Nanometre electron beam sculpting of suspended graphene and hexagonal boron nitride heterostructures

Nick Clark$^{1,2}$, Edward A Lewis$^1$, Sarah J Haigh$^{1,2,3}$ and Aravind Vijayaraghavan$^{1,2,3}$

$^1$ School of Materials and University of Manchester, Oxford Road, Manchester M13 9PL, United Kingdom
$^2$ National Graphene Institute, University of Manchester, Oxford Road, Manchester M13 9PL, United Kingdom
$^3$ Author to whom any correspondence should be addressed.

E-mail: aravind@manchester.ac.uk

Keywords: graphene, boron nitride, nanolithography, heterostructure, e-beam

Supplementary material for this article is available online

Open Access citation and DOI.

Abstract

Nano-patterned and suspended graphene membranes find applications in electronic devices, filtration and nano-pore DNA sequencing. However, the fabrication of suspended graphene structures with nanoscale features is challenging. We report the direct patterning of suspended membranes consisting of a graphene layer on top of a thin layer of hexagonal boron nitride which acts as a mechanical support, using a highly focused electron beam to fabricate structures with extremely high spatial resolution within the scanning transmission electron microscope (STEM). The boron nitride support enables the fabrication of stable graphene geometries by preventing intrinsic strain in graphene membranes from distorting the patterned features. Line cuts with widths below 2 nm are reported. We also demonstrate that the extent of cutting can be monitored in situ utilising electron energy loss spectroscopy (EELS).

Graphene is an atomically thin [1] carbon allotrope with exceptional mechanical strength [2] and electron conductivity [3]. Patterning graphene at extremely small length scales can be used to create interesting physical systems, such as nanogap electrodes for molecular electronics [4], molecular separation membranes [5], nanopores for DNA sequencing [6], and nanoribbons [7] for studying electron transport phenomena.

Molecular electronics [8, 9] represents the ultimate in device miniaturisation, where individual molecules are used as circuit components. The overall performance of such systems is largely determined by the electrode/molecule interface (anchor) and ensuring the stability of such contacts is a major challenge. Thiol-gold based systems are limited by the variability in electrode geometry between repeat devices and by the instability of thiol-gold bonding resulting from the mobility of gold atoms at ambient temperatures [10]. Systems based on sp$^3$ bonded carbon electrodes formed from graphene [11] or carbon nanotubes (CNTs) [12] have been suggested as alternative materials systems which are predicted to have a number of advantages [13] as long as the gap can be made narrow enough to be bridged by molecules with lengths up to a few nm. Carbon electrodes also have the advantage of being atomically stable at room temperature [14]. Molecules terminated with polyaromatic hydrocarbon end groups can bridge to carbon electrodes in a secure and conductive way via mechanical attachment of the aromatic moieties either side of the gap [15–17].

Solid state nanopores are of interest for filtration and biomolecular analysis, particularly DNA sensing [6]. Typically, a DNA molecule is threaded through a hole in a thin membrane using an applied voltage, and the ionic current across the membrane is monitored. It has been suggested that different bases along the DNA molecule block the ionic current through the pore by differing amounts, in principle allowing determination of the base order. Nanopores in large graphene sheets have demonstrated DNA detection [18–20], although in all cases the translocation has occurred too quickly to resolve signals from individual bases. An alternative strategy is measuring the conductance of a graphene nanoribbon patterned with a nanopore in the centre. Theoretical simulations have suggested that electron transport through the ribbon is strongly affected by interactions with nucleotides in the pore [21], with different bases having differing signals due to variations in coupling strength. Detection using this mechanism has
been reported [22], although sequencing remains elusive due to the high speed of the translocation and low signal to noise ratios seen in realised devices, most likely as a result of the fabrication challenges.

Graphene nanoribbons themselves are of scientific and technological interest, particularly for electronic applications. As the ribbon width is reduced, there are significant structure dependent (primarily edge orientation) changes to the electronic, magnetic and optical properties [23–26], including the opening of a band gap [7,27], due to the quasi 1-dimensional (1D) spatial confinement at few nanometre scale widths [28]. They are therefore promising candidates for building electrical interconnects, photodetectors, transistors, and sensors [29–32]. Nonetheless, fabrication limitations mean that the properties of graphene nanoribbons with widths below around 10 nm have not been fully experimentally explored [33].

Graphene is typically patterned using electron beam lithography (EBL) and subsequent plasma etching into structures such as Hall bars [34], nanoribbons [23], quantum dots [35] and nanogap electrodes [12, 36]. However, fundamental limitations of electron scattering [37] mean that this technique cannot be used to pattern features at length scales of less than ~10 nm. Scanning probe-based lithography on graphene has also been demonstrated [38, 39], but is slow and unwieldy. Carbon nanogap electrodes are typically fabricated using feedback controlled electroburning [14, 40] to open a narrow physical gap in a graphene ribbon. Direct write patterning of suspended graphene promises a far higher patterning resolution and control of the edge geometry. Patterning in focused ion beam (FIB) systems using gallium ions [41] or helium ions [42–45] has been reported, but can cause damage to the neighbouring graphene ribbon, reducing the potential of this approach for electronic applications [46].

