Published in final edited form as:

Nat Nanotechnol. 2010 December ; 5(12): 868–873. doi:10.1038/nnano.2010.213.

Single base resolution in tunneling reads of DNA composition

Shuo Huang^{1,2}, Jin He¹, Shuai Chang^{1,2}, Peiming Zhang¹, Feng Liang¹, Shengqin Li¹, Michael Tuchband¹, Alexander Fuhrman², Robert Ros², and Stuart Lindsay^{1,2,3}

¹Biodesign Institute, Arizona State University, Tempe, AZ 85287.

²Department of Physics, Arizona State University, Tempe, AZ 85287.

³Department of Chemistry and Biochemistry, Arizona State University, Tempe, AZ 85287.

Abstract

Single-molecule DNA sequencing based on measuring the physical properties of bases as they pass through a nanopore^{1,2} eliminates the need for the enzymes and reagents used in other approaches. Theoretical calculations indicate that electron tunneling could identify bases in single-stranded DNA, yielding long reads and eliminating enzymatic processing.^{3–5} It was shown recently that tunneling can sense individual nucleotides⁶ and nucleosides.⁷ Here, we show that tunneling electrodes functionalized with recognition reagents can identify a single base flanked by other bases in a short DNA oligomer. The residence time of a single base in a recognition junction is on the order of a second, but pulling the DNA through the junction with a force of tens of piconewtons would yield reading speeds of tens of bases per second.

Changes in the ion current through a nanopore can be used to identify translocating nucleotides. This opens the way to DNA sequencing if an exonuclease can pass each cleaved nucleotide into the pore sequentially.⁸ As an alternative, it has been proposed that the high spatial resolution of electron tunneling would allow direct reading of bases in an intact DNA polymer.^{3–5} Recent progress in measuring electron tunneling through nucleotides or nucleosides shows that they can be identified by means of characteristic current signals.^{6,7} Recognition tunneling^{7,9} is an approach in which electrodes are functionalized with reagents that bind the target DNA bases. Contact via molecular adsorbates has been used to produce extraordinarily high spatial resolution in atomic force microscopy¹⁰ and, as we show here, single bases can be resolved in a DNA polymer when read by means of a selective chemical contact.

Correspondence and requests for material should be addressed to SL.

Web Summary: Electron tunneling via functionalized electrodes can resolve and identify a single DNA base embedded in an oligomer.

Author Contributions

SH, SC and JH carried out tunneling measurements and characterized the samples. PZ, FL and Sq. L designed, synthesized and characterized reagents. MT prepared tunneling probes. AF and RR carried out force spectroscopy. SL designed experiments, analyzed data and wrote the paper.

Competing Financial Interests

SL, PZ and JH are named as inventors in patent applications.

To extend recognition tunneling to reads in buffered aqueous electrolyte, we synthesized the reagent 4-mercaptobenzamide (Fig. 1a and Methods) which presents two hydrogen-bond donor sites (on the nitrogen) and one hydrogen-bond acceptor site (the carbonyl). Likely binding modes to the four bases are shown in Fig. 2a.⁷ A gold (111) substrate and a partially-insulated gold STM probe were functionalized with this reagent (Methods and online supporting information) and characterized in an electron tunneling junction formed in a scanning tunneling microscope (PicoSPM, Agilent, Chandler, AZ). Fig. 1a shows a d(CCACC) oligomer trapped in a tunnel gap through hydrogen bonding to one mercaptobenzamide molecule on the probe and another on the substrate. In reality, the oligomer is probably held by many contacts, but only those that complete a short tunneling path (highlighted) will contribute significantly to the current. In our measurements, the probe is not deliberately scanned, but moves over the substrate as the microscope drifts. Alternatively, molecules may diffuse through the gap. Characteristic bursts of current are observed, and an example is shown in Fig. 1b. As we show below, the low frequency, large amplitude pulses indicate a C, while the high frequency, small amplitude pulses signal an A. Fig. 1c shows a sliding average of the spike amplitudes – values below the red line identify an A base unambiguously. Figure 1d shows a sliding average over the pulse frequencies (as defined for each adjacent pair of spikes) – the low frequency regions at each end enhance the confidence with which those regions can be assigned to a C base. The probability of an assignment to A (red line) or C (blue line) is shown in Fig. 1e. Calculation of these probabilities is based on our study of nucleotides, homopolymers and heteropolymers as described below. This example clearly shows that a single A base can be identified with high confidence when flanked by C bases in an intact DNA molecule.

