0.82}$ , which means that the relative separation of the histogram peaks (the absolute separation

divided by the width of the peaks) for adjacent elements decreases as  $Z^{-0.18}$ . This is a rather weak dependence: the relative separation of Pt ( $Z=78$ ) and Au ( $Z=79$ ) histogram peaks is about  $(78/6)^{-0.18} = 63\%$  of the relative separation of C and N peaks. Provided that no influences other than the finite image statistics limit the precision with which individual atomic intensities can be measured, distinguishing isolated Pt atoms from Au ones with high confidence level should therefore be possible with an electron dose that is  $(1/0.63)^2 = 2.5\times$  times higher than the one used here (i.e., at about  $2\times 10^7$  electrons per  $\text{\AA}^2$ ), and distinguishing iridium ( $Z=77$ ) from Au should be possible with a slightly smaller dose than used here. Another requirement will be that atoms of a known and similarly high  $Z$  be imaged at the same time, so that  $Z$ -dependence data will not have to be extrapolated to very different  $Z$ s.

2) Extending the SNR considerations towards lighter elements shows that individual atoms of all elements down to H present in monolayer samples should now be identifiable by ADF imaging with a slightly smaller dose than the one used in this work, provided that they remain stationary while the electron beam is scanning over them. Unfortunately, hydrogen itself is unlikely to remain stationary at the high doses used in this kind of imaging, and may therefore continue to be hard to detect directly, even though the detection has become possible in principle.

3) The high SNR of MAADF imaging makes it possible to distinguish whether an individual atom was in place for each pixel in an atomic image spanning an area consisting of 100 or more pixels, for atoms as light as carbon. This is allowing atomic motions to be studied on a time scale corresponding to the per-pixel dwell time, in our case 10  $\mu\text{s}$  (for single Er atoms) and 64  $\mu\text{s}$  (for single carbon atoms).

4) Heavy atoms in nanotubes and  $Z \sim 13$  atoms on graphene were seen to be mobile even when the beam was not directly over them. The atoms tended to be more stationary when the electron dose was smaller, and this suggests that the beam had to be in the general vicinity in order for the atoms to move. Previous studies of atomic motion with Crewe's original 30 kV STEM (Isaacson et al. 1977, Crewe 1978) have suggested that much of the atomic motion is thermal in origin, but the higher primary energy used here (60 rather than 30 keV), may make the beam-induced effects much more important. In the future, especially when probe correctors of chromatic aberration have become widely available, 30 keV may become a widely used operating energy.

5) An energy loss spectrum from a single atom can now be collected with excellent SNR, provided that the atom has a large cross section and remains stationary during the acquisition. Nevertheless, there are several factors that make atomic identification by EELS more complementary than competitive with atomic identification using MAADF imaging:

- a) Much higher doses and correspondingly longer acquisition times need to be used for the EELS, making it likely that the atoms of interest will simply run away,

- b) The spatial resolution of the EELS elemental map is typically not as good as the probe size. For energy losses  $< 500$  eV, it will typically not be good enough to resolve the nearest neighbors in closely packed crystals.
- c) EELS edges suitable for elemental mapping, i.e. edges with energies between about 100 and 2000 eV, with sharp thresholds and sufficiently large cross-sections, are only available for about half the elements in the periodic table. This means that EELS mapping cannot become a general technique applicable to all atomic species. EELS therefore needs to be supplemented either by ADF imaging, or by other spectroscopic techniques such as energy-dispersive X-ray spectroscopy (EDXS).

6) Holes were made in BN away from its edges, even though the electron energy used here was considerably below the theoretical knock-on displacement threshold of 78 kV. Two explanations appear possible:

- a) an intermediary agent may be able to transfer more energy from an incident electron to a B or N atom than can be transferred in a direct electron – B or electron – N collision. Hydrogen could be such an agent: it is known to be able to lower the knock-on threshold energy of its neighbours by acting as an impedance-matching medium for the electron impact, whereby the fast electron impacts the hydrogen and the hydrogen subsequently impacts an atom in the lattice. This mechanism can transfer an energy to a lattice atom that is nearly 4x higher than the maximum amount that can be transferred directly from the fast electron (Bond et al. 1987). Hydrogen was quite likely to be present as a migrating adatom species on the BN monolayer surface.
- b) an electronic transition may be responsible for the ejection of the initial atom, e.g a double or even higher ionization.

Given the fact that hole creation was a rare event, and that it was seen to occur mostly in areas close to the hydrocarbon overlayers, the first explanation appears more likely. Work at different energies and with different types of light  $Z$  materials is likely to clarify the mechanism in the future.