Recently, the highly focussed electron beam of the transmission electron microscope (TEM) has been used to pattern [47] and generate nanopores [48–50] and nanoribbons [51–54] in suspended graphene sheets. The use of this technique in patterning carbon nanotubes has previously been reported [55, 56], but geometry means that it can be more usefully applied to planar structures. Nanopatterning in the scanning transmission electron microscope (STEM) can be combined with simultaneous electron diffraction or atomic resolution imaging, enabling the fabrication of nanostructures with controlled edge geometries [57]. Spectroscopic techniques such as electron energy loss spectroscopy (EELS) can be used while patterning [58] to monitor the process. In situ holders also allow patterning to be performed at a range of temperatures and patterning at 600 °C has been found to induce self-repair of the graphene lattice [59]. However, the fact that the graphene structure needs to be self-supporting limits the type of structures that can be produced. For example, use of this technique to produce nanogap electrodes is extremely challenging, as the electrodes need to be electrically isolated from each other, meaning they must be mechanically separated. Once cut, they would be unlikely to maintain the required sub-5 nm gap due to intrinsic tension and mechanical instability.

In order to overcome these limitations, we demonstrate the direct write electron beam nanopatterning of two layer heterostructures consisting of monolayer or bilayer graphene on top of a thin (6–9 nm) hexagonal boron nitride (hBN) layer. hBN is commonly used as a substrate for graphene due to its insulating nature and atomic flatness [60] (for thick flakes). As a result, graphene devices on hBN substrates have shown exceptional carrier mobilities [61]. Here we propose its use as an insulating mechanical support in suspended devices, to prevent distortion of the membrane after partial patterning. We report successful patterning of these heterostructures, reproducibly generating line cuts in graphene with lengths of hundreds of nanometres and widths less than 2 nm.

Figure 1 compares the patterning of thin lines in unsupported and hBN supported monolayer graphene. Figures 1(c) and (d) show that cuts in the unsupported graphene ‘bow’ outwards such that the width of the central part of the cut depends on the cut length, rather than any patterning conditions. In contrast, we observe that cuts in the hBN supported graphene show a much narrower cut profile, as the hBN prevents distortion of the membrane (figure 1(e)). Minimum cut widths in the heterostructure were less than 2 nm, whereas in unsupported graphene widths were limited to greater than 5 nm.

A two-step approach whereby conventional EBL and oxygen plasma etching are used to pattern the graphene-hBN heterostructure into nanoribbons and then followed by TEM nanosculpting is a viable route to the fabrication of extremely narrow nanogap electrodes. When applied creatively, this approach could be applied to pattern a wide array of geometries which are not achievable when the graphene is required to be self-supporting.

A significant challenge in generating ultra-narrow patterns is the verification of the effectiveness of the patterning—whether the cut has entirely separated the graphene on both sides, and if so what the true width of the cut is. This has been typically done by imaging the graphene at atomic resolution after the cutting, which necessitates a high electron flux and may cause further damage or edge reconstruction to the graphene, and can induce broadening or closure of the cut [59]. In order to overcome this, we demonstrate here that it is possible to monitor the cutting progression in real time as well as the width of the cut using EELS, removing the need for subsequent atomic resolution imaging.

**Results**

Imaging whilst cutting was performed in STEM mode, which allowed accurate control over the electron flux
incident on the sample in order to avoid unwanted beam induced damage [18]. The probability of knock-on damage from the beam is dependent on the incident electron energy, the beam current, and the electron flux. The threshold for knock-on damage for graphene has been theoretically predicted to be ~113 keV [62], but experimental measurements observe some electron sputtering at high magnification for lower voltages between 80–100 keV, as a result of catalytic impurities or beam induced lattice vibrations [63, 64]. For hBN monolayers the knock-on threshold energies are 80 keV and 120 keV for B and N sites respectively [65, 66]. Experimentally, it is challenging to frequently modify the TEM accelerating voltage. Therefore, a constant accelerating voltage of 200 keV provides a suitable compromise between the energy needed to reliably perform patterning with that required to preserve the structure while imaging. We find that no damage is detectable after repeated imaging by employing low magnifications and short dwell times (~2 µs). However, longer dwell times did result in visible damage to the sample.

Figure 2 shows the series of EEL spectra obtained over repeated line scans whilst patterning a straight cut in a graphene monolayer supported by a 6.5 nm thick hBN flake. In total, 11 line scans were used at the same location on the flake, with each scan corresponding to a linear electron fluence of 100 mC m\(^{-1}\). The probe current was 1 nA. The individual edges shown in figure 2 were obtained from the captured core loss EELS spectra integrated along the length of the line scan to reduce noise. The same features are observed in all individual point spectra within these scans (even in the centre of the widest part of the cut). Furthermore, as shown in supplementary information figure S7 (stacks.iop.org/TDM/6/025032/mmedia), the shape of the edges does not change significantly during the cutting process, although the position of the inelastic scattering peak in the low loss spectrum reduces from ~23 eV to ~18 eV. The progression of the peak intensities with subsequent line scans, is plotted in figure 3(a), and allows us to track the relative quantities of boron, nitrogen, and carbon atoms within the interaction volume of the probe. An equivalent quantitative plot normalised by the effective cross section is presented in figure S17 in the supplementary information. The B and N signals decline more quickly than C signal, which is expected as the hBN is thicker than the graphene so the rate of hBN removal is expected to be faster (due to its far greater thickness leading to an increased likelihood of interaction with the transmitted electron beam). However, the greater hBN thickness also means that the graphene may be completely removed before the hBN slab. The carbon signal initially reduces more slowly as mobile contamination on the surface migrates towards the beam, before reducing faster than boron or carbon as