We first characterized the tunnel gap using doubly-distilled water and 0.1 mM phosphate buffer (PB – pH=7.4). Small signals were observed from buffer alone with bare electrodes, but they were much rarer when both electrodes were functionalized and the tunnel gap conductance set to 20 pS or less. (Fig. 2b and online supporting information). The tunnel decay was much more rapid (decay constant, $\beta = 14.2 \pm 3.2 \text{ nm}^{-1}$) with both electrodes functionalized than is the case in water alone ($\beta \sim 6.1 \pm 0.7 \text{ nm}^{-1}$ – ¹¹ and online supporting information) and we estimate that the tunnel gap at $i=10 \text{ pA}$ and $V = +0.5 \text{ V}$ is a little over the length of two benzamide molecules (i.e. a little greater than 2 nm).

Introducing DNA nucleotides (10 μM in PB) into the tunnel gap yielded characteristic noise spikes as shown in Figs. 2c–f. The signal count rate (defined in Fig. 2k) varied considerably from 25 counts/s (5-methyl-deoxycytidine 5'-monophosphate, d^mCMP) to less than 1 c/s (deoxycytidine 5'-monophosphate, dCMP). No signals were recorded at all with thymidine 5'-monophosphate (dTMP), the signal looking exactly like the control (Fig. 2b). STM images suggest that this nucleotide binds to the surface (and presumably the probe) very strongly, blocking interactions in which a single molecule spans the junction.

The current occurs in bursts of spikes (longer signal runs are given in online supporting information) and distributions of the spike heights were quite well fitted with two Gaussians distributions of the logarithm of current⁷ as shown in Figs 2 g–j (fitting parameters are given

Supplementary Information accompanies this paper.

in the online supporting information). These histograms were generated by counting only pulses that exceeded $1.5\times$ the SD of the local noise background – i.e., typically pulses above 6 pA (a full description of the analysis procedure is given by Chang et al.⁷).

dCMP generates the highest signals and the lowest count rate while deoxyadenosine 5'-monophosphate (dAMP) and d^mCMP produce the smallest signals and the highest count rate (we found little difference between cytidine and 5-methylcytidine in organic solvent⁷ – supporting online information). The three bases with narrower pulse height distributions (dAMP, d^mCMP and GMP) often show bursts of “telegraph-noise” characteristic of sources that fluctuate between two levels⁹ (particularly marked for dAMP). Such a two-level distribution is a strong indication that the tunneling signals are generated by a single molecule trapped in the tunnel junction.⁹ The characteristics of the tunneling noise from the nucleotides are summarized in Table 1.

dAMP signals are well-separated from dCMP signals, and d^mCMP signals are well separated from dCMP signals in spike amplitude and in the time distribution of their signals (Table 1 and online supporting information). For this reason, we chose to investigate DNA oligomers composed of A, C and ^mC bases.

Figs. 3a,c and e show representative tunneling noise traces for d(A)₅, d(C)₅ and d(^mC)₅ with the corresponding current peak distributions shown in Figs. 3b, d and f. Comparing Fig. 3b (d(A)₅) with Fig. 2g (dAMP), Fig. 3d (d(C)₅) with Fig. 2h (dCMP) and Fig. 3f (d(^mC)₅) with Fig. 2i (d^mCMP) leads to the following startling conclusion: *most of the polymer binding events in the tunnel junction generate signals that resemble those generated by single nucleotides*. That this should be so is not obvious. It requires (1) that single bases are being read and (2) that steric constraints owing to the polymer backbone do not prevent base-binding events from dominating the signals.

There are some (small) differences between nucleotide and oligomer signals: (1) Peak positions, widths and relative intensities are altered somewhat (see online supporting information for details of the fits, and the also the nucleotide distributions which have been replotted on top of the homopolymer distributions as the black lines on Figs 3b,d and f.). (2) Almost all of the signals generated by nucleotides are less than 0.1 nA at 0.5V bias (Table 1). In contrast, 20% of the total signals generated by d(A)₅ and d(^mC)₅ are larger than 0.1 nA at this bias (Table 2 - this is not obvious in Figure 2 where distributions are plotted only up to 0.1nA – the high current regions are shown in the online supporting information). These high current (>0.1 nA) features in d(A)₅ and d(C)₅ are continuously distributed so they do not represent parallel reads of more than one base at a time (where currents would be distributed in multiples of the single molecule values¹²). Rather, they are new features associated with the presence of the polymeric structure in the tunnel gap. Such a non-specific, large amplitude spike is labeled by an asterisk in Fig. 1b.