## 4.2 Future directions

In pre-aberration correction days, a STEM that could form a probe whose size was  $50 \lambda$  ( $1.3 \text{ \AA}$  at 200 keV) was “top of the line”. With aberration correction, we have progressed to probe sizes of the order of  $20 \lambda$  ( $0.74 \text{ \AA}$  at 100 kV,  $1 \text{ \AA}$  at 60 kV). In order to progress to  $0.5 \text{ \AA}$  probe size at 60 keV, or to  $1 \text{ \AA}$  at 20 keV, the probe size will have to come down to around  $10 \lambda$ . This is a tall order even by aberration correction standards, but it should be reachable with chromatic correction and a further reduction of instrumental instabilities.

The principal goal of making the probe size smaller is to concentrate the signal from each atom into a smaller area. This results in a better signal-to-noise ratio for individual atoms, and in an improved ability to resolve atoms lying close to each other. At resolutions better than 1 Å, we will be taking advantage of the fact that ADF STEM basically images the atomic nucleus, which is very much smaller than the electron orbitals around it. The potential ADF resolution is therefore considerably higher than the resolution of techniques that image the outer electron orbitals, such as STM and AFM.

The main reason for lowering the operating energy further is to avoid radiation damage and other sample instabilities to such an extent that complex structures can be analyzed with irradiation doses large enough to allow atom-by-atom imaging and analysis. A secondary reason is to improve the resolution of EELS elemental mapping, provided of course that it is confirmed that the resolution does improve at lower primary energies.

The analysis that we have been able to perform on monolayer BN with impurities shows that in favourable circumstances, every single atom in a small area of a sample can be resolved and individually identified. The challenge posed by this success is to see whether this approach can be made applicable to small molecules of unknown structure. This will amount to extending the tantalizing glimpses of 3-D sample structures that we saw in the thicker sample areas of Figs 11 and 13 to a full 3-D characterization of non-periodic structures, such as molecules of unknown shapes.

Some molecules may be able to withstand the high dose we have used here, but many others will require doses of less than 100 electrons / Å<sup>2</sup>. The best way to achieve 3-D atomic resolution will then probably be to combine the low-dose techniques developed for determining macromolecular structures (Frank 2006) with ADF imaging. The molecules will probably be supported on hydrophilic substrates such as monolayer graphane oxide (Pantelic et al. 2010), and they could also be encapsulated in nanotubes. The image contribution from these kinds of support structures can in principle be modeled and subtracted, rendering them nearly invisible, apart from an increase in statistical noise.

Imaging a large number of identical and separated molecules of random orientations may then lead to an atomically resolved 3D structure of the molecule, at illumination doses low enough to avoid serious damage, even in biological molecules containing hydrocarbon chains. The task should be made easier by using chemical information about the molecule to narrow down the search among candidate structures. Since we often know the amino acid sequence making up a particular protein, but do not know its precise structure, this kind of capability should find a very wide range of applications.

The sample temperature may need to be lowered for this work, without sacrificing the sample stability. This is, however, a problem that has been solved several times before, as much of biological microscopy is carried out only at low temperatures.

Another promising avenue for the capabilities demonstrated here will be to analyze and track the motion of individual atoms of various species. Our improving ability to image single atoms has brought many insights to catalysis (e.g. Rashkeev et

al. 2007) in which individual atoms such as Au, Pt and Ru exhibit markedly different catalytic properties from atomic aggregates. This kind of ability can now be extended to lighter atoms, and it is a safe bet that it will lead to new insights on both natural and man-made materials.

Determining which atom is which using the atom's ADF intensity would be a lot easier if one had the ability to put down atomic markers of known species. This could take the form of a simple evaporator or an ion beam deposition system, preferably in-situ or designed so that it's easy to go back and forth between the deposition system and the microscope. We plan on constructing such a system and using it to measure experimentally the scattering cross-sections for atoms across the periodic table.

Systematic measurement of experimental cross sections should help resolve a long-standing controversy about the ratio of the ADF cross sections of U and C. On theoretical grounds, one would expect the  $I_{\text{uranium}}/I_{\text{carbon}}$  to be  $\sim (92/6)^{1.5} = 60$ . However, the ratio has been measured to be  $9\pm 4$  in a 40 keV STEM (Wall, 1979), and the measurement so far appears to have stood the test of time. If it is correct, our understanding of useful approximations that model the scattering process is clearly rather limited, and it is high time we revisited the subject.