Figure 1. Patterning suspended monolayer graphene and hBN supported monolayer graphene using a 1.4 nA STEM beam. (a) schematic (not to scale) showing the membrane structure when viewed in cross section at the position indicated by the dotted line superimposed on the image in (b). (b)–(e) HAADF STEM images illustrating the patterning of this structure. (b) Shows a low magnification image of the sample. The dark pill shape in the centre is the hole in the silicon nitride membrane. A 9 nm thick hBN flake covers the majority of the membrane and is false coloured blue. A monolayer graphene flake covers the entire membrane, and is false coloured red. Where the graphene is supported on the hBN is purple. The scale bar is 5 µm. (c) Shows a STEM cut from the suspended region of the graphene (left) to the supported region (right). The cut has a width of ~7 nm on the unsupported graphene, and a width of just ~2 nm on the hBN supported area. The linear electron fluence for this cut was 625 mC m\(^{-1}\). The scale bar is 100 nm. The inset shows a higher magnification image of the indicated area. Small bridges over the cut are visible. The scale bar is 20 nm. (d) Shows 2 cuts in unsupported monolayer graphene with lengths of 161 and 332 nm, using equivalent linear fluences of 311 and 151 mC m\(^{-1}\) for the short and long cuts respectively. The cut features are broadened by tension within the membrane, with the widest areas in the centre of the cut. The scale bar is 50 nm. (e) Shows a long cut on a hBN supported region of monolayer graphene, with a linear fluence of 1.8 mC m\(^{-1}\). In contrast to (d) no bowing or deformation of the membrane after patterning is visible. The scale bar is 50 nm.
the area is depleted of mobile carbon. As can be seen from figure 3(a) (and the original spectra in figures 2 and S7), the ionization edges are still present even when a continuous cut has been formed through the entire thickness of the graphene-hBN stack.

The low loss spectral region is widely used to monitor specimen thickness by estimating a value for $t/\lambda$ from the relative intensities of the parts of the spectrum corresponding to elastic and inelastic scattering and using the log ratio technique [67], where $t$ is the sample thickness and $\lambda$ is the mean free path for inelastic scattering. Hence we can use the $t/\lambda$ value obtained during successive scans to monitor cutting progression. Figure 3(a) shows the $t/\lambda$ value obtained alongside the normalised ionization edge strengths for the three main atomic species present. While a significant drop in the $t/\lambda$ value is observed during cutting, it does not drop to zero after patterning is finished and sees a smaller relative decline than the core loss signals in the high loss region of the EEL spectra.

The cut progression can also be monitored by looking at the change in the high-angle annular dark field imaging (HAADF) intensity of the cut feature, as shown in the survey images (left hand column of figure 2). Figure 3(b) shows intensity line profiles extracted from the successive survey images perpendicular to the cut and integrated along the cut direction. The changes in relative HAADF intensity and full width at half maximum (FWHM) of the patterned line as the cutting proceeds are plotted in figure 3(c).

The final resultant cut is pictured in figure 3(d) (rotated 90° clockwise from the low dose images in figure 2). The width of the feature is around 2 nm for most of its length but broadens to 3.5 nm at one end. When raster scanning the electron beam, several distortions can arise where the true probe position does not match the desired position. One of the most common of these is fly-back distortion resulting from the sudden change of beam location from the end of a line to the beginning of the next [68]. Fly-back distortion can be reduced by pausing at the starting point of the scan, but this delay results in the first point scanned receiving a slightly higher electron dose and produces the broadening observed in the patterned line in figure 3(d).
Figure 4 examines the compositional and structural changes that result from prolonged imaging of a line cut. The central square area of figure 4(a) was scanned continuously for 344 s with a total electron fluence of $8.7 \times 10^6$ e Å$^{-2}$ and the resulting series of HAADF STEM images shown as SI Video 1. During imaging the original horizontal cut is observed to increase in width to ~7 nm wide and becomes bordered by an area of >5 nm width which has a lower HAADF intensity than the surrounding material. As shown in figure 4(a), acquiring the video has produced a vertical feature on the left of the scanned area, which is where the greater electron fluence resulting from video fly-back correction has caused the electron beam to etch completely through the sample. Subsequent EEL spectrum image of the specimen region in figure 4(a) allows compositional mapping and spectral analysis of the areas with different HAADF intensity (figures 4(b)–(g)).

