LETTERS
PUBLISHED ONLINE: 6 FEBRUARY 2011 | DOI: 10.1038/NNANO.2010.283
nature nanotechnology

Fast DNA sequencing with a graphene-based nanochannel device

Seung Kyu Min‡, Woo Youn Kim†‡, Yeonchoo Cho and Kwang S. Kim*

Devices in which a single strand of DNA is threaded through a nanopore could be used to efficiently sequence DNA[1–9]. However, various issues will have to be resolved to make this approach practical, including controlling the DNA translocation rate, suppressing stochastic nucleobase motions, and resolving the signal overlap between different nucleobases[4,7]. Here, we demonstrate theoretically the feasibility of DNA sequencing using a fluidic nanochannel functionalized with a graphene nanoribbon. This approach involves deciphering the changes that occur in the conductance of the nanoribbon[10,11] as a result of its interactions with the nucleobases via π–π stacking[12,13]. We show that as a DNA strand passes through the nanochannel[14], the distinct conductance characteristics of the nanoribbon[15–17] (calculated using a method based on density functional theory coupled to non-equilibrium Green function theory[18–20]) allow the different nucleobases to be distinguished using a data-mining technique and a two-dimensional transient autocorrelation analysis. This fast and reliable DNA sequencing device should be experimentally feasible in the near future.

In spite of the significant promise of pioneering work on the development of rapid and inexpensive DNA sequencing[3–9], a practical and reliable method to guarantee single-base resolution has yet to be devised. The stochastic motion of a single-stranded DNA (ssDNA) through a nanopore lacks well-defined interactions between the nanopore and the nucleobases[4]. Orientational fluctuations of nucleobases cause overlaps of the current distributions between different bases[7], so it would be advantageous to develop a device that could hold each base firmly while the base was being interrogated. We therefore designed a hybrid device composed of a nanoscale fluidic channel functionalized by a graphene nanoribbon (GNR) that interacts with each nucleobase by means of a π–π interaction (Fig. 1). The GNR forms a bridge across a nanochannel[14] between opposing tip electrodes (in the same way in which a nanolens can be attached to a near-field scanning optical microscope tip[21]). The nanochannel provides a pathway along which an ssDNA can be driven by an external field, and the graphene controls the orientational fluctuation of individual bases by stacking them on its surface by means of the dispersion force between π-systems. The π–π interaction is sufficiently strong to sustain the stacking structures for current measurements, while being weak enough to allow the ssDNA to be translocated by an external force.

Here, we use a narrow GNR[15] that can be synthesized using a recently developed bottom-up approach. As a DNA passes through a nanochannel, each nucleobase interacts with the GNR via π–π stacking. The ballistic conductance of the GNR diminishes at specific energies corresponding to the characteristic π-molecular orbitals (MO) via the Fano resonance[22], which enables different bases to be distinguished. The conductance of graphene is highly sensitive to a small change on its surface due to its unrivalled high surface-to-volume ratio[23]. Indeed, such sensitivity allows the measurement of Hall conductance in graphene even at room temperature[24]. In contrast to narrow GNRs, carbon nanotubes[25] show more conductance channels as well as double to multiple stacking, which significantly reduces the characteristic perturbations of the electric current towards the noise level. This means that GNRs are superior to carbon nanotubes (see Supplementary Information).

To verify our design hypothesis, we performed comprehensive state-of-the-art transport simulations[18,19]. The trajectory of molecular structures for the transport calculations was obtained using molecular dynamics (MD) simulations at room temperature (the NAMD package[26]) for a ssDNA composed of eight bases (5′-GCATCGCT-3′) located with solvents and counterions in a nanochannel. The nanochannel was composed of silicon nitrate[14], and was equipped with a narrow GNR with a width of ~1 nm (Supplementary Movie)[15]. The ssDNA was dragged with a force of 30 kcal mol^{-1} Å^{-1} (~0.25 nN per base) to mimic an applied strong electric field. This strong force was chosen simply for the sake of computational efficiency, and provided a favourable outcome, but could also be reduced substantially (to an