## 5 Conclusion

A highly respected colleague of ours recently remarked that with aberration correction, it is as if a veil of fog has finally lifted from the objects we look at with electron microscopes. This is indeed very much the case already now, and the fog will undoubtedly lift even more in the future. The dreams of electron microscopy pioneers such as Ruska, Scherzer, Gabor and Crewe about being able to see matter clearly with the electron microscope are thus being realized, step by step. It is an exciting time to be active in this field, and to be helping to advance it further.

## 6 Acknowledgement

We are grateful to Dr. S.J. Pennycook and the Oak Ridge National Laboratory for the use of the Nion UltraSTEM after its installation, to Drs D. Geohegan, V. Nicolosi and K. Suenaga for the provision of samples, to G.J. Corbin, J. Woodruff and Drs C.S. Own and Z. Szilagyι for their part in the design and construction of the Nion UltraSTEM, and to Prof. D.J.H. Cockayne and Drs N.J. Bacon, P.D. Nellist and S.J. Pennycook for useful discussions. Research at Oak Ridge National Laboratory (MFC) was sponsored by the Division of Materials Sciences and Engineering of the U.S. Department of Energy.

## References

- Alem N., Erni R., Kisielowski C., Rossell M. D., Gannett W., and Zettl A. (2009). Atomically thin hexagonal boron nitride probed by ultrahigh-resolution transmission electron microscopy, *Phys. Rev. B* 80, 155425-1 (7 pp.).
- Archard G.D. (1955). Two new simplified systems for the correction of spherical aberration in electron lenses. *Proc. Roy. Soc. B* 68, 156–164.
- Ardenne, M. von (1938). Das Elektronen-Raster Mikroskop (The scanning electron microscope). *Z. Tech. Phys* 19, 407-416.
- Ardenne, M. von (1940). *Elektronen übermikroskopie*. Springer, Berlin.
- Ardenne, M. von (1985). On the history of scanning electron microscopy, of the electron microprobe, and of early contributions to transmission electron microscopy. *Adv. Electronics and Electron Physics*. Supplement 16, 1-21.
- Bacon N.J. (2008). private communication.
- Bacon N.J., Corbin G.J., Dellby N., Hrnčirik P., Krivanek O.L., Murfitt M.F. and Szilagyí Z.S. (2010). A High-Performance 200 kV Cold Field Emission Electron Source. *Proceedings International Microscopy Congress (IMC-17, Rio)*, to be published.
- Batson P.E. (1987) Low-workfunction field-emission source for high-resolution EELS. *Proceedings 45<sup>th</sup> EMSA Meeting* (G.W. Bailey, ed., San Francisco Press, San Francisco) 132-133.
- Batson P.E. (2006). Characterizing probe performance in the aberration corrected STEM *Ultramicroscopy* 106, 1104-1114.
- Batson P.E., Dellby N. and Krivanek O.L. (2002). Sub-ångstrom resolution using aberration corrected electron optics. *Nature* 418, 617–620.
- Beck V. (1979). Hexapole spherical-aberration corrector. *Optik* 53, 241–245.
- Beck V. and Crewe A.V. (1975). High resolution imaging properties of the STEM. *Ultramicroscopy* 1, 137-144.
- Beck V. and Crewe A.V. (1976) *Proceedings 32nd EMSA meeting*, 578.
- Bond, G.M., Robertson, I.M., Zeides, F.M. and Birnbaum, H.K. (1987). Sub-threshold electron irradiation damage in hydrogen-charged aluminium. *Phil. Mag.* A55, 669.
- Bosman M., Keast V.J., García-Muñoz J.L., D’Alfonso A.J., Findlay S.D. and Allen L.J. (2007). Two-dimensional mapping of chemical information at atomic resolution. *Phys. Rev. Letts* 99, 086102-1 to 086102-4.
- Crewe, A.V. (1979). Direct imaging of single atoms and molecules using the STEM. *Proceedings 47th Nobel Symposium, Chemica Scripta* 14, 17-20.
- Crewe A.V. (2009). The work of Albert Victor Crewe on the scanning transmission electron microscope and related topics. *in: Cold field emission and the scanning transmission electron microscope*. *Advances in imaging and electron physics* vol. 159. (P.W. Hawkes, editor, Elsevier, Amsterdam), 1-61.
- Crewe A.V., Eggenberger D.N., Wall J. and Welter L.M. (1968b). Electron Gun Using a Field Emission Source, *Rev. Sci. Inst.* 39, 576-583.
- Crewe A.V. and Kopf D. (1980). A sextupole system for the correction of spherical aberration. *Optik* 5, 1–10.
- Crewe A.V., Wall J. and Welter L.M. (1968a). A high-resolution scanning transmission electron microscope. *J. Appl. Phys.*, 39, 5861-5868.
- Crewe A.V., Wall J. and Langmore J. (1970). Visibility of single atoms. *Science*, 168, 1338-1340.
- Dellby N., Krivanek O.L., Nellist P.D., Batson P.E. and Lupini A.R. (2001). Progress in aberration-corrected scanning transmission electron microscopy. *J. Electron Microscopy* 50, 177–185.

