Digital Data Storage Using DNA Nanostructures and Solid-State Nanopores

Kaikai Chen, [†,‡] Jinglin Kong, [†,‡] Jinbo Zhu, [†] Niklas Ermann, [†] Paul Predki, [§] and Ulrich F. Keyser [†,]*

[†]Cavendish Laboratory, University of Cambridge, JJ Thomson Avenue, Cambridge, CB3 0HE, United Kingdom

[§]Iridia, Inc. 3156 Lionshead Ave, Suite 1, Carlsbad, CA 92010, USA

*Email: ufk20@cam.ac.uk

[‡]These authors contributed equally.

ABSTRACT

Solid-state nanopores are powerful tools for reading out the three-dimensional shape of molecules allowing for the translation of molecular structure information into electric signals. Here, we show a high-resolution integrated nanopore system for identifying DNA nanostructures that has the capability of distinguishing attached short DNA hairpins with only a stem length difference of 8 bp along a DNA double strand named DNA carrier. Using our platform, we can read up to 112 DNA hairpins with a separating distance of 114 bp attached on a DNA carrier which carries digital information. Our encoding strategy allows for the creation of a library of molecules with a size of up to 5×10^33 (2^112) which is only built from a few hundred types of base molecules for data storage and has the potential to be extended by linking multiple DNA carriers. Our platform provides a nanopore- and DNA nanostructure-based data storage method with convenient access and the potential for miniature scale integration.

KEYWORDS: Solid-state nanopores, nanopore sensing, single-molecule, DNA storage, DNA nanotechnology

Nanopore-based single molecule sensing has demonstrated its applicability for the analysis of single molecules in electrolyte solutions.[1, 2] The principle relies on the modulation of ionic current flow as molecules pass through a nanopore driven by an applied electric field.[3] The current signal reflects information about the passing objects, especially their three-dimensional shapes.[4-6] Thus the nanopore sensing method directly transforms structural information into electric signals and can probe individual molecules at the single-molecule level giving access to analysis with high throughput.[7] Sensors

based on nanopores are especially attractive as they can be used for multiplexed sensing and integrated into low-volume fluidic chips, suitable for miniature portable devices reducing material requirements.[8] In the literature, nanopores have shown the ability to detect objects distributed along a linear DNA strand,[9-13] including DNA nanostructures and proteins, for DNA or protein sensing.[14-16] The DNA nanostructures along a strand can be used to encode digital information.[14] For their utility in data storage, the number of bits stored on one DNA strand can be improved by increasing the density of the structures. The foremost goal is the reduction of the size of the structures and their separating distance on the path to achieving realistic data storage applications.

Digital data can be directly stored in a DNA sequence at high data density.[17, 18] This method depends on the synthesis and sequencing of DNA molecules that need complex devices and high costs for large amounts of data.[19-21] As an alternative, storage of information in three-dimensional DNA nanostructures is an attractive strategy[22, 23] which can be easily integrated with solid-state nanopores. Combined with DNA nanotechnology, nanopores can be used to build devices with exceptional readout capabilities.

In this study, we demonstrate a solid-state nanopore platform for digital data storage. The route is to use high-density DNA hairpins as digital bits along a strand named a DNA carrier. The carriers can be built by mixing DNA scaffolds and oligonucleotides to create a library of a large number of different DNA molecules to store information. We show our nanopore has the capability of detecting DNA hairpins down to a length of around 3 nm to decode the digital information. The mixing method for DNA structure synthesis and nanopore sensing allows high-density data storage and has great potential for miniature scale integration.

First, we establish our system by building DNA hairpins as digital bits along DNA carriers and examine the sensing capability of our nanopore platform. We used an empirical approach to optimise the structure size while maximising the success rate of our nanopore measurements. The most important design goal was to ensure that the nanostructures generate observable secondary current drops while being small enough to ensure smooth DNA translocations. In previous studies structures consisting of 299 bp protrusions[12] and a group of DNA dumbbells[14] were measured with 10 nm and 14 nm diameter nanopores, respectively. These structures allow for easy detection but the data density is severely limited due to a bit size of up to 100 nm. We started by testing ~20 bp ds DNA protrusions which turned out to be undetectable using our established ~14 nm nanopores.[24, 25] As smaller nanopores increase the signal-to-noise ratio but may lead to stalled translocations[26, 27] we decided to decrease the diameter of our glass nanopores to ~5 nm. These smaller pores allowed for detection of dsDNA protrusions consisting of hairpins of less than 25bp and difference in stem length of 8bp. In addition, translocations are observed with almost constant velocity which simplifies data analysis. In this paper we show that 5 nm glass nanopores are suitable for the detection of structures differing in molecular volume by 8bp with good success rate in a salt concentration of 4M LiCl. Future development of lower noise nanopores and machine learning[28] will allow for further reduction of the bit size and increased data density.

We synthesised the DNA carrier with protruding DNA hairpins by hybridizing a 7228-base single strand (linearised M13mp18 DNA,[14] New England Biolabs) as the scaffold and short oligonucleotides (Integrated DNA Technologies) with complementary sequences, as shown in Figure 1a. They were mixed at a ratio of 1/6-1/3 (the scaffold to the

oligonucleotides), annealed for 50 min, and filtered to remove the excess oligonucleotides.[14] The oligonucleotides consist of 38 bp complementary strands and hairpin-linked strands. We built DNA hairpins with different stem lengths of 8 bp and 16 bp (Figure 1a). The full sequences are shown in Supporting Information Section 1. We fabricated nanopores by directly pulling quartz capillaries (outer diameter 0.5 mm and inner diameter 0.2 mm) with a laser-heated pipette puller (P2000, Sutter Instrument). An example Scanning Electron Microscope (SEM) image of the glass nanopores with diameter down to ~5 nm is shown in Figure 1b. The characterisation is shown in Supporting Information Section 2.

The glass nanopores were integrated into our PDMS chips (Figure 1b) and then filled with 4 M LiCl solution[29]. We diluted DNA samples into 4 M LiCl with final concentrations of 0.1-1 nM and added them to the *cis* side (the chamber outside the capillary). 400 mV voltage was applied to drive DNA molecules through nanopores (Figure 1c). The ionic current across the nanopore was recorded with an amplifier (Axopatch 200B, Molecular Devices), filtered with an external 49.9 kHz Bessel filter (Frequency Devices) and digitised at a 250 kHz sampling rate with a data card (PCI 6251, National Instruments). We used custom-made LabVIEW algorithms to collect the data and analyse the current traces to find single translocation signals (named events). Unfolded events[30] were used to identify substructures with their positions along the linear strands.[31]

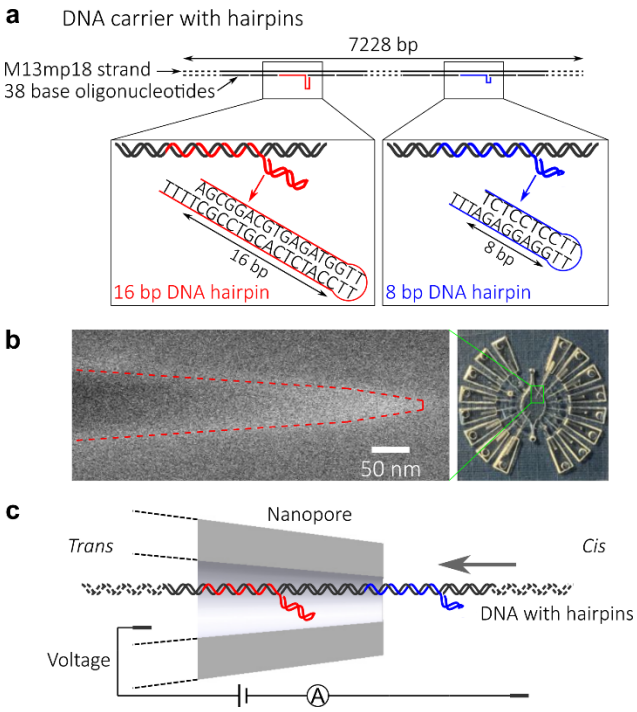

Figure 1. Design of DNA carriers with attached DNA hairpins and nanopore measurement. (a) A DNA carrier with 16 bp (red) and 8 bp (blue) DNA hairpins. Structures and sequences of the hairpins are shown at the bottom. (b) The SEM image of a glass nanopore and the integrated chip. Dashed lines indicate the outer edges (scale bar = 50 nm). (c) Schematic of the measurement of the DNA carrier with a nanopore.