In the centre of the ~7 nm wide cut no EELS edges are visible (region 2 in figure 4(a) and red spectra in figure 4(b)). Comparison of the spectra at the regions of the flake with different HAADF intensity shows that the darker region bordering the cut contains only B and N edges (region 3 in figure 4(a) and blue spectra in figure 4(b)), while the rest of the substrate contains the expected B, N, and C (region 4 in figure 4(a) and pink spectra in figure 4(b)). This suggests that the darker region bordering the cut is where prolonged irradiation has completely etched the graphene away at the edges but the thicker hBN substrate remains. The areas where the graphene has been etched away, including holes separate from the main cut feature, are clearly visible in the C map (figure 4(d)). The O map is unclear due to the relatively low signal to noise level, but the relative oxygen percentage map in figure 4(g) shows that the O signal is only significant around the edges of the main cut feature. This distribution of O signal is also seen in elemental maps of cut features acquired using energy dispersive x-ray spectroscopy (EDXS) (as illustrated in figure S16 in the supplementary information). This is consistent with theoretical predictions [69], and previous atomic resolution EELS studies of hBN monolayers which show oxygen atoms can substitute into nitrogen sites [70]. The EELS mapping in figure 4 demonstrates that the loss of graphene causes a measurable decrease in the HAADF STEM intensity. Thus for cut features where the HAADF intensity is constant up to the cut edge we conclude that the graphene reaches to the edge of the cut feature within our measurement accuracy of ±0.1 nm.

Comparison of an area of the substrate where the video was acquired (region 4) with an area outside the videoed region (region 5) demonstrates a further effect of prolonged imaging; region 5 contains a high percentage of oxygen and carbon, likely to be
the result of beam induced fixation of mobile surface contamination. Additional oxygen is also observed in the vertical cut formed at the left-hand edge of the video area (region 1 in figure 4(a) and green spectra in figure 4(b)) in the spectra. Inhomogeneous hydrocarbon contamination is a ubiquitous problem in electron microscopy [71], especially with organic samples where plasma cleaning is challenging. For our samples the mobile surface contamination is likely to consist of polymeric residue remaining from flake transfer processes.

The cut feature in figure 3 was not completely continuous, with bridges crossing the gap especially in the centre of the line. Examination of EELS line scans parallel and perpendicular to the cut edge reveals that the core-loss signals never drop to zero (see supplementary information figure S5). In contrast, the horizontal cut feature after video acquisition in figure 4 is sufficiently large that the EELS core-loss signals for B, C, and N all drop to zero in the centre. Examination of EELS line scans perpendicular to the cut edge reveals that the core-loss signals only become negligible when the probe is $>2$ nm from the cut edge, giving us an estimate of our effective EELS interaction radius. The low loss spectra show inelastic scattering events are visible in the spectra even in the centre of the 7 nm wide cut feature; (calculated $t/\lambda$ values are superimposed in a table inset in figure 4(a)). This data is consistent with the behaviour observed for the thinner cut features shown in figures 2 and 3 and indicates that the effective area under test when using the log ratio [67] method to specify the sample thickness has a radius of at least 3.5 nm under our patterning conditions. The localization distance [72, 73] for inelastic scattering processes is inversely proportional to the energy loss involved, therefore higher energy loss processes (such as the core loss ionization peaks) are more spatially constrained than lower energy processes, such as the inelastic scattering events in the low loss spectrum. Thus, it is to be expected that the low loss features will persist further from the sample edge than the core-loss.

Although the effective EELS interaction area can be affected by the accelerating voltage, the interaction energy loss, and beam broadening (in thicker samples), it is dominated by the effective electron probe size [73]. The presence of significant core loss peak intensities 2 nm from the sample edge suggests the electron probe has tails which extend to this radius, although the main probe peak is estimated to be $<1$ nm. This is larger than would normally be used in an aberration corrected STEM but is unsurprising given the exceptionally high probe current (1 nA) used here in order to ensure effective patterning in a timely manner.

We also used EDXS to characterise patterned features, as discussed in the supplementary information. When capturing EDXS line scans perpendicular to previously patterned line features, we note that the EDXS elemental signals for B, C, and N drop to zero even when the original cuts were less than 2 nm in diameter (as shown in Figure S10 in the supplementary information). EDXS maps of cut features also show an absence of elemental counts within the cut

![Figure 4. EELS mapping of large cut features in monolayer graphene on 6.5 nm thick hBN. The horizontal feature is the cut corresponding to that shown in figure 3 after progressive imaging (the acquired images are shown in supplementary video 1). The vertical feature corresponds to the left edge of the frame during the video acquisition and results from fly-back corrections during scanning. (a) Shows a HAADF STEM image of the area. (b) Shows the EELS spectra integrated over the areas highlighted by the corresponding colours in (a). The inset is a zoomed in region showing the location of the O edge. (c)-(f) Show maps of the core loss peak intensities for B, C, N and O edges respectively. (g) Shows the elemental O percentage. All maps are normalised so that a black pixel represents no signal, and a white pixel represents the highest signal level in the image. All scale bars represent 20 nm.](image-url)
features. However, for light elements like C, N and B, the signal to noise ratio of EDXS is poorer than EELS, therefore EELS was preferred for monitoring cutting progression. Nonetheless, both techniques provide a useful means of assessing elemental distribution after patterning and simultaneous acquisition of both spectroscopic signals is preferred.