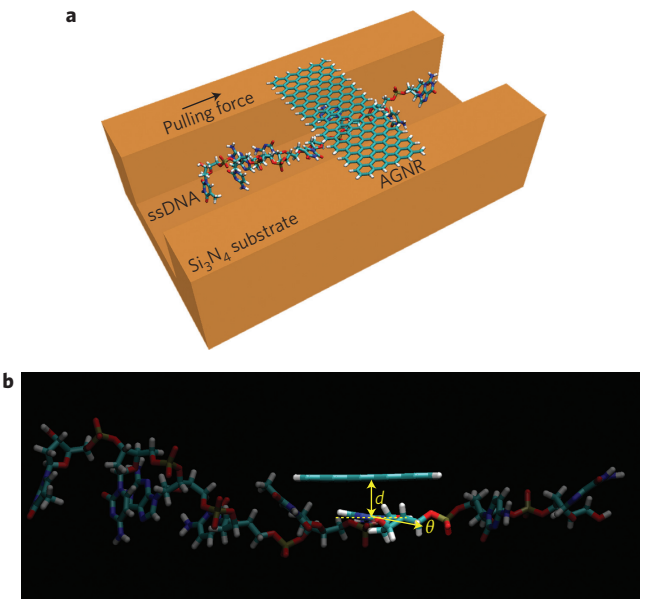

Figure 1 | DNA base stacking on a graphene nanodevice during its passage through a fluidic nanochannel. a, Schematic of a nanochannel device with an armchair GNR (AGNR) through which a ssDNA passes. The water molecules and counterions in the nanochannel are not depicted. **b**, Instantaneous snapshot from a simulation (d, stacking distance; θ, tilt angle).

Center for Superfunctional Materials, Department of Chemistry and Department of Physics, Pohang University of Science and Technology, Hyojadong, Namgu, Pohang 790-784, Korea; †Present address: Department of Chemistry, KAIST, Daejeon 305-701, Korea; ‡These authors contributed equally to this work. *e-mail: kim@postech.ac.kr

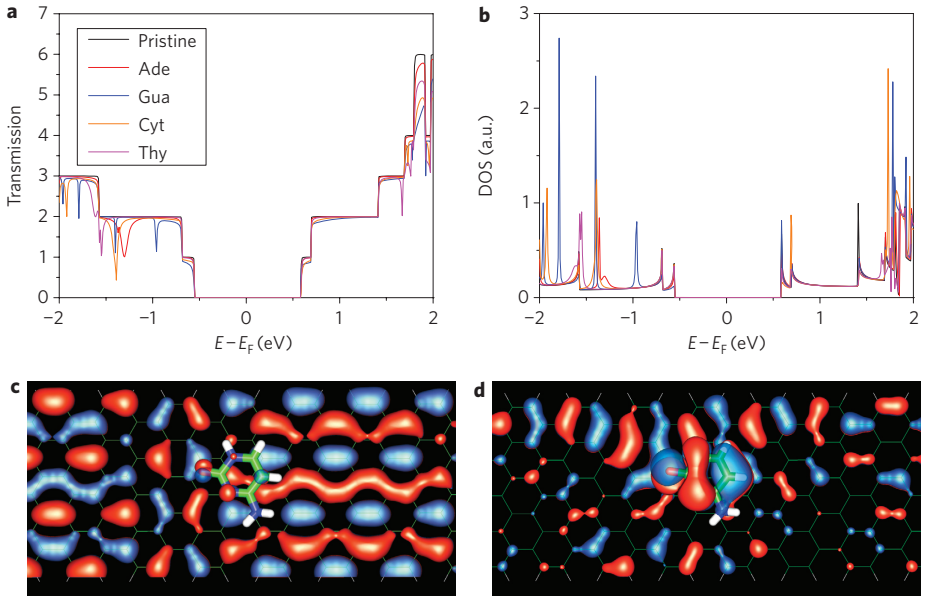

Figure 2 | Electronic and transport properties of a GNR device stacked optimally with DNA bases. a, Transmission curves (in G_0). **b**, Density of states. **c**,**d**, MOs of a GNR with stacked Cyt at energies ($E - E_F$) of −0.9 eV (**c**) and −1.4 eV (**d**). The sharp dip in transmission of Cyt at −1.4 eV occurs due to the strong Fano resonance between its MO and graphene (**d**).

experimental value of ∼0.25 pN per base[14]) to provide an increased and steadier base–GNR holding time for more reliable measurement. An analysis of structural changes confirmed that each base formed a stable stacking structure over the entire transit. Such π-stacking is much stronger than the H–π interactions between water molecules and graphene, so the stacking structures are not significantly influenced by solvent effects. The π-stacking energies (∼20 kcal mol^{-1})[27] are sufficiently strong to form stable stacked structures of nucleobases on graphene in water at room temperature.