- Deltrap J.H.M. (1964a). Correction of spherical aberration with combined quadrupole-octopole units. *Proc. EUREM-3 (Prague)*, 45.
- Deltrap J.H.M. (1964b). Correction of spherical aberration of electron lenses. Ph. D. dissertation, University of Cambridge.
- Egerton R.F. (1996). *EELS in the electron microscope*, (2nd edition, Plenum Press, New York).
- Egerton, R.F. (2006). Limits to the spatial, energy and momentum resolution of electron energy-loss spectroscopy. *Ultramicroscopy* 107, 575-586.
- Erni R., Rossell M.D., Kisielowski C., and Dahmen U. (2009). Atomic-resolution imaging with a sub-50-pm electron probe. *Phys. Rev. Lett.* 102, 096101 (4 pages).
- Fejes P.L., Editor (1978). *Proceedings of Cornell Specialist Workshop in Analytical Electron Microscopy (July 1978)*.
- Fitting-Kourkoutis L. (2010). From imaging individual atoms to atomic resolution 2D mapping of bonding. *Proceedings 12<sup>th</sup> Interantional Microscopy Congress (Rio de Janeiro)*, to be published.
- Frank J., (2006). *Three-Dimensional Electron Microscopy of Macromolecular Assemblies* (Oxford University Press, Oxford).
- Fraser H.L., Geiss R.H., Isaacson M.S., Joy D.C., Maher D.M. and Silcox J., editors (1976). *Proceedings of Cornell Specialist Workshop in Analytical Electron Microscopy (August 1976)*.
- Girit Ç.Ö., Meyer J.C., Erni R., Rossell M.D., Kisielowski C., Yang L., Park C-H., Crommie M.F., Cohen M.L., Louie S.G. and Zettl A. (2009). Graphene at the Edge: Stability and Dynamics, *Science* 323, 1705-1708.
- Haider M., Uhlemann S., Schwan E., Rose H., Kabius B. and Urban K. (1998). Electron microscopy image enhanced. *Nature* 392, 768–769.
- Hardy D.F. (1967). *Combined Magnetic and Electrostatic Quadrupole Electron Lenses*. Ph. D. dissertation, University of Cambridge.
- Hartel P., Rose H. and Dignes C. (1996). Conditions and reasons for incoherent imaging in STEM, *Ultramicroscopy* 63, 93–114.
- Hawkes P.W., editor (2008). *Aberration-Corrected Electron Microscopy*, *Advances in Imaging and Electron Physics*, volume 153 (Elsevier, Amsterdam).
- Hawkes P.W., editor (2009). *Cold field emission and the scanning transmission electron microscope*. *Advances in Imaging and Electron Physics*, volume 159 (Elsevier, Amsterdam).
- Hawkes P., editor (2009). *Cold field emission and the scanning transmission electron microscope*. *Advances in imaging and electron physics vol. 159*. (Elsevier, Amsterdam).
- Hernandez Y., Nicolosi V., Lotya M., Blighe F.M., Sun Z., De S., McGovern I.T., Holland B., Byrne M., Gun'ko Y., Boland J., Niraj P., Duesberg G., Krishnamurthy S., Goodhue R., Hutchison J., Scardaci V., Ferrari A.C. and Coleman J.N. (2008). High yield production of graphene by liquid phase exfoliation of graphite, *Nature Nanotechnology* 3, 563 – 568.
- Isaacson, M. and Johnson, D. (1975) Microanalysis of light-elements using transmitted energy-loss electrons. *Ultramicrosc.*, 1, 33-52.
- Isaacson M., Kopf D., Ohtsuki M. and Utlaut M. (1979). Atomic imaging using the dark field annular detector in the STEM, *Ultramicroscopy* 4, 101-104.
- Isaacson M., Kopf D., Utlaut M., Parker N.W. and Crewe A.V. (1977). Direct observation of atomic diffusion by scanning transmission electron microscopy. *Proc Natl. Acad. USA* 74, 1802-1806.
- Jin C., Lin F., Suenaga K. and Iijima S. (2009). Fabrication of a freestanding boron nitride single layer and its defect assignments, *Phys. Rev. Letts* 102, 195505 (4 pages).
- Koshino M., Tanaka T., Solin N., Suenaga K., Isobe H. and Nakamura E. (2007). Imaging of single organic molecules in motion, *Science* 316, 85.