To test the wider use of this patterning approach on thicker samples we have also investigated patterning of bilayer graphene supported on a hBN flake substrate (7 nm thick). Figure 5 illustrates the importance of high electron flux rate in addition to overall electron fluence to ensure effective cutting. Both cuts received a total electron fluence of 2 C m$^{-1}$, but in figures 5(a) and (b) the cutting was performed with just 2 line scans, each with a line fluence of 1 C m$^{-1}$ per scan, while in figures 5(c) and (d) eight separate scans were used each at a lower fluence of 0.25 C m$^{-1}$ per scan. The slower scans successfully patterned the line feature which has a final measured width of ~2 nm, but the cutting failed when using a larger number of individual scans. The spectroscopic analysis of the region where patterning failed revealed a large growth in the carbon K-edge (figure 5(d)). Thus, the same build-up of mobile contamination species observed in figure 4 is deemed to be the reason for failure. Comparing EEL spectra acquired before and after patterning revealed that both patterning conditions cause a reduction in the normalised core loss edge intensities of B, C and N derived from the integrated energy loss spectra captured while patterning the two areas (left axes). The dose plotted on the horizontal axis corresponds to the total linear electron fluence received by the sample after the scan. The right axis and the green traces show the values of $t/\lambda$ extracted from the respective low loss EEL spectra.

Figure 5. Comparing the effect of patterning parameters for cutting of bilayer graphene. (a) and (c) Show HAADF STEM images of lines patterned in bilayer graphene supported on hBN (7 nm thick), using the same total linear fluence (2 mC m$^{-1}$), and same beam current (2 nA) but written using two and eight beam scans respectively (flux rates of 1 mC m$^{-1}$ per cut and 0.25 mC m$^{-1}$ per cut respectively). The scale bars are 20 nm. (b) and (d) Show the initial and final integrated core loss EEL spectra captured while patterning the lines in (a) and (c). Ionization edges for B, C and N are clearly visible at energies of 188, 284 and 401 eV respectively. The blue and green profiles represent the initial and final cuts respectively. (e) Shows the evolution of the normalised core loss intensity of mobile contamination species observed in figure 4 as a function of the received total fluence is shown in figure 5(b). For the 6 min (3.66 C m$^{-1}$) line, two values are given next to each cut. In contrast to the previously shown examples, there was no time delay between individual cuts. Only total cutting times of >2 min (>1.22 C m$^{-1}$) have produced a cut through in the sample, although the lines are not fully continuous until the cut time >6 min (>3.66 C m$^{-1}$). The FWHM of the lines at the ends and in the centre as a function of the received total fluence is shown in figure 5(b). For the 6 min (3.66 C m$^{-1}$) line, two values are given for the central width; the dark data-point indicates the width of the central cut feature, and the light data-point includes the width of the dark...
grey area above the line. Subsequent EDXS mapping reveals that this dark grey region responds to the local peeling of the graphene away from the cut, as was previously observed in figure 4. The small differences in the cutting behaviour along the length of the cut are most likely due to local variations in the level of contamination and could be reduced by careful specimen cleaning [74, 75]. In addition, although there is a clear increase in the width of the cut for longer cutting times there appears to be an optimal fluence (and flux) where effective cutting is achieved with a minimum cut width of <2 nm. The cutting behaviour also varies depending on the proximity to previously patterned features as discussed further in the supplementary information.

We also investigated the potential for patterning acute angles by cutting angled crosses. This is of particular interest for creating nanogap electrodes for molecular electronic applications; if nanoribbons were simply cut using a linear pattern, this would result in ‘long’ cuts (minimum nanoribbon width is limited by conventional EBL resolution) that could be bridged at any point along their length. However, by cutting a cross feature into the ribbon, it would be possible to define a set of nanogap contacts that are both narrow and ‘short’, so that they can only be bridged in the central area. Figure 7 shows an example of the patterning of angled crosses in a graphene monolayer supported by hBN (6.5 nm thick). An equivalent linear fluence of 1.64 C m⁻¹ was used for the 220 nm long horizontal...
line, and linear fluences of 1.7–2 C m$^{-1}$ were used for the crossing lines. The lines were cut sequentially from left to right without intermediate imaging. The resulting lines were continuous, and had widths below 2 nm. The ‘acute wedge peninsulas’ surrounding the crosses were found to be self-supporting during cutting but were more sensitive to damage during subsequent imaging, with additional etching of the graphene in these regions observed during acquisition of the subsequent EDXS maps (see figure S12 in the supplementary information). The ‘obtuse wedge peninsulas’ were more stable and are separated by just 3 nm at the apex of the cross as shown in the inset of figure 7(b) (although after EDXS mapping and TEM imaging, the distance across the crosses increased to around 4 nm due to the additional radiation dose received). These distances compare favourably with gap widths reported previously for carbon electrodes for molecular electronics [12, 36]. Unfortunately, the specimen area received a significant radiation dose between cutting and HRTEM imaging, which results in the graphene receding from the edge, as seen in the TEM images in figure 7(c) and in figure S12 in the supplementary information, although the width of the cut in the hBN did not appear to be affected.