With regard to ensuring experimental feasibility, we investigated how to control the sliding motion of the ssDNA on graphene. Sliding friction is uniform because of the non-local character of the dispersion force, as can be noted in the extremely smooth motion observed in carbon nanotubes[28,29]. The translational speed of a DNA strand in a nanochannel can be made steady by using another wide GNR (in addition to the GNR used to measure the currents) on which a number of bases are simultaneously stacked (frictional force is proportional to GNR width[30]). For practical experiments, the force used in the computations could be reduced by three orders of magnitude to an experimental value[14]. The frictional force of ssDNA on a large graphene sheet, which can be estimated from the frictional force between two neighbouring shells in a carbon nanotube[28], offsets a dragging force such as that exerted by an externally applied electric field. A DNA strand can then slide on a graphene surface smoothly and steadily with controlled speed. The negatively charged phosphate groups translocate the whole DNA against the direction of the electric field, while the counterions solvated in water move along the direction of the field.

To promote the identification of each base, we investigated the quantum conductance of the graphene-based device using our Postrans package[19]. A narrow armchair GNR with a width of ∼1 nm and a bandgap calculated to be 1.16 eV was considered. Figure 2a shows the transmission curves of the nucleobase–GNR complexes as a function of energy, $T(E)$. The pristine GNR shows integer transmission values corresponding to the number of conduction channels. However, the GNR with a stacked base presents a transmission curve with sharp drops at the characteristic MO energies of each base, due to the Fano resonance between the localized MOs of the nucleobase and the continuum π-band of the GNR (Fig. 2c,d), as can be noted from the conspicuous peaks in the density of states for the whole system (Fig. 2b). For example, the strong resonance between graphene and the Cyt MO at $E - E_F = -1.4$ eV (E_F is Fermi energy) partially blocks the ballistic conductance of the graphene (Fig. 2d).

In experiments, the transmission function can be replaced by conductance measured at a small bias as a function of gate voltage V_g:

$$G(V_g) = \frac{2e^2}{h} T(\mu)$$

where h is Planck's constant and $\mu = E_F - eV_g$ is the chemical potential adjusted by the gate voltage. Measuring conductance rather than current allows a clearer distinction of the change in the signal, because conductance through the GNR device changes by a unit of quantum conductance G_0 ($=2e^2/h$) in its interaction with DNA bases. For instance, the 50% relative reduction from $2G_0$ to $1G_0$ and the dip width of ∼30 meV in the conductance of this GNR device (see Supplementary Information) is experimentally measurable, being well above the noise level at room temperature. At each DNA translocation step, the two-dimensional conductance of the GNR has to be measured for a whole sweep of chemical potential, which is experimentally controlled via a backgate capacitively coupled to the GNR. The rate-determining step is then the voltage sweep. The transit time of each base should be controlled so as to be longer (∼10 μs) than the gate voltage sweeping time (∼0.1 μs, ∼10 MHz); this lies at the limit of present experimental capability. The transmission function in the energy range of interest is not affected significantly when the geometries of the bases on the GNR are changed, or in response to environmental effects (such as backbones, water, hydrated phosphate groups and counterions). This is because the other molecular groups have much larger energy gaps between the occupied and unoccupied MOs than nucleobases, and there are cancellation effects due to the similar background environments and the random motions of the environmental species.

The transmission data in Fig. 3a were obtained using density functional theory coupled to non-equilibrium Green's function (DFT-NEGF) method for the time-evolving series of base–GNR geometries, from the MD trajectories for the whole sequence of