We tested the current and spatial resolution of our detection system with two DNA carrier designs and analysed the translocation events (see data analysis in Supporting Information Section 3). First, we measured DNA carriers with one 16 bp hairpin and one 8 bp hairpin on each (Figure 2a). Two example translocation events are shown with either end of the DNA entering the pore first. The current blockade ΔI_0 caused by the double-stranded DNA backbone corresponds to the nearly constant level while the hairpin structures are indicated by secondary peaks. We analysed the secondary current drop ΔI for the 16 and 8 bp hairpins and plotted the statistics in Figure 2a (right panel). As

expected the Gaussian fits show that the 16 bp hairpin caused a deeper current drop than the 8 bp hairpin. Our glass nanopore system is therefore able to distinguish the protruding hairpins with a size difference of only 8 bp with data from tens of events. In the second design, we placed two 16 bp long hairpins at a distance of 76 bp to explore the minimum distance for bits that can be clearly distinguished with our nanopores (Figure 2b). We defined the resolved and unresolved events by analysing whether the peaks were separated in the time series with examples shown in Figure 2b. The fractions of resolved events for three independent pores show the two peaks were separated for nearly half of all unfolded events. The length of 76 base pairs is only ~26 nm, about ten times below the diffraction limit and comparable to the resolution limit of complex optical imaging systems.[32, 33] Our nanopore platform shows the capability of identifying DNA nanostructures by their size even at small distances of structures.

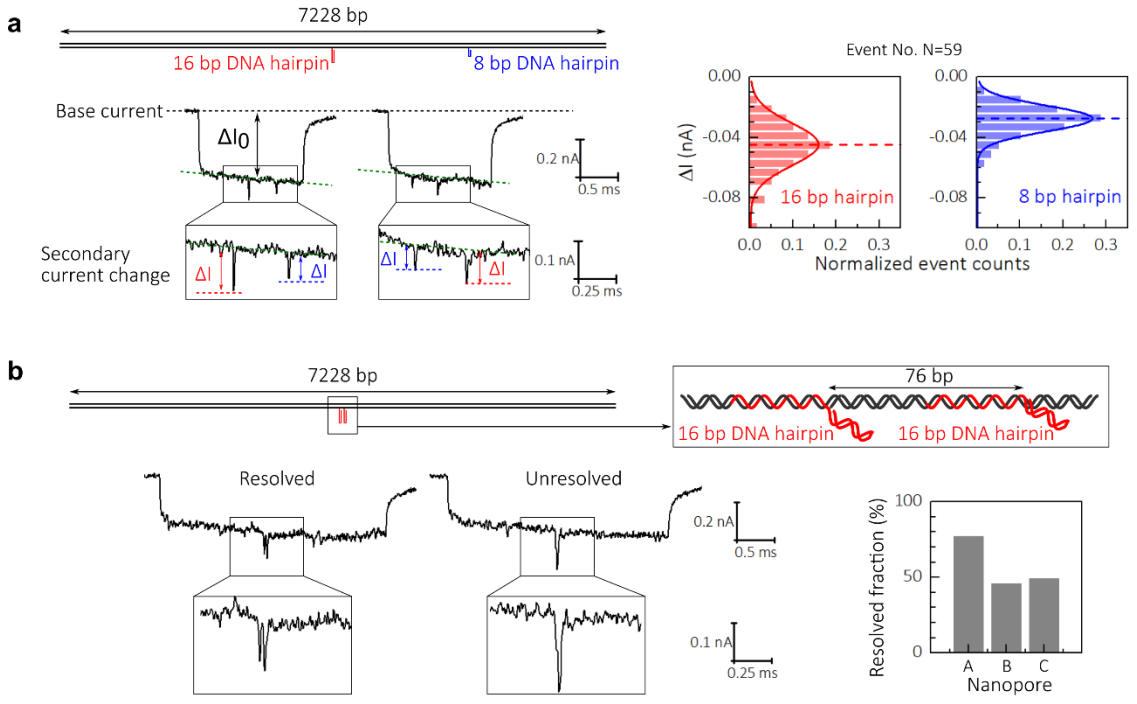

Figure 2. Resolution test of the nanopore system. (a) Discrimination of 16 bp and 8 bp DNA hairpins. The two hairpins are placed at 50% and 75% positions of the carrier. The

secondary current changes in the sample events correspond to the two hairpins attached on the DNA carrier. The right panel shows the histogram of Δ*I* from 59 events. (b) Detection of two close hairpins. Two 16 bp hairpins are placed at the middle position with a distance of 76 bp. The revolved events show two separated peaks while the unresolved events show a single large peak. The fractions of resolved events measured with 3 distinct nanopores are shown in the right panel. The detailed designs with DNA sequences are shown in Supporting Information Section 1.

We then designed dense hairpin arrays on the carriers to encode digital information. In our design, an 8 bp hairpin encodes a digital bit '0' while a 16 bp hairpin a '1'. The distance between bits is set to 114 bp to aid data analysis, and each DNA carrier accommodates 56 hairpins (Figure 3a). The detailed design and sequences are shown in Supporting Information Section 1. The sample was measured with the nanopores shown in Figure 1b. 56 peaks (bits) were fully resolved as shown in the example event in Figure 3b. Using the knowledge about the structure we decoded the information by analysing the current drop amplitudes using a Bayesian inference method[34] (see the description in Supporting Information Section 4). For the single event shown in Figure 3b, we detected 7 errors (marked in grey) in the 56 bits as expected due to the overlap in the current drops of the 16 and 8 bp hairpins (see Figure 2a).

More independent events of the same molecules were used to improve the accuracy of identification of each bit (see examples in Supporting Information Section 4). We aggregated statistics from 58 events and calculated the natural logarithm of the probability ratio of the respective bit being a '1' versus a '0': $\ln\left(\frac{P(H_{1,m}|\{D\})}{P(H_{0,m}|\{D\})}\right)$, where *m* is the bit number

(1-56). Figure 3c shows the analysed results where a large $\left|\ln\left(\frac{P(H_{1,m}|\{D\})}{P(H_{0,m}|\{D\})}\right)\right|$ indicates a high confidence in our estimate. The aggregated analysis of 58 events reduced the number of errors by more than half to only 2. It should be noted that when aggregating the estimate for a particular bit using many events, we might use the values of neighbouring bits in some events where the orders of bits were not correctly counted, introducing a disturbance. This is also reflected by the high confidence with consecutive '0' or '1' and a lower confidence for alternating patterns of '0' and '1'. The confidence in the bit estimation increased with more events used, as shown in Figure 3d, meaning that errors can be reduced by aggregating more events. However, the correct reading of such data can be improved through established encoding strategies such as sub blocks or checksums in the future. Data analysis relying on machine learning can also improve the decoding accuracy.