Features at $I > 0.1$ nA appear much less frequently in oligomers of mixed sequence, suggesting that they are associated with base-stacking in the homopolymers. Fig. 3h shows a current distribution for d(ACACA) where 95% of events are below 0.1 nA. Fig. 3j shows a current distribution for d(C^mCC^mCC) where 99% of events are below 0.1 nA. The solid red

lines are the sums of the distributions measured for the homopolymers corresponding to the constituents with, scaling aside, only one fitting parameter. This parameter is the ratio, r_{fit} , of the A/C ($r_{\text{fit}}=0.48$) or $^m\text{C}/\text{C}$ ($r_{\text{fit}}=0.66$) contributions. These values differ from the known composition ratios (0.6 for ACACA and 0.4 for $\text{C}^m\text{CC}^m\text{CC}$) but are surprising in as much as the spike rate for dCMP alone is very small, yet C appears to be quite well represented in the mixed sequence oligomer data. This suggests that Cs surrounded by As are read more frequently, possibly because the C-containing oligomer is better attached to the substrate than the isolated dCMP.

Most importantly, mixed oligomers generate signals that are largely described as the sum of the individual base signals. (Some intermediate current reads, labeled “1” in Figs 3h and j, and a small number of additional high current features – labeled “2” – show that sequence context plays a small role.)

In these experiments, the probe drifts randomly over the samples, so the sequence is not “read” deterministically. Nonetheless we can readily find traces in which the signals alternate between “A-like” and “C-like” (Fig. 3g) and “ ^mC -like” and “C-like” (Fig 3i). The duration of these “bursts” (see Fig. 2k) of signals is long ($0.14 \pm 0.02\text{s}$ in ACACA and $0.15 \pm 0.02\text{s}$ in $\text{C}^m\text{CC}^m\text{CC}$). Similar bursts are seen in the homopolymers (Table 2) and the nucleotides (Table 1). This leads us to our second unexpected conclusion: *the lifetime of the bound complex in the tunnel gap is very long* (fraction of a second) compared to either the interval between noise spikes (ms) or the lifetime of the bound-state in solution (very short – online supporting information).

We have used dynamic force spectroscopy as an independent test of the unexpectedly long lifetime of the benzamide-base-benzamide complex confined to a nanoscale gap. In these measurements (Fig. 4a) one of the recognition molecules was bound to an AFM probe via a 34 nm long polyethyleneglycol (PEG) linker (Methods) while the other formed a monolayer on an Au(111) substrate. dAMP was used as the target analyte to bridge the gap. In the absence of dAMP, adhesion between probe and substrate was extremely small, presumably because the hydrogen bonding sites on the benzamide recognition molecules were stably bound by water molecules. Adhesion features were observed in the presence of a small amount of dAMP, falling as the concentration of dAMP increased (resulting in binding of both probe and substrate by dAMP - online supporting information). Stretching of the PEG tether generated a characteristic signal that permitted multiple binding events (Fig 4b (i)) to be separated from single molecule events (Fig. 4b (ii)) so that only single-molecule bond-breaking events were analyzed.¹³ Single molecule bond-breaking forces as a function of pulling speed are summarized in Fig. 4c (solid lines are maximum likelihood fits to a heterogeneous bond model^{13,14}) and the bond survival probability as a function of bond-breaking force is shown in Fig. 4d. The solid lines are fits to the same heterogeneous bond model.¹⁴ They yield an off-rate at zero force, $K_{\text{off}}^0=0.28\text{ s}^{-1}$. *Thus the intrinsic (zero-force) survival time of this complex is on the order of seconds, not milliseconds.* The analysis also yields the distance to the transition state for dissociation, $\alpha = 0.78\text{ nm}$ (as well as its variance, $\sigma=0.19\text{ nm}$). We conclude that each base resides in the tunnel junction for a significant fraction of a second, while generating tunneling signals at kHz rates. Thus the

entire cluster of signals that occur in one burst (burst durations are listed in Tables 1 and 2) can be used to characterize a base.