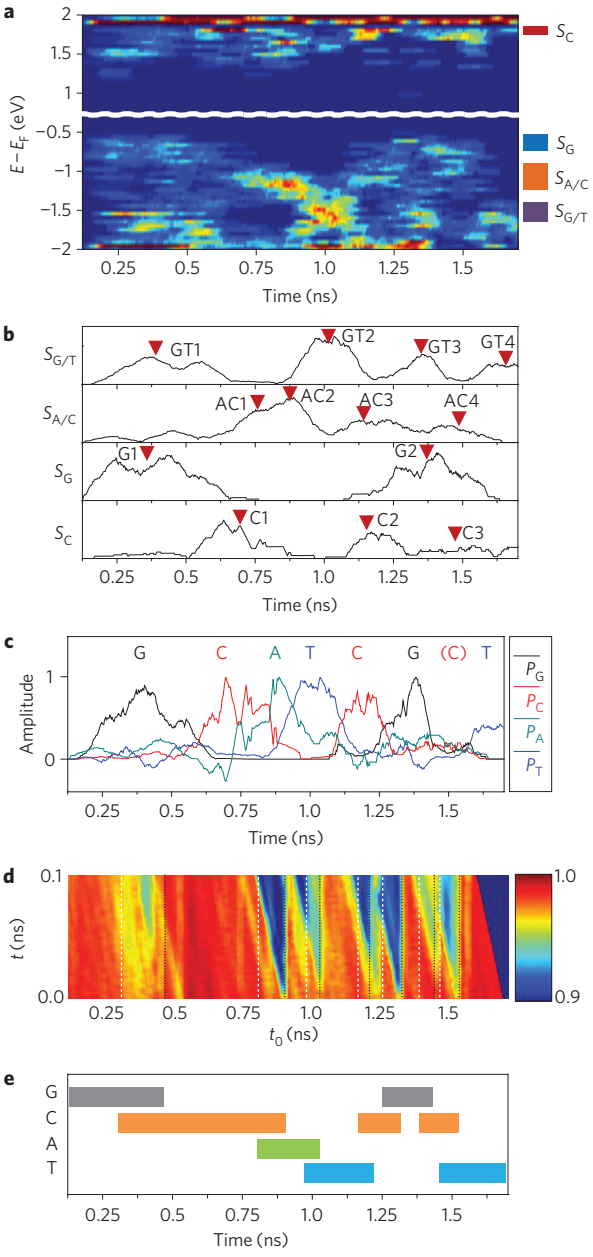

Figure 3 | Simulation results of the transport properties of 5′-GCATCGCT-3′. **a**, Time-dependent histogram of the transmission peak positions (red, maximal values; blue, minimal values). The histogram shows features of Cyt (S_C), Gua (S_G), Ade/Cyt ($S_{A/C}$) and Gua/Thy ($S_{G/T}$) corresponding to the band centred around $E - E_F = E_s = 1.8$, -0.65, -1.2 and -1.65 eV, respectively. The height of each box to the right of the histogram represents the energy range of integration for each S_{base}. **b**, Plots of $S_{base}(t)$ versus time as an ssDNA passes. The sequence GCATCGCT can be inferred from the four curves. **c**, Resolving of the sequence by $P_{base}(t)$, as a function of $S_{base}(t)$. **d**, Plot of two-dimensional TACF, $C(t, t_0; \tau = 0.025$ ns). The black dotted lines each indicate an instant at which a (subsequent) DNA base passes out; white dashed lines each indicate an instant at which a (subsequent) DNA base comes in. **e**, Final base sequence obtained from analysis of **c** and **d**. Colours and heights are given in relative intensity.

5′-GCATCGCT-3′ interacting with the GNR, including all surrounding water molecules and counterions in the nanochannel. To identify each base, we manipulated the transmission raw data using our data-mining technique. In the first step, we introduced a time-dependent histogram $H(E, t; \Delta)$ to represent the number of transmission drops appearing at a specific energy E within a tiny time interval Δ around time t:

$$H(E, t; \Delta) = \int_{t-\Delta}^{t} D(E, t')\mathrm{d}t'$$

where $D(E, t)$ is the number of peaks in the energy interval $(E, E + \varepsilon)$ at time t (ε is the histogram energy interval), and Δ characterizes the sampling size of the time data set.

However, the histogram in Fig. 3a, taken alone, is not sufficient to achieve single base resolution. For automatic sequencing, we devised a data-mining approach for more simplistic identification (see Methods). We then plotted the probability function projected onto each base, $P_x(t)(x = \mathrm{G/T/C/A})$ (Fig. 3b,c). In this way, each base is uniquely separated from the others, allowing single base resolution in the test sequencing for 5′-GCATCGCT-3′ (see Supplementary Information).

Much more reliable sequencing is achieved by clear recognition of the interaction interval of each individual base with the GNR, removing the possibility of the introduction of ambiguities if a series of bases of the same type pass beneath the GNR. To this end, we have developed a normalized two-dimensional transient autocorrelation function (TACF) for the given time span τ:

$$C(t, t_0; \tau) = \int_{t_0}^{t_0+\tau} \mathrm{d}t' \int_{-\infty}^{\infty} \mathrm{d}E\, J(E, t')J(E, t + t')/$$
$$\int_{t_0}^{t_0+\tau} \mathrm{d}t' \int_{-\infty}^{\infty} \mathrm{d}E\, [J(E, t')]^2$$

This gives the autocorrelation of a given event function $J(E, t)$ between two time intervals of duration τ separated by time t. Here, τ is smaller than the passage time of any single base. In the situation where the transient correlation between different bases is zero, the two-dimensional TACF shows a distinct triangular pattern when one base passes out and another base comes into the interaction region (Fig. 3d; Supplementary Fig. S2). This helps us determine the passage time of each base. Then, from Fig. 3c,d, base sequencing is possible (Fig. 3e).