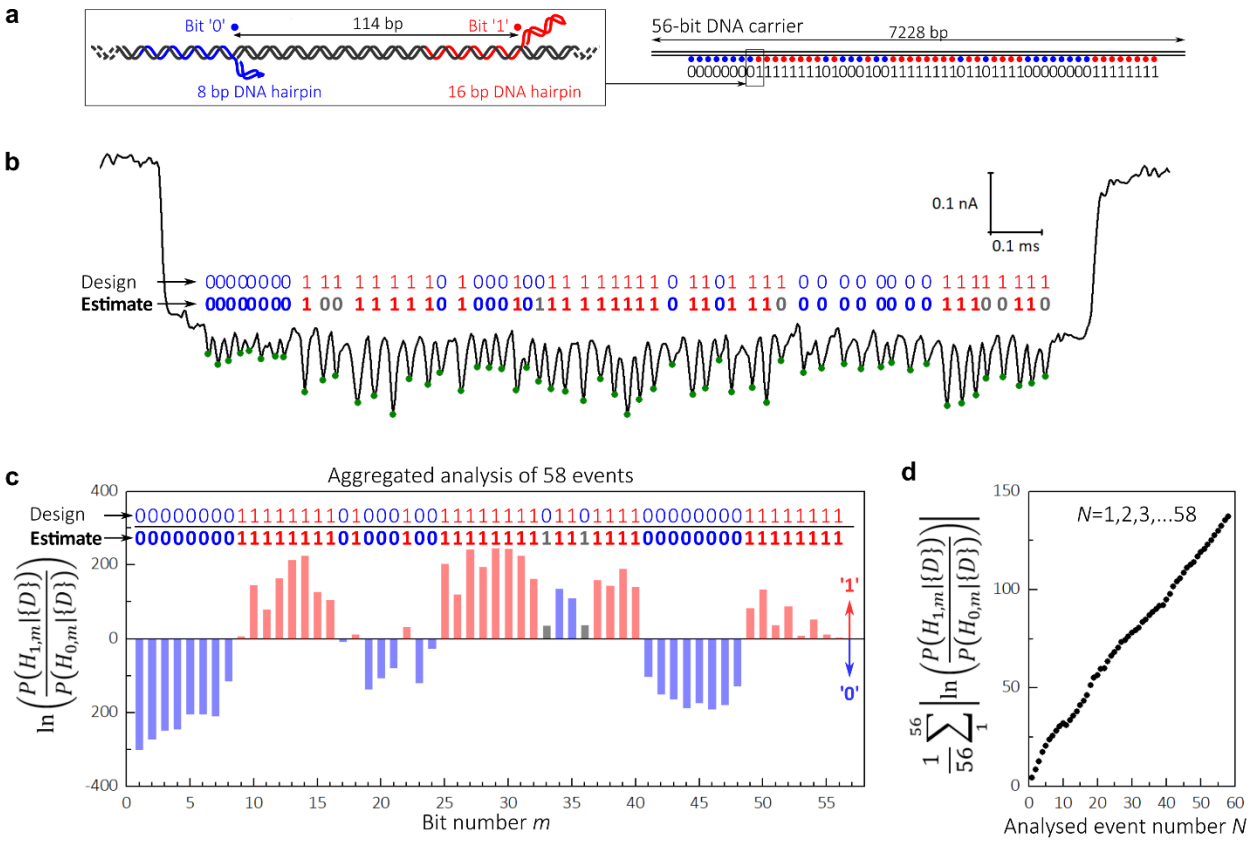

Figure 3. Design and decoding of a 56-bit DNA carrier. (a) Design of the 56-bit DNA carrier with 8 bp (blue) and 16 bp (red) DNA hairpins representing '0' and '1' respectively. (b) An example current signal measured with our nanopores. The grey values indicate misclassifications. (c) Information decoding by aggregating multiple events. The '0' and '1' were extracted by analysing 58 events. $\ln\left(\frac{P(H_{1,m}|\{D\})}{P(H_{0,m}|\{D\})}\right)$ shows the log ratio of the probability of a bit (at number m) being a '1' versus a '0'. (d) Confidence of estimation as a function of analysed event number. $\frac{1}{56}\sum_{1}^{56}\left|\ln\left(\frac{P(H_{1,m}|\{D\})}{P(H_{0,m}|\{D\})}\right)\right|$ shows the average confidence values of the 56 bits.

With even smaller nanopores[35-37] and nanopores on 2D materials[38, 39], one can decrease bit size and consequently increase the data density. In the current stage of the development, our 56-bit DNA carrier allows for libraries with ~7.2×10^16 (2^56) different

molecules with the possibility to encode data by mixing the scaffold and hundreds of oligonucleotides. The data density could be further enhanced by implementing a system with more bits with three or more corresponding current levels. Another route to increase storage capacity is to use longer DNA molecules with more bits.

To this end, we now show the possibility of expanding the bit number on a DNA carrier by linking 7228 bp DNA carriers using complementary single-stranded ends. The design of a dimer and one example event are shown in Figure 4a and Figure 4b. For this designed carrier, we had 12 errors out of the 112 bits by aggregating data from 19 events (marked in grey, Figure 4c). It should be noted that using this strategy, one can significantly increase the number of bits encoded on a DNA strand for larger libraries with 2^(56×n) different molecules, where *n* is the number of monomers linked together. This linking method shows the potential of expanding the maximum information stored using our system.

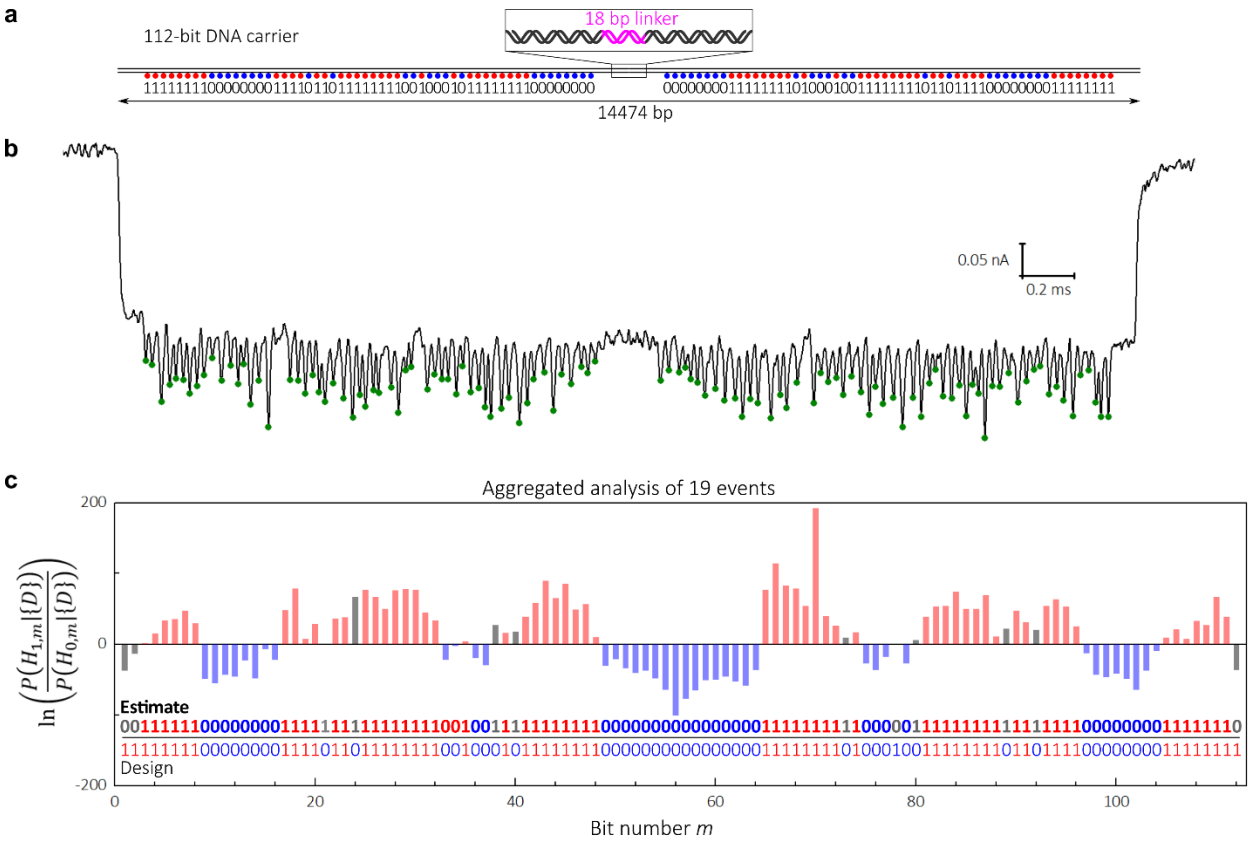

Figure 4. Extendibility of the storage system by linking 7228 bp DNA carriers. (a) Design of a dimer carrying 112-bit information by linking two 7228 bp carriers. (b) An example event caused by the dimer in the nanopore measurement. Green dots show recognised peaks in the analysis. (c) Aggregated analysis of 19 events for decoding the 112-bit information. Grey estimate values show the errors in our estimation (bold font) compared to the design.