Long-bound-state life-times accompanied by rapid fluctuations in electronic signatures have been reported previously in STM images¹⁵ and in the effect of single-molecule reactions on transport in carbon nanotubes.¹⁶ The origin of this noise is unclear, save that it appears to be very temperature sensitive, indicative of small energy barriers to the motion that causes the noise.¹⁵ Following Goldsmith et al.¹⁶ we have analyzed (online supporting information) the distribution of “on” and “off” times (see Fig. 2k). In a limited time range of times, determined by the amplifier response at one end, and the servo response time at the other,⁷ these distributions are exponential (as expected for a Poisson source) and the 1/e times (τ_{on} and τ_{off}) are listed in Tables 1 and 2. They do not differ much, and calculating an energy difference, G , between the on and off states from $G = kT_B \ln(\tau_{\text{off}}/\tau_{\text{on}})$ yields the values listed in the Tables (in units of thermal energy, $k_B T$, at 300K). These values are all a fraction of $k_B T$. Thus the “switching” cannot represent thermal activation over a significant barrier (the normal source of two-level noise). One possible explanation is Brownian motion in a bound state sampled by an exponentially-sensitive matrix element (online supporting information).

The “on” and “off” times are so broadly distributed that they are not very useful for identifying base-signals. However, the frequency within a burst (f_S Tables 1 and 2) is a much simpler parameter. Figs S19a and S19b show the current distributions and frequency distributions for the three homopolymers, normalized so that the area under each curve is unity. The frequency distribution for $d(^m\text{C})_5$ is bimodal, with many reads in the “C” frequency range and a number at the very fast rate (ca. 1300 Hz) observed for ^mC MP alone (labeled $f(^m\text{CMP})$ on the figure). This suggests that the binding modes of ^mC are altered significantly in a polymer context (consistent with the larger shift of the polymer signal compared to the nucleotide signal, Fig. 3f) so we chose to analyze oligomers containing A and C, in particular the $d(\text{CCACC})$ sequence shown in Fig. 1a.

Given an average current in a burst, $\langle i \rangle$ and frequency, $\langle f \rangle$, the distributions shown in Fig. S20, $I_{A,C}(\langle i \rangle)$ (Fig. S20a) and $F_{A,C}(\langle f \rangle)$ (Fig. S20b) determine independent probabilities

that a base is an A or a C: $P_{A,C}^i = \frac{I_{A,C}(\langle i \rangle)}{I_A(\langle i \rangle) + I_C(\langle i \rangle)}$ and $P_{A,C}^f = \frac{F_{A,C}(\langle f \rangle)}{F_A(\langle f \rangle) + F_C(\langle f \rangle)}$. The current distribution from $d(\text{CCACC})$ (inset in Fig. S9a) is almost completely dominated by A spikes (the component of the C distribution in this fit is 7% or less). This is a surprising result, that more C's in the sequence give a smaller number of C spikes. But it is consistent with our hypothesis that the frequency of C reads is increased when the base is flanked by A's (c.f. the increase in C reads in $d(\text{ACACA})$ compared to the $d\text{CMP}$ vs. $d\text{AMP}$ count rate).

Armed with our analysis of the burst signals, we can now make quantitative assignments of mixed signals (this was done “by eye” in Figs 3g and i). $d(\text{C})_5$ produces no signals below 0.015 nA, so bursts of current below this level (but above the noise) can be unambiguously assigned to A. For larger amplitude signals we use both the frequency and amplitude data as described in the supporting information. The result is the pair of curves shown in Fig. 1e.

Using this approach to sequence DNA requires several further developments. Firstly, the polymer must be pulled through a tunnel junction at a controlled speed, particularly if homopolymer runs are to be read. Since DNA passes through unfunctionalized nanopores too rapidly to be read² the long residence time of bases in a functionalized tunnel junction is an asset. At present, movement from one site to another is driven by uncontrolled mechanical drift that generates unknown forces on the reading complex. Our force spectroscopy data can be used to give a crude estimate of the “pulling” force that would be needed to achieve a given read rate (assuming the measured off-rate for dAMP to be representative for all bases). The Bell equation gives the off rate at a force F as

$K_{off} = K_{off}^0 \exp\left(\frac{F\alpha}{k_B T}\right)$ so, with $K_{off}^0 = 0.28 \text{ s}^{-1}$ and $\alpha = 0.78 \text{ nm}$, 19 pN would result in passage of 10 bases per second. A rate of 10 bases s^{-1} gives about 30 data spikes (on average) for a “C” read, enough to generate an assignment with a reasonable level of confidence. A force of 19 pN can be generated by a bias of just 80 mV across a nanopore¹⁷ so read rates of 10 bases per second per tunnel junction seem feasible.