In summary, we have reported an ultrasensitive graphene-embedded nanochannel device that effectively controls the motion of nucleobases via π–π interaction and deciphers the ultrasensitive characteristic Fano resonance-driven conductance of individual bases, one by one, in real time. In practice, the resolution of sequencing patterns will be much clearer as the DNA translocation slows down with a dragging force much weaker than that used in the calculations. Any further ambiguities can be resolved by either parallel or repetitive serial measurements, because the ssDNA is reusable. The experimental implementation of this device would allow ultrafast and reliable DNA sequencing, which is essential for collecting wide-ranging data related to differences/similarities between diverse DNA sequences in the field of bioinformatics.

Methods

The DFT-NEGF transport calculations[18–20] based on the local density approximation and using double-ζ plus polarization basis sets were performed in a fully self-consistent way. These calculations were carried out for all the time-evolving geometries and the entire sequence of the DNA interacting with the GNR, using room-temperature MD simulation trajectories including all surrounding water molecules and counterions in the nanochannel, which was composed of silicon nitrate and with a narrow GNR (width, ~1 nm).

For automatic sequencing analysis using the above electric conductance data, a data-mining approach was required. As specific drops in the transmission curves at 1.8, –0.9, –1.4 and –1.6 eV in the absence of water (Fig. 2a) indicated stacking of the corresponding nucleobases on the graphene, the histogram $H(E, t; \Delta)$ showed similar characteristic patterns in a realistic GNR–nanochannel device containing whole DNA in water with counterions (Fig. 3a). $H(E, t; \Delta)$ was integrated over a specified energy range $[E_s - \Gamma_s/2, E_s + \Gamma_s/2]$ to obtain the number of transmission

drops appearing within the given energy range and time interval Δ around t:

$$S(E_s, t; \Delta) = \int_{E_s-\Gamma_s/2}^{E_s+\Gamma_s/2} H(E, t; \tau)\mathrm{d}E$$

The energy values E_s (1.8, −0.65, −1.2 and −1.65 V for Cyt, Gua, Ade/Cyt and Gua/Thy, respectively) were chosen from the histogram in Fig. 3a. The small coloured bars at the right side of Fig. 3a show the ranges of Γ_s at a given energy E_s. To unravel some partial overlaps, we used the probability projecting each base. For example, the probability of Gua (Cyt) is obtained by projecting $S_{G,T}$ ($S_{A,C}$) into S_G (S_C):

$$P_G(t) = S_G S_{G,T};\ P_C(t) = S_C S_{A,C}$$

The probability of Thy (Ade) can then be obtained by projecting out P_G (P_T) from $S_{G,T}$ ($S_{A,C}$):

$$P_T(t) = S_{G,T} - P_G;\ P_A(t) = S_{A,C} - P_C$$

Thus, the data-mining technique, including two-dimensional TACF analysis of $C(t, t_0; \tau)$, for the time evolution of the conductance renders the automatic sequencing feasible.

Received 7 October 2010; accepted 22 December 2010; published online 6 February 2011