To conclude, we developed a platform based on the combination of nanopore sensing and DNA nanotechnology for digital data storage. Our solid-state nanopore platform outperforms previous ones in identifying DNA nanostructures along a DNA strand in terms of resolution and data quality. We were able to detect short 16 bp and 8 bp DNA hairpins and even distinguish between them on the same molecule. Hairpins with a distance of

114 bp were clearly separated in current signals of nanopore measurements and could be resolved with a distance of 76 bp in up to ~50% of the events. We built 56 hairpins on a 7228 bp DNA carrier to store 56-bit information and used a Bayesian inference method with tens of events to reduce the reading error to 2 out of 56 bits. By linking two 7228 bp carriers we encoded 112-bit data onto a single DNA molecule to show the potential of reading longer molecules.

Our digital data storage method shows an alternative to information storage in the DNA base sequence. It has advantages due to the ease of encoding and decoding by combining the simple mixing of oligonucleotides with nanopore sensing, which avoids complex devices and the use of enzymes and can be integrated at miniature scale. In future work, the technique can be improved in a number of ways. For instance, scalable synthesis of carriers can be achieved with micro- and nanofluidics to build a large number of DNA carriers to encode data with the mixing method.[40] The nanopore decoding errors can be reduced by using encoding strategies and further optimising the size of DNA nanostructures and nanopores when parallel measurement systems become available. Even smaller nanopores should allow the design of DNA carriers with more densely spaced bits and more than two different hairpin lengths can be used to encode three or more values per position. Our platform is readily extensible and shows great potential for the integration of molecular information storage into nanoscale systems.

ASSOCIATED CONTENT

Supporting Information

Details on the synthesis and design of DNA carriers, nanopore characterization and measurement, data analysis for the DNA carriers with two hairpins, and data analysis for the 56-bit and 112-bit DNA carriers (PDF).

AUTHOR INFORMATION

Corresponding Author

*Email:ufk20@cam.ac.uk

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript. ‡These authors contributed equally.

Notes

The authors declare the following competing financial interest: the study was partly supported by Iridia Inc. through funding of the postdoc position of Dr Jinglin Kong.

ACKNOWLEDGMENTS

We thank the lab of Mark Howarth at Oxford University for kindly providing monovalent streptavidin. The authors thank Dr K. Misiunas, Dr J. Cama, and Dr A. Thorneywork for discussions and suggestions. K.C. and U.F.K. acknowledge funding from an ERC Consolidator Grant (Designerpores No. 647144). J.K. acknowledges funding from Cambridge Trust and Iridia Inc.. J.Z. acknowledges funding from an EPSRC grant (EP/M008258/1). N.E. acknowledges funding from the EPSRC, Cambridge Trust and Trinity Hall, Cambridge.

REFERENCES

1. Storm, A. J.; Chen, J. H.; Ling, X. S.; Zandbergen, H. W.; Dekker, C. *Nat. Mater.* **2003,** 2, (8), 537-40.
2. Li, J.; Gershow, M.; Stein, D.; Brandin, E.; Golovchenko, J. A. *Nat. Mater.* **2003,** 2, (9), 611-5.
3. Dekker, C. *Nat. Nanotechnol.* **2007,** 2, (4), 209.
4. Waduge, P.; Hu, R.; Bandarkar, P.; Yamazaki, H.; Cressiot, B.; Zhao, Q.; Whitford, P. C.; Wanunu, M. *ACS Nano* **2017,** 11, (6), 5706-5716.
5. Yusko, E. C.; Bruhn, B. R.; Eggenberger, O. M.; Houghtaling, J.; Rollings, R. C.; Walsh, N. C.; Nandivada, S.; Pindrus, M.; Hall, A. R.; Sept, D.; Li, J.; Kalonia, D. S.; Mayer, M. *Nat. Nanotechnol.* **2017,** 12, (4), 360-367.
6. Hu, R.; Rodrigues, J. V.; Waduge, P.; Yamazaki, H.; Cressiot, B.; Chishti, Y.; Makowski, L.; Yu, D.; Shakhnovich, E.; Zhao, Q.; Wanunu, M. *ACS Nano* **2018,** 12, (5), 4494-4502.
7. Gooding, J. J.; Gaus, K. *Angew. Chem. Int. Ed.* **2016,** 55, (38), 11354-11366.
8. Mikheyev, A. S.; Tin, M. M. *Mol. Ecol. Resour.* **2014,** 14, (6), 1097-1102.
9. Singer, A.; Wanunu, M.; Morrison, W.; Kuhn, H.; Frank-Kamenetskii, M.; Meller, A. *Nano Lett.* **2010,** 10, (2), 738-742.
10. Bell, N. A.; Keyser, U. F. *J. Am. Chem. Soc.* **2015,** 137, (5), 2035-41.
11. Plesa, C.; Ruitenberg, J. W.; Witteveen, M. J.; Dekker, C. *Nano Lett.* **2015,** 15, (5), 3153-8.
12. Plesa, C.; van Loo, N.; Ketterer, P.; Dietz, H.; Dekker, C. *Nano Lett.* **2015,** 15, (1), 732-7.
13. Bulushev, R. D.; Marion, S.; Petrova, E.; Davis, S. J.; Maerkl, S. J.; Radenovic, A. *Nano Lett.* **2016,** 16, (12), 7882-7890.
14. Bell, N. A. W.; Keyser, U. F. *Nat. Nanotechnol.* **2016,** 11, (7), 645-651.
15. Sze, J. Y. Y.; Ivanov, A. P.; Cass, A. E. G.; Edel, J. B. *Nat. Commun.* **2017,** 8, (1), 1552.
16. Zahid, O. K.; Wang, F.; Ruzicka, J. A.; Taylor, E. W.; Hall, A. R. *Nano Lett.* **2016,** 16, (3), 2033-9.
17. Cox, J. P. *Trends Biotechnol.* **2001,** 19, (7), 247-250.
18. Bancroft, C.; Bowler, T.; Bloom, B.; Clelland, C. T. *Science* **2001,** 293, (5536), 1763-1765.
19. Goldman, N.; Bertone, P.; Chen, S.; Dessimoz, C.; LeProust, E. M.; Sipos, B.; Birney, E. *Nature* **2013,** 494, (7435), 77-80.
20. Erlich, Y.; Zielinski, D. *Science* **2017,** 355, (6328), 950-954.
21. Organick, L.; Ang, S. D.; Chen, Y.-J.; Lopez, R.; Yekhanin, S.; Makarychev, K.; Racz, M. Z.; Kamath, G.; Gopalan, P.; Nguyen, B. *Nat. Biotechnol.* **2018,** 36, (3), 242.
22. Tumpane, J.; Sandin, P.; Kumar, R.; Powers, V. E. C.; Lundberg, E. P.; Gale, N.; Baglioni, P.; Lehn, J.-M.; Albinsson, B.; Lincoln, P.; Wilhelmsson, L. M.; Brown, T.; Nordén, B. *Chem. Phys. Lett.* **2007,** 440, (1-3), 125-129.
23. Chandrasekaran, A. R.; Levchenko, O.; Patel, D. S.; MacIsaac, M.; Halvorsen, K. *Nucleic Acids Res.* **2017,** 45, (19), 11459-11465.
24. Kong, J.; Zhu, J.; Keyser, U. F. *Chem. Commun.* **2017,** 53, (2), 436-439.