The second requirement for a practical sequencing system is a better recognition chemistry in which there is a much larger separation of the current distributions from all five bases. New compounds are presently under study in our lab.

Methods

Nucleoside 5'-monophosphates (from Sigma-Aldrich) were used as supplied. DNA oligomers were synthesized and characterized by IDT and used without further purification. Synthesis and characterization of other materials are described in online supporting information. Gold probes were etched as described previously⁷ and coated with high-density polyethylene¹⁸ to leave a fraction of a micron of exposed gold (optical and TEM characterization is described in online supporting information). These probes gave no measureable DC leakage, important as this can be a source of distortion of the tunneling signal.⁷ Capacitive coupling of 120 Hz switching signals was a problem (Fig. 2b) minimized by careful control of the coating profile. The gaps were characterized by recording current decay curves as a function of distance starting at 20 pS, a distance that gave no signals in buffer alone with functionalized electrodes. Current signals were recorded using an Agilent PicoSPM together with a digital oscilloscope controlled by a custom Labview program. The servo response time was set to about 30 ms as described previously.⁷ This places an upper limit on undistorted measurements of pulse widths of a few ms. Analysis of current distributions was automated using the software described elsewhere.⁷ Force spectroscopy was carried out with a MFP3D AFM (Asylum Research, Santa Barbara). Heterobifunctional PEG linkers (MAL-SVA 3400 from Lysan Bio) of 34 nm extended length were attached at one end to silicon nitride AFM probe (Veeco MSNL - spring constant = 0.02 N/m) and mercapto-benzamide molecules attached to the remaining maleimide as described elsewhere¹⁹. Force curves were taken in 1 mM PB buffer with an initial 10 μM concentration of dAMP in the gap adjusted by rinsing (online supporting information). Force curves were analyzed using custom software.²⁰

Acknowledgments

We acknowledge useful discussions with Otto Sankey, Predrag Krstic and Brett Gyarfus. Phil Collins made helpful comments on an earlier version of this manuscript. Hao Liu composed the graphic for Fig. 1a. This work was supported by a grant from the Sequencing Technology Program of the National Human Genome Research Institute (HG004378). RR and AF were supported by a grant from the National Cancer Institute (U54CA143682).