**Discussion**

We demonstrate that it is possible to reproducibly pattern cut lengths of $>100$ nm whilst maintaining cut widths below 2 nm. Thinner line widths of $\sim 1.5$ nm were possible but these lines were not consistently continuous and were highly dependent on the level of local contamination. We hypothesise that even thinner cut widths could be achievable for cleaner samples. Analysis of the EELS spectra obtained during cutting allowed us to plot the evolution of the relative quantities of different elements in the probe area and hence monitor the sputtering or deposition of B, C and N during cutting. We found that it was possible to use EELS to successfully monitor the patterning process and to infer complete cutting using calibrated baseline values for the EELS core loss peak intensities without directly imaging the sample in a ‘blind’ patterning procedure [58].

Our results demonstrate that a critical linear fluence of 1.5 C m$^{-1}$ with a beam current over 0.5 nA was consistently able to cut through clean areas of suspended membranes consisting of monolayer graphene on 6.5 nm thick hBN in the STEM operated at a 200 kV accelerating voltage. We find that the patterning is relatively robust for monolayer and bilayer graphene with line widths of 2 nm achievable for a range of fluences up to twice the threshold fluence. We observe a weak dependence on total fluence of the cut quality, with slower scan rates providing more effective cutting as a result of the balance between sputtering from the electron beam and the beam driven ingress of mobile carbon contaminants. Inhomogeneous contamination is a ubiquitous problem for graphene research and here significant contamination has been observed to negatively affect cut quality and can even lead to cuts being repaired [76].

These proof of principle results offer the promise of a two stage patterning process for the manufacture of parallel graphene nanodevices with few nanometre resolution. Firstly, the stack is transferred onto a SiN$_x$ membrane. Conventional EBL and oxygen plasma etching can then be used to pattern the graphene into isolated structures on top of the hBN. As the etch rate of hBN is relatively slow and easy to control, this could also be done with an additional protective hBN sheet.
on top of the graphene layer, although it would need to be relatively thin. EBL, metallization and lift-off can also be used at this point to connect the graphene devices to metallic electrodes. This would lead to a linear array of graphene nanoribbons which could then be individually patterned (narrowed or cut) using a STEM beam in a second patterning step, while the whole membrane remains mechanically stable. Ideally, this process would be performed at elevated temperature [59] to enable lattice self-repair, or using a pattern generator such as those commonly seen on EBL scanning electron microscope systems to control the exact position of the beam and avoid exposure of the active elements. Etched or deposited features created near the membrane during the EBL step could be used to work out the transform between the sample and beam coordinate systems and produce the desired pattern without exposing any of the remaining graphene. This would enable current annealing of the nanoribbons (ideally in situ in the TEM) [18] to reduce the level of surface contamination. Other strategies to reduce contamination after transfer could include high temperature vacuum annealing before [75], or even during patterning [59], or annealing in the presence of platinum to help catalytically break down the PMMA residue at high temperature [74].

The nanosculpting results we demonstrate here could be extended to other 2D materials heterostructures. One application is the patterning of materials encapsulated in hBN using the hot stamping transfer technique [77], where the residual contamination layers around the central material are much lower. It could also allow sculpting of air and beam sensitive 2D materials such as black phosphorus [78,79], which is of technological interest due to its layer number dependent band gap.

Conclusion

In summary, we have demonstrated that it is possible to pattern thin suspended graphene and graphene-hBN heterostructures using a high current 200 keV aberration corrected STEM probe. The addition of the thin hBN substrate allows thinner cuts to be achieved by preventing ‘bowing’ of the monolayer due to tension by acting as a mechanical support. We have demonstrated the nanosculpting of cuts with widths of less than 2 nm. We identified a set of optimal cutting parameters which enable ‘blind’ patterning for device fabrication and show that the EELS signal acquired during patterning can be used to track the cutting process until a continuous cut is produced. This approach allows the cut nanostructures to remain un-imaged until they have been electrically characterised, thus preserving the electronic properties and very narrow linewidth structures. Performing this technique on pre-patterned graphene (using conventional EBL and plasma etching) on hBN is suggested as a convenient route for the fabrication of near-atomic precision graphene device structures, such as nanoribbon and nanogap electrode arrays.

Methods

Sample fabrication

The sample fabrication process is illustrated in figure S1 and discussed in detail in the supplementary information. Briefly, the flakes were deposited on patterned membranes of SiNₓ supported by 2.8 mm diameter Si chips. The Si chips were 200 μm thick and coated with 100 nm thick low stress SiNₓ. The membrane was produced by removing the silicon from the central area using a masked KOH etch, and then patterning the membrane using masked reactive ion etching. Flakes of hBN were exfoliated from a bulk crystal on the oxidised surface of silicon wafers. Monolayer graphene flakes were exfoliated onto a PMMA film, which was then used as a carrier film to transfer the graphene onto the hBN flake using a dry transfer [60] process. Flake thicknesses were measured using AFM for the hBN, and optically identified for the graphene [80]. The hBN/graphene stack was then transferred from the oxide coated silicon wafer to the patterned SiNₓ membrane using a conventional wet transfer process [81]. After polymer removal, it was annealed at 280 °C in a H/Ar atmosphere to remove as much polymer contamination as possible. Immediately prior to insertion into the TEM, samples were baked on a hotplate at 300 °C. The produced membranes were monolayer and bilayer graphene on hBN substrates 6–9 nm thick.