25. Kong, J.; Zhu, J.; Chen, K.; Keyser, U. F. *Adv. Funct. Mater.* **2018**, 1807555.
26. Beamish, E.; Tabard-Cossa, V.; Godin, M. *ACS Sensors* **2017,** 2, (12), 1814-1820.
27. Karau, P.; Tabard-Cossa, V. *ACS Sensors* **2018,** 3, (7), 1308-1315.
28. Misiunas, K.; Ermann, N.; Keyser, U. F. *Nano Lett.* **2018**.
29. Kowalczyk, S. W.; Wells, D. B.; Aksimentiev, A.; Dekker, C. *Nano Lett.* **2012,** 12, (2), 1038-1044.
30. Mihovilovic, M.; Hagerty, N.; Stein, D. *Phys. Rev. Lett.* **2013,** 110, (2), 028102.
31. Chen, K.; Juhasz, M.; Gularek, F.; Weinhold, E.; Tian, Y.; Keyser, U. F.; Bell, N. A. W. *Nano Lett.* **2017,** 17, (9), 5199-5205.
32. Jeffet, J.; Kobo, A.; Su, T.; Grunwald, A.; Green, O.; Nilsson, A. N.; Eisenberg, E.; Ambjörnsson, T.; Westerlund, F.; Weinhold, E. *ACS Nano* **2016,** 10, (11), 9823-9830.
33. Dai, M.; Jungmann, R.; Yin, P. *Nat. Nanotechnol.* **2016,** 11, (9), 798-807.
34. Feroz, F.; Hobson, M.; Bridges, M. *Mon. Not. of the R. Astron. Soc.* **2009,** 398, (4), 1601-1614.
35. Venta, K.; Shemer, G.; Puster, M.; Rodriguez-Manzo, J. A.; Balan, A.; Rosenstein, J. K.; Shepard, K.; Drndic, M. *ACS Nano* **2013,** 7, (5), 4629-4636.
36. Rodriguez-Manzo, J. A.; Puster, M.; Nicolai, A.; Meunier, V.; Drndic, M. *ACS Nano* **2015,** 9, (6), 6555-6564.
37. Carson, S.; Wilson, J.; Aksimentiev, A.; Wanunu, M. *Biophys. J.* **2014,** 107, (10), 2381-93.
38. Garaj, S.; Hubbard, W.; Reina, A.; Kong, J.; Branton, D.; Golovchenko, J. *Nature* **2010,** 467, (7312), 190.
39. Feng, J.; Liu, K.; Bulushev, R. D.; Khlybov, S.; Dumcenco, D.; Kis, A.; Radenovic, A. *Nat. Nanotechnol.* **2015,** 10, (12), 1070.
40. Zhou, X.; Cai, S.; Hong, A.; You, Q.; Yu, P.; Sheng, N.; Srivannavit, O.; Muranjan, S.; Rouillard, J. M.; Xia, Y. *Nucleic Acids Res.* **2004,** 32, (18), 5409-5417.

Supporting Information

Digital Data Storage Using DNA Nanostructures and Solid-State Nanopores

Kaikai Chen, Jinglin Kong, Jinbo Zhu, Niklas Ermann, Paul Predki, and Ulrich F. Keyser

Table of contents

Section 1. Design and sequence of DNA carriers

1.1 Synthesis of DNA carriers

We used linearized M13mp18 single-stranded DNA (ssDNA, 7228 bp after linearization, from New England Biolabs) and complementary oligonucleotides (from Integrated DNA Technologies, abbreviated as oligos in the following) to synthesise DNA carriers. 190 complementary oligos fully occupy the 7228 bp distance with each length of 38 bp (except for the 46 bp two ends, with 42 bp complementary sequences and 'TTTT' as the beginning and end). The oligos are numbered from 1 to 190 with sequences listed in Table S-1. Oligos with additional hairpins or ss DNA overhangs were used instead at designed positions to make protrusions on the carriers. Sequences start at 5' ends in this Supporting Information.

Oligo No.	Sequence	Length (bp)
1	TTTTCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATC	46
2	CGCTCACAATTCCACACAACATACGAGCCGGAAGCATA	38
3	AAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACT	38
4	CACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGT	38
5	CGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCGGC	38
6	CAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCCA	38
7	GGGTGGTTTTTCTTTTCACCAGTGAGACGGGCAACAGC	38
8	TGATTGCCCTTCACCGCCTGGCCCTGAGAGAGTTGCAG	38
9	CAAGCGGTCCACGCTGGTTTGCCCCAGCAGGCGAAAAT	38
10	CCTGTTTGATGGTGGTTCCGAAATCGGCAAAATCCCTT	38
11	ATAAATCAAAAGAATAGCCCGAGATAGGGTTGAGTGTT	38
12	GTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGT	38
13	GGACTCCAACGTCAAAGGGCGAAAAACCGTCTATCAGG	38
14	GCGATGGCCCACTACGTGAACCATCACCCAAATCAAGT	38
15	TTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAA	38
16	CCCTAAAGGGAGCCCCCGATTTAGAGCTTGACGGGGAA	38
17	AGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAAAGCG	38
18	AAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGT	38
19	CACGCTGCGCGTAACCACCACACCCGCCGCGCTTAATG	38
20	CGCCGCTACAGGGCGCGTACTATGGTTGCTTTGACGAG	38
21	CACGTATAACGTGCTTTCCTCGTTAGAATCAGAGCGGG	38
22	AGCTAAACAGGAGGCCGATTAAAGGGATTTTAGACAGG	38
23	AACGGTACGCCAGAATCCTGAGAAGTGTTTTTATAATC	38
24	AGTGAGGCCACCGAGTAAAAGAGTCTGTCCATCACGCA	38
25	AATTAACCGTTGTAGCAATACTTCTTTGATTAGTAATA	38
26	ACATCACTTGCCTGAGTAGAAGAACTCAAACTATCGGC	38
27	CTTGCTGGTAATATCCAGAACAATATTACCGCCAGCCA	38
28	TTGCAACAGGAAAAACGCTCATGGAAATACCTACATTT	38
29	TGACGCTCAATCGTCTGAAATGGATTATTTACATTGGC	38
30	AGATTCACCAGTCACACGACCAGTAATAAAAGGGACAT	38
31	TCTGGCCAACAGAGATAGAACCCTTCTGACCTGAAAGC	38
32	GTAAGAATACGTGGCACAGACAATATTTTTGAATGGCT	38
33	ATTAGTCTTTAATGCGCGAACTGATAGCCCTAAAACAT	38
34	CGCCATTAAAAATACCGAACGAACCACCAGCAGAAGAT	38