References

1. Collins, F. S. *et al.* A vision for the future of genomics research. *Nature* **422,** 835–847 (2003).
2. Service, R. F. The race for the $1000 genome. *Science* **311,** 1544–1546 (2006).
3. Shendure, J. & Ji, H. Next-generation DNA sequencing. *Nature Biotechnol.* **26,** 1135–1145 (2008).
4. Branton, D. *et al.* The potential and challenges of nanopore sequencing. *Nature Biotechnol.* **26,** 1146–1153 (2008).
5. Garaj, S. *et al.* Graphene as a subnanometre trans-electrode membrane. *Nature* **467,** 190–193 (2010).
6. Meller, A., Nivon, L. & Branton, D. Voltage-driven DNA translocations through a nanopore. *Phys. Rev. Lett.* **86,** 3435–3438 (2001).
7. Zwolak, M. & Di Ventra, M. Colloquium: physical approaches to DNA sequencing and detection. *Rev. Mod. Phys.* **80,** 141–165 (2008).
8. Lagerqvist, J., Zwolak, M. & Di Ventra, M. Fast DNA sequencing via transverse electronic transport. *Nano. Lett.* **6,** 779–782 (2006).
9. Zwolak, M & Di Ventra, M. Electronic signature of DNA nucleotides via transverse transport. *Nano. Lett.* **5,** 421–424 (2005).
10. Novoselov, K. S. *et al.* Electric field effect in atomically thin carbon film. *Science* **306,** 666–669 (2004).
11. Kim, K. S. *et al.* Large-scale pattern growth of graphene films for stretchable transparent electrodes. *Nature* **457,** 706–710 (2009).
12. Kim, K. S., Tarakeshwar, P. & Lee, J. Y. Molecular clusters of π-systems: theoretical studies of structures, spectra and origin of interaction energies. *Chem. Rev.* **100,** 4145–4185 (2000).
13. Lee, E. C. *et al.* Understanding of assembly phenomena by aromatic–aromatic interactions: benzene dimer and the substituted systems. *J. Phys. Chem. A* **111,** 3446–3457 (2007).
14. Liang, X. & Chou, S. Y. Nanogap detector inside nanofluidic channel for fast real-time label-free DNA analysis. *Nano Lett.* **8,** 1472–1476 (2008).
15. Cai, J. *et al.* Atomically precise bottom-up fabrication of graphene nanoribbon. *Nature* **466,** 470–473 (2010).
16. Kosynkin, D. V. *et al.* Longitudinal unzipping of carbon nanotubes to form graphene nanoribbons. *Nature* **458,** 872–876 (2009).
17. Jiao, L., Zhang, L., Wang, X., Diankov, G. & Dai, H. Narrow graphene nanoribbons from carbon nanotubes. *Nature* **458,** 877–880 (2009).
18. Kim, W. Y. & Kim, K. S. Tuning molecular orbitals in molecular electronics and spintronics. *Acc. Chem. Res.* **43,** 111–120 (2010).
19. Kim, W. Y. & Kim, K. S. Prediction of very large values of magnetoresistance in a graphene nanoribbon device. *Nature Nanotech.* **3,** 408–412 (2008).
20. Kim, W. Y. & Kim, K. S. Carbon nanotube, graphene, nanowire, and molecule based electron and spin transport phenomena using the nonequilibrium Green's function method at the level of first principles theory. *J. Comput. Chem.* **29,** 1073–1083 (2008).
21. Lee, J. Y. *et al.* Near-field focusing and magnification through self-assembled nanoscale spherical lenses. *Nature* **460,** 498–501 (2009).
22. Fano, U. Effects of configuration interaction on intensities and phase shifts. *Phys. Rev.* **124,** 1866–1878 (1961).
23. Schedin, F. *et al.* Detection of individual gas molecules adsorbed on graphene. *Nature Mater.* **6,** 652–655 (2007).
24. Novoselov, K. S. *et al.* Room-temperature quantum hall effect in graphene, *Science* **315,** 1329 (2007).
25. Meng, S. *et al.* DNA nucleoside interaction and identification with carbon nanotubes. *Nano Lett.* **7,** 45–50 (2007).
26. Phillips, J. C. *et al.* Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26,** 1781–1802 (2005).
27. Antony, J. & Grimme, S. Structures and interaction energies of stacked graphene–nucleobase complexes. *Phys. Chem. Chem. Phys.* **10,** 2722–2729 (2008).
28. Cummings, J. & Zettl, A. Low-friction nanoscale linear bearing realized from multiwall carbon nanotubes. *Science* **289,** 602–605 (2000).
29. Hong, B. H. *et al.* Extracting subnanometer single shells from ultralong multi-walled carbon nanotubes. *Proc. Natl Acad. Sci. USA* **102,** 14155–14158 (2005).
30. Mo, Y., Turner, K. T. & Szlufarska, I. Friction laws at the nanoscale. *Nature* **457,** 1116–1119 (2009).

Acknowledgements

This work was supported by the National Research Foundation (National Honor Scientist program: 2010-0020414, WCU:R32-2008-000-10180-0, EPB Center: 2009-0063312, GRL) and KISTI (KSC-2008-K08-0002). The authors thank D. R. Mason and N. Kim for discussions.

Author contributions

S.K.M. and W.Y.K. worked together. Y.C. assisted in the calculations and analysis. K.S.K. supervised the project. S.K.M., W.Y.K. and K.S.K. wrote the paper together.

Additional information

The authors declare no competing financial interests. Supplementary information accompanies this paper at www.nature.com/naturenanotechnology. Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/. Correspondence and requests for materials should be addressed to K.S.K.