References

1. Zwolak M, Di Ventra M. Physical approaches to DNA sequencing and detection. *Reviews of Modern Physics*. 2008; 80:141–165.
2. Branton D, et al. Nanopore Sequencing. *Nature Biotechnology*. 2008; 26:1146–1153.
3. Lagerqvist J, Zwolak M, Di Ventra M. Influence of the Environment and Probes on Rapid DNA Sequencing via Transverse Electronic Transport. *Biophys J*. 2007; 93:2384–2390. [PubMed: 17526560]
4. Zwolak M, Di Ventra M. Electronic Signature of DNA Nucleotides via Transverse Transport. *Nano Lett*. 2005; 5:421–424. [PubMed: 15755087]
5. Krstic PS, Wells JC, Fuentes-Cabrera M, Xu D, Lee JW. Toward Electronic Conductance Characterization of DNA Nucleotide Bases. *Solid State Phenomena*. 2007; 121–123:1387–1390.
6. Tsutsui M, Taniguchi M, Yokota K, Kawai T. Identification of Single Nucleotide via Tunnelling Current. *Nature Nanotechnology*. 2010; 5:286–290.
7. Chang S, Huang S, He J, Liang F, Zhang P, Li S, Chen X, Sankey OF, Lindsay SM. Electronic Signature of all four DNA Nucleosides in a Tunneling Gap. *Nano Letters*. 2010; 10:1070–1075. [PubMed: 20141183]
8. Clarke J, Wu H-C, Jayasinghe L, Patel A, Reid S, Bayley H. Continuous base identification for single-molecule nanopore DNA sequencing. *Nature Nanotechnology*. 2009; 4:265–270.
9. Lindsay S, He J, Sankey O, Hapala P, Jelinek P, Zhang P, Chang S, Huang S. Recognition Tunneling. *Nanotechnology*. 2010; 21:262001–262013. [PubMed: 20522930]
10. Gross L, Mohn F, Moll N, Liljeroth P, Meyer G. The Chemical Structure of a Molecule Resolved by Atomic Force Microscopy. *Science*. 2009; 325:1110–1114. [PubMed: 19713523]
11. Vaught A, Jing TW, Lindsay SM. Non-exponential tunneling in water near an electrode. *Chemical Physics Letters*. 1995; 236:306–310.
12. Lindsay SM, Ratner MA. Molecular Transport Junctions: Clearing Mists. *Advanced Materials*. 2007; 19:23–31.
13. Fuhrmann A, Anselmetti D, Ros R, Getfert S, Reimann P. Refined procedure of evaluating experimental single-molecule force spectroscopy data. *Phys Rev. E*. 2008; 77:03912–03922.
14. Getfert S, Reimann P. Optimal evaluation of single-molecule force spectroscopy experiments. *Phys Rev E*. 2007; 76:052901–052905.
15. Ramachandran GK, Hopson TJ, Rawlett AM, Nagahara LA, Primak A, Lindsay SM. A Bond-Fluctuation Mechanism for Stochastic Switching in Wired Molecules. *Science*. 2003; 300:1413–1415. [PubMed: 12775835]
16. Goldsmith BR, Coroneus JG, Kane AA, Weiss GA, Collins PG. Monitoring single-molecule reactivity on a carbon nanotube. *Nano Letters*. 2008; 8(1):189–194. [PubMed: 18088152]
17. Keyser UF, Koelman BN, van Dorp S, Krapf D, Smeets RMM, Lemay SG, Dekker NH, Dekker C. Direct force measurements on DNA in a solid-state nanopore. *Nature Physics*. 2006; 2:473–477.
18. Visoly-Fisher I, Daie K, Terazono Y, Herrero C, Fungo F, Otero L, Durantini E, Silber JJ, Sereno L, Gust D, Moore TA, Moore AL, Lindsay SM. Conductance of a biomolecular wire. *Proc. Nat. Acad. Sci*. 2006; 103:8686–8690. [PubMed: 16728508]
19. Ashcroft B, Spadola Q, Qamar S, Zhang P, Kada G, Bension R, Lindsay SM. An AFM/Rotaxane Molecular Reading Head for Sequence-Dependent DNA Structure. *Small*. 2008; 4:1468–1475. [PubMed: 18680093]
20. Fuhrmann, A. PhD thesis in Physics. Arizona State University; 2010.

Figure 1.

Reading a single base within a heteropolymer. (a) Benzamide groups on the probe and substrate bind bases in the polymer to give a signal dominated by the shortest tunneling path (highlighted for the connection to the single A in d(CCACC)). (b) Characteristic bursts of tunneling noise with large, infrequent spikes signaling C and smaller, more frequent spikes signaling A. (Background tunnel current is 10 pA, bias + 0.5V). The spike labeled * (off-scale at 0.11 nA) is non specific, and rejected from the analysis. (c) Rolling average of the spike height (0.25s window, 0.125 s steps) and spike frequency (d). C bases generate a negligible number of spikes below 0.015 nA (red line). (e) Probability that the signal comes from a A (shown by the red line on panel c) or a C (blue line). For signal amplitudes >0.015 nA, probabilities are calculated as described in the supporting information (values are not normalized to add to 1 in these regions). This burst of signal was chosen to show a clear example of transitions between A and C bases. Longer time traces (Fig. S9a are dominated by signals from A's which are preferentially trapped in the junction.

Figure 2.

Tunneling signals from nucleotides trapped in a functionalized tunnel gap. (a) Proposed hydrogen bonding modes for all four bases. In practice water must play a role because the observed difference between C and ^{5me}C would not be accounted for by these structures alone. In phosphate buffered saline, but in the absence of analyte, a 20 pS gap ($i = 10$ pA, $V = +0.5$ V) gave a signal free of features, except for some AC coupled line-noise pointed by arrows in (b). (c) – (f) Characteristic current spikes produced when nucleotides dAMP, dCMP, d^mCMP and dGMP were introduced (longer signal runs are given in the online supporting information). dTMP gave no signals. (g) – (j) corresponding distribution of pulse heights. Red lines are fits to two Gaussian distributions in the logarithm of current. (k)

Definition of the parameters used to characterize the tunneling signals. Spikes are counted if they exceed a threshold equal to $1.5 \times$ the standard deviation of the noise on the local background. The signals occur in bursts (duration T_B , frequency f_B) each containing current spikes at a frequency f_S . The spikes stay high for a period t_{on} and low for a period t_{off} . The total count rate (inset in g–j) is the number of spikes in all bursts divided by the measurement time.