35	AAAACAGAGGTGAGGCGGTCAGTATTAACACCGCCTGC	38
36	AACAGTGCCACGCTGAGAGCCAGCAGCAAATGAAAAAT	38
37	CTAAAGCATCACCTTGCTGAACCTCAAATATCAAACCC	38
38	TCAATCAATATCTGGTCAGTTGGCAAATCAACAGTTGA	38
39	AAGGAATTGAGGAAGGTTATCTAAAATATCTTTAGGAG	38
40	CACTAACAACTAATAGATTAGAGCCGTCAATAGATAAT	38
41	ACATTTGAGGATTTAGAAGTATTAGACTTTACAAACAA	38
42	TTCGACAACTCGTATTAAATCCTTTGCCCGAACGTTAT	38
43	TAATTTTAAAAGTTTGAGTAACATTATCATTTTGCGGA	38
44	ACAAAGAAACCACCAGAAGGAGCGGAATTATCATCATA	38
45	TTCCTGATTATCAGATGATGGCAATTCATCAATATAAT	38
46	CCTGATTGTTTGGATTATACTTCTGAATAATGGAAGGG	38
47	TTAGAACCTACCATATCAAAATTATTTGCACGTAAAAC	38
48	AGAAATAAAGAAATTGCGTAGATTTTCAGGTTTAACGT	38
49	CAGATGAATATACAGTAACAGTACCTTTTACATCGGGA	38
50	GAAACAATAACGGATTCGCCTGATTGCTTTGAATACCA	38
51	AGTTACAAAATCGCGCAGAGGCGAATTATTCATTTCAA	38
52	TTACCTGAGCAAAAGAAGATGATGAAACAAACATCAAG	38
53	AAAACAAAATTAATTACATTTAACAATTTCATTTGAAT	38
54	TACCTTTTTTAATGGAAACAGTACATAAATCAATATAT	38
55	GTGAGTGAATAACCTTGCTTCTGTAAATCGTCGCTATT	38
56	AATTAATTTTCCCTTAGAATCCTTGAAAACATAGCGAT	38
57	AGCTTAGATTAAGACGCTGAGAAGAGTCAATAGTGAAT	38
59	TTATCAAAATCATAGGTCTGAGAGACTACCTTTTTAAC	38
59	CTCCGGCTTAGGTTGGGTTATATAACTATATGTAAATG	38
60	CTGATGCAAATCCAATCGCAAGACAAAGAACGCGAGAA	38
61	AACTTTTTCAAATATATTTTAGTTAATTTCATCTTCTG	38
62	ACCTAAATTTAATGGTTTGAAATACCGACCGTGTGATA	38
63	AATAAGGCGTTAAATAAGAATAAACACCGGAATCATAA	38
64	TTACTAGAAAAAGCCTGTTTAGTATCATATGCGTTATA	38
65	CAAATTCTTACCAGTATAAAGCCAACGCTCAACAGTAG	38
66	GGCTTAATTGAGAATCGCCATATTTAACAACGCCAACA	38
67	TGTAATTTAGGCAGAGGCATTTTCGAGCCAGTAATAAG	38
68	AGAATATAAAGTACCGACAAAAGGTAAAGTAATTCTGT	38
69	CCAGACGACGACAATAAACAACATGTTCAGCTAATGCA	38
70	GAACGCGCCTGTTTATCAACAATAGATAAGTCCTGAAC	38
71	AAGAAAAATAATATCCCATCCTAATTTACGAGCATGTA	38
72	GAAACCAATCAATAATCGGCTGTCTTTCCTTATCATTC	38
73	CAAGAACGGGTATTAAACCAAGTACCGCACTCATCGAG	38
74	AACAAGCAAGCCGTTTTTATTTTCATCGTAGGAATCAT	38
75	TACCGCGCCCAATAGCAAGCAAATCAGATATAGAAGGC	38
76	TTATCCGGTATTCTAAGAACGCGAGGCGTTTTAGCGAA	38
77	CCTCCCGACTTGCGGGAGGTTTTGAAGCCTTAAATCAA	38
78	GATTAGTTGCTATTTTGCACCCAGCTACAATTTTATCC	38
79	TGAATCTTACCAACGCTAACGAGCGTCTTTCCAGAGCC	38
80	TAATTTGCCAGTTACAAAATAAACAGCCATATTATTTA	38
81	TCCCAATCCAAATAAGAAACGATTTTTTGTTTAACGTC	38
82	AAAAATGAAAATAGCAGCCTTTACAGAGAGAATAACAT	38
83	AAAAACAGGGAAGCGCATTAGACGGGAGAATTAACTGA	38
84	ACACCCTGAACAAAGTCAGAGGGTAATTGAGCGCTAAT	38
85	ATCAGAGAGATAACCCACAAGAATTGAGTTAAGCCCAA	38
86	TAATAAGAGCAAGAAACAATGAAATAGCAATAGCTATC	38
87	TTACCGAAGCCCTTTTTAAGAAAAGTAAGCAGATAGCC	38
88	GAACAAAGTTACCAGAAGGAAACCGAGGAAACGCAATA	38
89	ATAACGGAATACCCAAAAGAACTGGCATGATTAAGACT	38
90	CCTTATTACGCAGTATGTTAGCAAACGTAGAAAATACA	38
91	TACATAAAGGTGGCAACATATAAAAGAAACGCAAAGAC	38
92	ACCACGGAATAAGTTTATTTTGTCACAATCAATAGAAA	38
93	ATTCATATGGTTTACCAGCGCCAAAGACAAAAGGGCGA	38
94	CATTCAACCGATTGAGGGAGGGAAGGTAAATATTGACG	38
95	GAAATTATTCATTAAAGGTGAATTATCACCGTCACCGA	38
96	CTTGAGCCATTTGGGAATTAGAGCCAGCAAAATCACCA	38
97	GTAGCACCATTACCATTAGCAAGGCCGGAAACGTCACC	38
98	AATGAAACCATCGATAGCAGCACCGTAATCAGTAGCGA	38
99	CAGAATCAAGTTTGCCTTTAGCGTCAGACTGTAGCGCG	38
100	TTTTCATCGGCATTTTCGGTCATAGCCCCCTTATTAGC	38
101	GTTTGCCATCTTTTCATAATCAAAATCACCGGAACCAG	38
102	AGCCACCACCGGAACCGCCTCCCTCAGAGCCGCCACCC	38
103	TCAGAACCGCCACCCTCAGAGCCACCACCCTCAGAGCC	38
104	GCCACCAGAACCACCACCAGAGCCGCCGCCAGCATTGA	38
105	CAGGAGGTTGAGGCAGGTCAGACGATTGGCCTTGATAT	38