Imaging and spectroscopy

High resolution TEM and STEM imaging with simultaneous EDXS and EELS was performed using a probe side aberration-corrected FEI Titan G2 80–200kV with an x-FEG electron source operated at 200kV. High angle annular dark field (HAADF) STEM imaging was performed using a probe current of 0.5–2 pA, a convergence angle of 21 mrad and a HAADF inner collection angle of 54 mrad. EELS was carried out using a Gatan Quantum ER spectrometer with an energy dispersion of 0.25 eV and a collection angle of 60 mrad. EELS was acquired in Dual EELS mode where the low loss region and high loss region were collected separately with different acquisition times (typical acquisition times were 1 × 10⁻⁵ s for the low loss region and 0.01 s for the high loss) and an energy offset of ~150 eV. EDX spectrum imaging was performed using the Titan’s Super-X four silicon drift EDX detector system with a total collection solid angle of 0.7 srad.

Beam patterning

Cutting was performed by acquiring intense line scans using Gatan’s Digital Micrograph or Bruker’s Esprit v1.9 software. This approach enabled simultaneous cutting and chemical analysis of the heterostructures using EELS and EDXS to monitor the progress of the
cutting. The linear fluence (mC/m) experienced by the sample can be adjusted by altering the pixel step size (and number of pixels), the beam current, the pixel dwell time, and/or the number of repeat scans in the same location. In this work we investigated the effect of the dwell time, beam current, and number of repeat scans but kept the step size below 0.5 Å in all cases. Beam currents between 0.5 and 2 nA were used. Full patterning parameters are listed in the supplementary information.

Analysis of EELS spectra

EELS core loss edge intensities were obtained from the captured spectra by using a model based fitting routine contained in the open source python package Hyperspy [82] to split the spectra into its various components. The total intensity of each peak (and its resultant tail) were extracted and normalised using the tabulated Hartree-Slater cross sections from Gatan. Hyperspy [82] to split the spectra into its various components. The total intensity of each peak (and its resultant tail) were extracted and normalised using the tabulated Hartree-Slater cross sections from Gatan.

Acknowledgment

This work was supported by the Engineering and Physical Sciences Research Council (EPSRC) (grants EP/G03737X/1, EP/G035954/1, EP/P009050/1, and EP/L01548X/1), the Defense Threat Reduction Agency (HDTRA1-12-1-0013) and the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation program (Grant Agreement ERC-2016-STG-EvoluTEM-715502 and ERC Synergy HETERO2D). The authors declare no competing financial interest.

Supplementary information available

The supplementary information contains detailed sample fabrication procedures, detailed beam patterning parameters, additional details on the analysis of EELS spectra, results from additional patterned areas and videos of patterning progress.

Notes

The authors declare no competing financial interest.

ORCID iDs

Nick Clark https://orcid.org/0000-0003-3351-5628
Sarah J Haigh https://orcid.org/0000-0001-5509-6706
Aravind Vijayaraghavan https://orcid.org/0000-0001-8289-2337