- Krivanek O.L., Dellby N., Spence A.J., Camps R.A. and Brown L.M. (1997). Aberration correction in the STEM. *Inst. Phys. Conf. Ser.* vol. 153 (Proceedings 1997 EMAG meeting, Rodenburg J.M., editor), 35–40.
- Krivanek O.L., Dellby N. and Lupini A.R. (1999). Towards sub-Å electron beams. *Ultramicroscopy* 78, 1–11.
- Krivanek, O.L. Dellby N., Keyse R.J., Murfitt M.F., Own C.S. and Szilagy Z.S. (2008a). Aberration-corrected STEM and EELS, in: Hawkes P.W. (Ed.), *Advances in imaging and electron physics* (pp. 121-155) London: Academic Press.
- Krivanek O.L., Corbin G.J., Dellby N., Elston B.F., Keyse R.J., Murfitt M.F., Own C.S., Szilagy Z.S. and Woodruff J.W. (2008b). An electron microscope for the aberration-corrected era. *Ultramicroscopy* 108, 179–195.
- Krivanek, O.L., Dellby N. and Murfitt M.F. (2009). Aberration correction in electron microscopy. in: Orloff J. (Ed.), *Handbook of charged particle optics*, (2nd ed., Boca Raton: CRC Press) 601-640.
- Krivanek O.L., Chisholm M.F., Nicolosi V., Pennycook T.J., Corbin G.J., Dellby N., Murfitt M.F., Own C.S., Szilagy Z.S., Oxley M.P., Pantelides S.T. and Pennycook S.J. (2010a). Atom-by-atom structural and chemical analysis by annular dark-field electron microscopy *Nature* 464, 571-574. See also the on-line supplementary information available at <http://www.nature.com/nature/journal/v464/n7288/supinfo/nature08879.html>.
- Krivanek O.L., Dellby N., Murfitt M.F., Chisholm M.F., Pennycook T.J., Suenaga K., Nicolosi V. (2010b). Gentle STEM: ADF imaging and EELS at low primary energies. *Ultramicroscopy*, in press (doi:10.1016/j.ultramic.2010.02.007).
- Krivanek O.L., Dellby N., Murfitt M.F., Szilagy Z.S., Chisholm M.F., and Suenaga K. (2010c). Slow and fast atomic motion observed by aberration-corrected STEM. *Proceedings MSA meeting (Portland)*, to be published.
- Liu Z., Yanagi K., Suenaga K., Kataura H. and Iijima S. (2007). Imaging the dynamic behaviour of individual retinal chromophores confined inside carbon nanotubes. *Nature Nanotechnology* 2, 422-425.
- Martin E.E., Trolan J.K. and Dyke W.P. (1960). Stable, High Density Field Emission Cold Cathode. *Journal of Applied Physics* 31, 782-789.
- Meyer J.C., Chuvilin A., Algara-Siller G., Biskupek J. and Kaiser U. (2009). Selective Sputtering and Atomic Resolution Imaging of Atomically Thin Boron Nitride Membranes, *Nano Lett.* 9, 2683-2689.
- Muller D.A., Fitting-Kourkoutis L., Murfitt M.F., Song J.H., Hwang H.Y., Silcox J., Dellby N. and Krivanek O.L. (2008). Atomic-scale chemical imaging of composition and bonding by aberration-corrected microscopy. *Science* 319, 1073–1076.
- Müller H., Uhlemann S., Hartel P. and Haider M. (2006). Advancing the Hexapole C<sub>s</sub>-Corrector for the Scanning Transmission Electron Microscope. *Microscopy and Microanalysis* 12, 442–455.
- Nellist P.D., Chisholm M.F., Dellby N., Krivanek O.L., Murfitt M.F., Szilagy Z.S., Lupini A.R., Borisevich A., Sides W.H. and Pennycook S.J. (2004). Direct sub-angstrom imaging of a crystal lattice. *Science* 305, 1741–1742.
- Pantelic R.S., Meyer J.C., Kaiser B., Baumeister W. and Pitzko J.M. (2010). Graphene oxide: A substrate for optimizing preparations of frozen-hydrated samples. *Journal of Structural Biology* 170, 152–156.
- Pennycook S.J. (2010) A Scan through the History of STEM. This volume.
- Rashkeev S.N., Lupini A.R., Overbury S.H., Pennycook S.J. and Pantelides S.T. (2007). Role of the nanoscale in catalytic CO oxidation by supported Au and Pt nanostructures. *Phys. Rev. B* 76, 035438 (8 pages).
- Rose H. (1971) *Abbildungseigenschaften sphärisch korrigierter elektronenoptischer Achromate.* *Optik* 33, 1-24.