106	TCACAAACAAATAAATCCTCATTAAAGCCAGAATGGAA	38
107	AGCGCAGTCTCTGAATTTACCGTTCCAGTAAGCGTCAT	38
108	ACATGGCTTTTGATGATACAGGAGTGTACTGGTAATAA	38
109	GTTTTAACGGGGTCAGTGCCTTGAGTAACAGTGCCCGT	38
110	ATAAACAGTTAATGCCCCCTGCCTATTTCGGAACCTAT	38
111	TATTCTGAAACATGAAAGTATTAAGAGGCTGAGACTCC	38
112	TCAAGAGAAGGATTAGGATTAGCGGGGTTTTGCTCAGT	38
113	ACCAGGCGGATAAGTGCCGTCGAGAGGGTTGATATAAG	38
114	TATAGCCCGGAATAGGTGTATCACCGTACTCAGGAGGT	38
115	TTAGTACCGCCACCCTCAGAACCGCCACCCTCAGAACC	38
116	GCCACCCTCAGAGCCACCACCCTCATTTTCAGGGATAG	38
117	CAAGCCCAATAGGAACCCATGTACCGTAACACTGAGTT	38
118	TCGTCACCAGTACAAACTACAACGCCTGTAGCATTCCA	38
119	CAGACAGCCCTCATAGTTAGCGTAACGATCTAAAGTTT	38
120	TGTCGTCTTTCCAGACGTTAGTAAATGAATTTTCTGTA	38
121	TGGGATTTTGCTAAACAACTTTCAACAGTTTCAGCGGA	38
122	GTGAGAATAGAAAGGAACAACTAAAGGAATTGCGAATA	38
123	ATAATTTTTTCACGTTGAAAATCTCCAAAAAAAAGGCT	38
124	CCAAAAGGAGCCTTTAATTGTATCGGTTTATCAGCTTG	38
125	CTTTCGAGGTGAATTTCTTAAACAGCTTGATACCGATA	38
126	GTTGCGCCGACAATGACAACAACCATCGCCCACGCATA	38
127	ACCGATATATTCGGTCGCTGAGGCTTGCAGGGAGTTAA	38
128	AGGCCGCTTTTGCGGGATCGTCACCCTCAGCAGCGAAA	38
129	GACAGCATCGGAACGAGGGTAGCAACGGCTACAGAGGC	38
130	TTTGAGGACTAAAGACTTTTTCATGAGGAAGTTTCCAT	38
131	TAAACGGGTAAAATACGTAATGCCACTACGAAGGCACC	38
132	AACCTAAAACGAAAGAGGCAAAAGAATACACTAAAACA	38
133	CTCATCTTTGACCCCCAGCGATTATACCAAGCGCGAAA	38
134	CAAAGTACAACGGAGATTTGTATCATCGCCTGATAAAT	38
135	TGTGTCGAAATCCGCGACCTGCTCCATGTTACTTAGCC	38
136	GGAACGAGGCGCAGACGGTCAATCATAAGGGAACCGAA	38
137	CTGACCAACTTTGAAAGAGGACAGATGAACGGTGTACA	38
138	GACCAGGCGCATAGGCTGGCTGACCTTCATCAAGAGTA	38
139	ATCTTGACAAGAACCGGATATTCATTACCCAAATCAAC	38
140	GTAACAAAGCTGCTCATTCAGTGAATAAGGCTTGCCCT	38
141	GACGAGAAACACCAGAACGAGTAGTAAATTGGGCTTGA	38
142	GATGGTTTAATTTCAACTTTAATCATTGTGAATTACCT	38
143	TATGCGATTTTAAGAACTGGCTCATTATACCAGTCAGG	38
144	ACGTTGGGAAGAAAAATCTACGTTAATAAAACGAACTA	38
145	ACGGAACAACATTATTACAGGTAGAAAGATTCATCAGT	38
146	TGAGATTTAGGAATACCACATTCAACTAATGCAGATAC	38
147	ATAACGCCAAAAGGAATTACGAGGCATAGTAAGAGCAA	38
148	CACTATCATAACCCTCGTTTTACCAGACGACGATAAAAA	38
149	CCAAAATAGCGAGAGGCTTTTGCAAAAGAAGTTTTGCC	38
150	AGAGGGGGTAATAGTAAAATGTTTAGACTGGATAGCGT	38
151	CCAATACTGCGGAATCGTCATAAATATTCATTGAATCC	38
152	CCCTCAAATGCTTTAAACAGTTCAGAAAACGAGAATGA	38
153	CCATAAATCAAAAATCAGGTCTTTACCCTGACTATTAT	38
154	AGTCAGAAGCAAAGCGGATTGCATCAAAAAGATTAAGA	38
155	GGAAGCCCGAAAGACTTCAAATATCGCGTTTTAATTCG	38
156	AGCTTCAAAGCGAACCAGACCGGAAGCAAACTCCAACA	38
157	GGTCAGGATTAGAGAGTACCTTTAATTGCTCCTTTTGA	38
158	TAAGAGGTCATTTTTGCGGATGGCTTAGAGCTTAATTG	38
159	CTGAATATAATGCTGTAGCTCAACATGTTTTAAATATG	38
160	CAACTAAAGTACGGTGTCTGGAAGTTTCATTCCATATA	38
161	ACAGTTGATTCCCAATTCTGCGAACGAGTAGATTTAGT	38
162	TTGACCATTAGATACATTTCGCAAATGGTCAATAACCT	38
163	GTTTAGCTATATTTTCATTTGGGGCGCGAGCTGAAAAG	38
164	GTGGCATCAATTCTACTAATAGTAGTAGCATTAACATC	38
165	CAATAAATCATACAGGCAAGGCAAAGAATTAGCAAAAT	38
166	TAAGCAATAAAGCCTCAGAGCATAAAGCTAAATCGGTT	38
167	GTACCAAAAACATTATGACCCTGTAATACTTTTGCGGG	38
168	AGAAGCCTTTATTTCAACGCAAGGATAAAAATTTTTAG	38
169	AACCCTCATATATTTTAAATGCAATGCCTGAGTAATGT	38
170	GTAGGTAAAGATTCAAAAGGGTGAGAAAGGCCGGAGAC	38
171	AGTCAAATCACCATCAATATGATATTCAACCGTTCTAG	38
172	CTGATAAATTAATGCCGGAGAGGGTAGCTATTTTTGAG	38
173	AGATCTACAAAGGCTATCAGGTCATTGCCTGAGAGTCT	38
174	GGAGCAAACAAGAGAATCGATGAACGGTAATCGTAAAA	38
175	CTAGCATGTCAATCATATGTACCCCGGTTGATAATCAG	38
176	AAAAGCCCCAAAAACAGGAAGATTGTATAAGCAAATAT	38

177	TTAAATTGTAAACGTTAATATTTTGTTAAAATTCGCAT	38
178	TAAATTTTTGTTAAATCAGCTCATTTTTTAACCAATAG	38
179	GAACGCCATCAAAAATAATTCGCGTCTGGCCTTCCTGT	38
180	AGCCAGCTTTCATCAACATTAAATGTGAGCGAGTAACA	38
181	ACCCGTCGGATTCTCCGTGGGAACAAACGGCGGATTGA	38
182	CCGTAATGGGATAGGTCACGTTGGTGTAGATGGGCGCA	38
183	TCGTAACCGTGCATCTGCCAGTTTGAGGGGACGACGAC	38
184	AGTATCGGCCTCAGGAAGATCGCACTCCAGCCAGCTTT	38
185	CCGGCACCGCTTCTGGTGCCGGAAACCAGGCAAAGCGC	38
186	CATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCG	38
187	ATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGA	38
188	AAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACG	38
189	CCAGGGTTTTCCCAGTCACGACGTTGTAAAACGACGGC	38
190	CAGTGCCAAGCTTGCATGCCTGCAGGTCGACTCTAGAGGATCTTTT	46

Table S-1. Sequences of the 190 complementary oligos.

The DNA carrier synthesis is introduced in a former study.[1] The carrier was made by mixing together 8 μL M13mp18 DNA (100 nM), 20 μL oligonucleotide mix (each oligo 200 nM), 4 μL 100 mM MgCl2, 1.2 μL 100 mM Tris-HCl (pH=8), 10 mM EDTA, and 6.8 μL deionised water. Then the mixture was heated to 70°C followed by a linear cooling ramp to 25°C over 50 min. After annealing, these excess oligonucleotides were removed using Amicon Ultra 100 kDa filters. One tube annealed as above was added with 460 μL of 10mM Tris-HCl (pH=8), 0.5 mM $MgCl_2$ and centrifuged at 9000g for 10 min at 4°C. 460 μL more 10mM Tris-HCl (pH=8), 0.5 mM $MgCl_2$ was added and the sample was centrifuged again for 10 min. The sample was then recovered by turning the filter upside down and centrifuging for 2 min at 1000g. The solution was immediately added after filtering to make the final concentration 10 mM Tris-HCl, 100 mM NaCl, 2 mM $MgCl_2$.

Designs of protrusions on DNA carriers are shown in the following parts.

1.2 Design of the DNA carrier with one 16 bp and one 8 bp hairpins

This was made by replacing the Oligo 94 and Oligo 142 with Oligo 16H94 and Oligo 8H142 shown below with red and blue parts marked to form the hairpins.