References

[1] Novoselov K S, Geim A K, Morozov S V, Jiang D, Zhang Y, Dubonos S V, Grigorieva I V and Firsov A A 2004 Electric field effect in atomically thin carbon films Science 306 666–9
[2] Lee C, Wei X D, Kysar J W and Hone J 2008 Measurement of the elastic properties and intrinsic strength of monolayer graphene Science 321 385–8
[3] Bolotin K I, Sikes K J, Jiang Z, Klima M, Fudenberg G, Hone J, Kim P and Stormer H L 2008 Ultrahigh electron mobility in suspended graphene Solid State Commun. 146 351–5
[4] Prins F, Barreiro A, Ruitenbeek J W, Seldenthuis J S, Aliaga-Alcalde N, Vandervsypen L M K and van der Zant H S J 2011 Room-temperature gating of molecular junctions using few-layer graphene nanogap electrodes Nano Lett. 11 4607–11
[5] Joshi R K, Carbone P, Wang F C, Kravets V G, Su Y, Grigorieva I V, Wu H A, Geim A K and Nair R R 2014 Precise and ultrafast molecular sieving through graphene oxide membranes Science 343 752–4
[6] Heerema S I and Dekker C 2016 Graphene nanodevices for DNA sequencing Nat. Nanotechnol. 11 127–36
[7] Han M Y, Ozylilmaz B, Zhang Y and Kim P 2007 Energy bandgap engineering of graphene nanoribbons Phys. Rev. Lett. 98 206805
[8] Aviram A and Ratner M A 1974 Molecular rectifiers Chem. Phys. Lett. 29 277–83
[9] Xiang D, Wang X, Jia C, Lee T and Guo X 2016 Molecular-scale electronics: from concept to function Chem. Rev. 116 6318–440
[10] Strachan D R, Smith D E, Fischbein M D, Johnston D E, Guilton B S, Drndić M, Bonnell D A and Johnson A T 2006 Clean electromigrated nanogaps imaged by transmission electron microscopy Nano Lett. 6 441–4
[11] Cao Y, Dong S H, Lü S, He L, Gan L, Yu X M, Steigerwald M L, Wu X S, Liu Z F and Guo X F 2012 Building high-throughput molecular junctions using indented graphene point contacts Angew. Chem., Int. Ed 51 12228–32
[12] Marquardt C W, Grunder S, Blassczyk A, Dehm S, Henrich R, von Lohneysen H, Mayer M and Krupke R 2010 Electroluminescence from a single nanotube-molecule-nanotube junction Nat. Nanotechnol. 5 863–7
[13] Cho Y, Kim W Y and Kim K S 2009 Effect of electrodes on electronic transport of molecular electronic devices J. Phys. Chem. A 113 4100–4
[14] Burzuri E, Prins F and van der Zant H S J 2012 Characterization of nanometer-spaced few-layer graphene electrodes Graphene 2 26–9
[15] Peterfalvi C G and Lambert C J 2012 Suppression of single-molecule conductance fluctuations using extended anchor groups on graphene and carbon-nanotube electrodes Phys. Rev. B 86 085443
[16] Perez-Jimenez A J and Sancho-Garcia J C 2009 Conductance enhancement in nanographene-gold junctions by molecular pi-stacking J. Am. Chem. Soc. 131 14857–67
[17] He J, Chen K Q and Sun C Q 2012 The weak pi–pi interaction originated resonant tunneling and fast switching in the carbon based electronic devices AIP Adv. 2012137
[18] Puster M, Rodriguez-Manzo J A, Balan A and Drndić M 2013 Toward sensitive graphene nanoribbon–nanopore devices by preventing electron beam-induced damage ACS Nano 7 11283–9
[19] Schneider G F, Kowalczyk SW, Calado V E, Pandraug D, Zandbergen H W, Vandersypen L M K and Dekker C 2010 DNA Translocation through graphene nanoropes Nano Lett. 10 3163–7
[20] Merchant C A et al 2010 DNA translocation through graphene nanoropes Nano Lett. 10 2915–21
[21] Nelson T, Zhang B and Prezhdo O V 2010 Detection of nucleic acids with graphene nanoropes: ab initio characterization of a novel sequencing device Nano Lett. 10 3237–42
[22] Traversi F, Raillon C, Benamer S M, Liu K, Khlybov S, Tosun M, Krasnozhon D, Kis A and Radenovic A 2013 Detecting the translocation of DNA through a nanopore using graphene nanoribbons Nat. Nanotechnol. 8 839–45
[23] Han M Y, Brant J C and Kim P 2010 Electron transport in disordered graphene nanoribbons Phys. Rev. Lett. 104 056801
[24] Son Y W, Cohen M L and Louie S G 2006 Half-metallic graphene nanoribbons Nature 444 347–9
[69] Huber S P, Gullikson E, van de Kruijs R W E, Bijkerk F and Prendergast D 2015 Oxygen-stabilized triangular defects in hexagonal boron nitride Phys. Rev. B 92 245310

[70] Krivanek O L et al 2010 Atom-by-atom structural and chemical analysis by annular dark-field electron microscopy Nature 464 571–4

[71] Mitchell D R G 2015 Contamination mitigation strategies for scanning transmission electron microscopy Micron 73 36–46

[72] Muller D A and Silcox J 1995 Delocalization in inelastic scattering Ultramicroscopy 59 195–213

[73] Egerton R F 2007 Limits to the spatial, energy and momentum resolution of electron energy-loss spectroscopy Ultramicroscopy 107 575–86

[74] Longchamp J N, Escher C and Fink H W 2013 Ultraclean freestanding graphene by platinum-metal catalysis J. Vac. Sci. Technol. B 31 020605

[75] Zan R et al 2012 Scanning tunnelling microscopy of suspended graphene Nanoscale 4 3065–8

[76] Zan R, Ramasse Q M, Bangert U and Novoselov K S 2012 Graphene reknits its holes Nano Lett. 12 3936–40

[77] Pizzocchero F, Gammelgaard L, Jessen B S, Caridad J M, Wang L, Hone J, Boggild P and Booth T J 2016 The hot pick-up technique for batch assembly of van der Waals heterostructures Nat. Commun. 7 11894

[78] Clark N, Nguyen L, Hamer M J, Schedin F, Lewis E A, Prestat E, Garner A, Cao Y, Zhu M and Kashiwaba R 2018 Scalable patterning of encapsulated black phosphorus Nano Lett. 18 5373–81

[79] Mashis Das P et al 2016 Controlled sculpture of black phosphorus nanoribbons ACS Nano 10 5687–93

[80] Blake P, Hill E W, Castro Neto A H, Novoselov K S, Jiang D, Yang R, Booth T J and Geim A K 2007 Making graphene visible Appl. Phys. Lett. 91 063124

[81] Reina A, Son H, Xiao L, Fan B, Dresselhaus M S, Liu Z and Kong J 2008 Transferring and identification of single- and few-layer graphene on arbitrary substrates J. Phys. Chem. C 112 17741–4

[82] Fauske V T et al 2015 HyperSpy: Multi-Dimensional Data Analysis Toolbox v0.81 (https://doi.org/10.5281/zenodo.27735)