- Rose H. (1990). Outline of a spherically corrected semiplanatic medium-voltage transmission electron-microscope. *Optik* 85, 19–24.
- Saito R., Dresselhaus G. and Dresselhaus M.S. (1998). *Physical Properties of Carbon Nanotubes*. (World Scientific Publishing, Singapore).
- Sawada H., Hosokawa F., Kaneyama T., Ishizawa T., Terao M., Kawazoe M., Sannomiya T., Tomita T., Kondo Y., Tanaka T., Oshima Y., Tanishiro Y., Yamamoto N., and Takayanagi K. (2007). Achieving 63 pm Resolution in Scanning Transmission Electron Microscope with Spherical Aberration Corrector. *Jap. J. Appl. Phys.* 46, L568–L570.
- Sawada H., Tanishiro Y., Ohashi N., Tomita T., Hosokawa F., Kaneyama T., Kondo Y. and Takayanagi K. (2009). STEM imaging of 47-pm-separated atomic columns by a spherical aberration-corrected electron microscope with a 300-kV cold field emission gun. *Journal of Electron Microscopy* 58, 357–361.
- Suenaga K., Sato Y., Liu Z., Kataura H., Okazaki T., Kimoto K., Sawada H., Sasaki T., Omoto K., Tomita T., Kaneyama T. and Kondo Y. (2009). Visualizing and identifying single atoms using electron energy-loss spectroscopy with low accelerating voltage. *Nature Chemistry* 1, 415–418.
- Suenaga K., Tence M., Mory C., Colliex C., Kato H., Okazaki T., Shinohara H., Hirahara K., Bandow S. and Iijima S. (2000). Element-selective single atom imaging. *Science* 290, 2280–2282.
- Suenaga K., Wakabayashi H., Koshino M., Sato Y., Urita K. and Iijima S. (2007). Imaging active topological defects in carbon nanotubes. *Nature Nanotechnology* 2, 358 - 360. (see the supplementary materials).
- Shao Z. (1988). On the fifth order aberration in a sextupole corrected probe forming system. *Rev. Sci. Instrum.* 59, 2429–2437.
- Swanson L.W. and Martin N.A. (1975). Field electron cathode stability studies: Zirconium/tungsten thermal-field cathode. *Journal of Applied Physics* 46, 2029–2050.
- Swanson L.W. and Schwind G.A. (2009). Review of ZrO/W Schottky Cathode, *in: Handbook of Charged Particle Optics*, (Second Edition, Orloff J., Editor, CRC Baton Rouge).
- Thomson M.G.R. (1968). The aberrations of quadrupole electron lenses. Ph.D. dissertation, University of Cambridge.
- Varela M., Findlay S.D., Lupini A.R., Christen H.M., Borisevich A.Y., Dellby N., Krivanek O.L., Nellist P.D., Oxley M.P., Allen L.J. and Pennycook S.J. (2004). Spectroscopic Imaging of Single Atoms Within a Bulk Solid. *Phys. Rev. Lett.* 92, 095502.
- Voyles P.M., Muller D.A., Graul J.L., Citrin P.H. & Gossman H.-J.L. (2002). Atomic-scale imaging of individual dopant atoms and clusters in highly n-type bulk Si, *Nature* 416, 826–829.
- Wall J.S. (1979). Limits on visibility of single heavy atoms in the scanning transmission electron microscope – an experimental study. *Proceedings 47th Nobel Symposium, Chemica Scripta* 14, 271–278.
- Zach J. and Haider M. (1995). Aberration correction in a low-voltage SEM by a multipole corrector. *Nucl. Instr. Meth. A* 363, 316–325.
- Zobelli A., Gloter A., Ewels C.P., Seifert G. and Colliex C. (2007). Electron knock-on cross section of carbon and boron nitride nanotubes, *Phys Rev.* B75, 245402 (9 pages).