Figure 3.

Tunneling signal distributions from oligomers resemble those of the constituent nucleotides. (a, c, e) Representative current traces from $d(A)_5$, $d(C)_5$ and $d(^mC)_5$ with the corresponding distributions shown in b, d and f. Red lines are fits with parameters similar to those used for nucleotides (online supporting information). The black lines are the fits to the corresponding nucleotide distributions shown in Figs. 2 g,h and i. “HCT” labels some of the high current features seen in homopolymers but not nucleotides (b and f). (g and i) Current traces from mixed oligomers, $d(ACACA)$ and $d(C^mCC^mCC)$, with corresponding current distributions

(h and j). Red lines are scaled homopolymer fits, with the green-dashed line showing the “C” contribution, the orange-dashed line showing the “A” contribution and the purple-dashed line the ^mC contribution. The data are well described by the homopolymer parameters though some intermediate signals (“1”) and new high current features (2) show that the sequence context affects the reads a little. The colored bars on the current traces mark bursts of A-like signals (orange), C-like signals (green) and ^mC-like signals (purple).

Figure 4.

The lifetime of the reading complex is on the order of a second at zero force. (a) AFM gap functionalization where the blue line represents a 34 nm PEG linker. (b) representative force curves showing: (i) Pulling on more than one molecule at a time – the force baseline is not restored after each break, and the z-extension (corrected for tip displacement) is > 34 nm. (ii) A single molecule curve of the type accepted by the software. The force returns to the baseline after the bond breaks and the corrected extension is ~ 34 nm. (c) Histograms of bond breaking forces at the pulling speeds marked. The solid lines are maximum likelihood fits to the heterogeneous bond model. (d) Bond survival probability plotted versus bond breaking force for the four pulling speeds, fitted by the same heterogeneous bond model parameters (solid lines). These fits yield a zero-force off rate of 0.28 s^{-1} implying that the assembly lives for times on the order of seconds in a nanogap, much longer than the lifetime in solution. For details see ref. 14.

Table 1

Nucleotide tunneling noise characteristics. Parameters are defined in Figure 2k.

Nucleotide	dAMP	dGMP	dCMP	d ^m CMP
Burst Duration (T_B , s)	$0.19 \pm 0.05^*$	$0.13 \pm 0.02^*$	$0.12 \pm 0.02^*$	$0.06 \pm 0.01^*$
Burst Frequency (f_B , Hz)	$732 \pm 82^{\S}$	$574 \pm 67^{\S}$	$306 \pm 23^{\S}$	$1305 \pm 100^{\S}$
Fraction of reads > 0.1 nA	0.02	0.001	0.02	0.01
τ_{on} (ms)	$0.38 \pm 0.01^*$	$0.48 \pm 0.02^*$	$0.42 \pm 0.02^*$	$0.31 \pm 0.09^*$
τ_{off} (ms)	$0.35 \pm 0.01^*$	$0.56 \pm 0.04^*$	$0.71 \pm 0.06^*$	$0.41 \pm 0.11^*$
τ_{on}/τ_{off}	~ 1	0.9	0.6	0.8
G (kT units)	0	0.1	0.51	0.22

* Error in fit to exponential distribution.

^{\S} Standard error

Table 2

Oligomer tunneling noise characteristics. Parameters are defined in Figure 2k.

Oligomer	$d(A)_5$	$d(C)_5$	$d(mC)_5$
Burst Duration (T_B , s)	$0.14 \pm 0.02^*$	$0.15 \pm 0.03^*$	$0.41 \pm 0.03^*$
Burst Frequency (f_B Hz)	$738 \pm 100^{\S}$	$320 \pm 85^{\S}$	$662 \pm 116^{\S}$
Fraction of reads > 0.1 nA	0.20	0.0	0.23
τ_{on} (ms)	$0.33 \pm 0.01^*$	$0.34 \pm 0.02^*$	$0.26 \pm 0.01^*$
τ_{off} (ms)	$0.52 \pm 0.02^*$	$0.42 \pm 0.01^*$	$0.47 \pm 0.01^*$
τ_{on}/τ_{off}	0.6	0.8	0.6
G (kT units)	0.51	0.22	0.51

* Error in fit to exponential distribution.

§ Standard error