Oligo 16H94

CATTCAACCGATTGAGGGAGGGAAGGTAAATATTGACG

TTTTCGCCTGCACTCTACCTTTTGGTAGAGTGCAGGCGA

Oligo 8H142

GATGGTTTAATTTCAACTTTAATCATTGTGAATTACCT TTTAGAGGAGGTTTTCCTCCTCT

1.3 Design of the DNA carrier with two 16 bp hairpins with a distance of 76 bp

The Oligo 94 and Oligo 96 were replaced with Oligo 16H94 and Oligo 16H96 shown below.

Oligo 16H94

CATTCAACCGATTGAGGGAGGGAAGGTAAATATTGACG

TTTTCGCCTGCACTCTACCTTTTGGTAGAGTGCAGGCGA

Oligo 16H96

CTTGAGCCATTTGGGAATTAGAGCCAGCAAAATCACCA TTTCGCCTGCACTCTACCTTTTGGTAGAGTGCAGGCGA

1.4 Design of the 56-bit and 112-bit DNA carriers

In this design, the Oligos 13, 16, 19, … 178 (increase by 3, 56 in total) were replaced with strands consisting of 16 or 8 bp hairpins and 38 bp parts to bind to the scaffold. The sequences are given in Tables S-2 and S-3.

Name	Sequence	Replaced oligo No.
16H37	CTAAAGCATCACCTTGCTGAACCTCAAATATCAAACCC TTTTCGCCTGCACTCTACCTTTTGGTAGAGTGCAGGCGA	37
16H40	CACTAACAACTAATAGATTAGAGCCGTCAATAGATAAT TTTTCGCCTGCACTCTACCTTTTGGTAGAGTGCAGGCGA	40
16H43	TAATTTTAAAAGTTTGAGTAACATTATCATTTTGCGGA TTTTCGCCTGCACTCTACCTTTTGGTAGAGTGCAGGCGA	43
16H46	CCTGATTGTTTGGATTATACTTCTGAATAATGGAAGGG TTTTCGCCTGCACTCTACCTTTTGGTAGAGTGCAGGCGA	46
16H49	CAGATGAATATACAGTAACAGTACCTTTTACATCGGGA TTTTCGCCTGCACTCTACCTTTTGGTAGAGTGCAGGCGA	49
16H52	TTACCTGAGCAAAAGAAGATGATGAAACAAACATCAAG TTTTCGCCTGCACTCTACCTTTTGGTAGAGTGCAGGCGA	52
16H55	GTGAGTGAATAACCTTGCTTCTGTAAATCGTCGCTATT TTTTCGCCTGCACTCTACCTTTTGGTAGAGTGCAGGCGA	55
16H58	TTATCAAAATCATAGGTCTGAGAGACTACCTTTTTAAC TTTTCGCCTGCACTCTACCTTTTGGTAGAGTGCAGGCGA	58
16H64	TTACTAGAAAAAGCCTGTTTAGTATCATATGCGTTATA TTTTCGCCTGCACTCTACCTTTTGGTAGAGTGCAGGCGA	64
16H76	TTATCCGGTATTCTAAGAACGCGAGGCGTTTTAGCGAA TTTTCGCCTGCACTCTACCTTTTGGTAGAGTGCAGGCGA	76
16H85	ATCAGAGAGATAACCCACAAGAATTGAGTTAAGCCCAA TTTTCGCCTGCACTCTACCTTTTGGTAGAGTGCAGGCGA	85

16H88	GAACAAAGTTACCAGAAGGAAACCGAGGAAACGCAATA TTTTCGCCTGCACTCTACCTTTTGGTAGAGTGCAGGCGA	88
16H91	TACATAAAGGTGGCAACATATAAAAGAAACGCAAAGAC TTTTCGCCTGCACTCTACCTTTTGGTAGAGTGCAGGCGA	91
16H94	CATTCAACCGATTGAGGGAGGGAAGGTAAATATTGACG TTTTCGCCTGCACTCTACCTTTTGGTAGAGTGCAGGCGA	94
16H97	GTAGCACCATTACCATTAGCAAGGCCGGAAACGTCACC TTTTCGCCTGCACTCTACCTTTTGGTAGAGTGCAGGCGA	97
16H100	TTTTCATCGGCATTTTCGGTCATAGCCCCCTTATTAGC TTTTCGCCTGCACTCTACCTTTTGGTAGAGTGCAGGCGA	100
16H103	TCAGAACCGCCACCCTCAGAGCCACCACCCTCAGAGCC TTTTCGCCTGCACTCTACCTTTTGGTAGAGTGCAGGCGA	103
16H106	TCACAAACAAATAAATCCTCATTAAAGCCAGAATGGAA TTTTCGCCTGCACTCTACCTTTTGGTAGAGTGCAGGCGA	106
16H112	TCAAGAGAAGGATTAGGATTAGCGGGGTTTTGCTCAGT TTTTCGCCTGCACTCTACCTTTTGGTAGAGTGCAGGCGA	112
16H115	TTAGTACCGCCACCCTCAGAACCGCCACCCTCAGAACC TTTTCGCCTGCACTCTACCTTTTGGTAGAGTGCAGGCGA	115
16H121	TGGGATTTTGCTAAACAACTTTCAACAGTTTCAGCGGA TTTTCGCCTGCACTCTACCTTTTGGTAGAGTGCAGGCGA	121
16H124	CCAAAAGGAGCCTTTAATTGTATCGGTTTATCAGCTTG TTTTCGCCTGCACTCTACCTTTTGGTAGAGTGCAGGCGA	124
16H127	ACCGATATATTCGGTCGCTGAGGCTTGCAGGGAGTTAA TTTTCGCCTGCACTCTACCTTTTGGTAGAGTGCAGGCGA	127
16H130	TTTGAGGACTAAAGACTTTTTCATGAGGAAGTTTCCAT TTTTCGCCTGCACTCTACCTTTTGGTAGAGTGCAGGCGA	130
16H157	GGTCAGGATTAGAGAGTACCTTTAATTGCTCCTTTTGA TTTTCGCCTGCACTCTACCTTTTGGTAGAGTGCAGGCGA	157
16H160	CAACTAAAGTACGGTGTCTGGAAGTTTCATTCCATATA TTTTCGCCTGCACTCTACCTTTTGGTAGAGTGCAGGCGA	160
16H163	GTTTAGCTATATTTTCATTTGGGGCGCGAGCTGAAAAG TTTTCGCCTGCACTCTACCTTTTGGTAGAGTGCAGGCGA	163
16H166	TAAGCAATAAAGCCTCAGAGCATAAAGCTAAATCGGTT TTTTCGCCTGCACTCTACCTTTTGGTAGAGTGCAGGCGA	166
16H169	AACCCTCATATATTTTAAATGCAATGCCTGAGTAATGT TTTTCGCCTGCACTCTACCTTTTGGTAGAGTGCAGGCGA	169
16H172	CTGATAAATTAATGCCGGAGAGGGTAGCTATTTTTGAG TTTTCGCCTGCACTCTACCTTTTGGTAGAGTGCAGGCGA	172
16H175	CTAGCATGTCAATCATATGTACCCCGGTTGATAATCAG TTTTCGCCTGCACTCTACCTTTTGGTAGAGTGCAGGCGA	175
16H178	TAAATTTTTGTTAAATCAGCTCATTTTTTAACCAATAG TTTTCGCCTGCACTCTACCTTTTGGTAGAGTGCAGGCGA	178

Table S-2. Oligos with 16 bp hairpins for the 56-bit DNA carrier.

Name	Sequence	Replaced oligo No.
8H13	GGACTCCAACGTCAAAGGGCGAAAAACCGTCTATCAGG TTTAGAGGAGGTTTTCCTCCTCT	13
8H16	CCCTAAAGGGAGCCCCCGATTTAGAGCTTGACGGGGAA TTTAGAGGAGGTTTTCCTCCTCT	16
8H19	CACGCTGCGCGTAACCACCACACCCGCCGCGCTTAATG TTTAGAGGAGGTTTTCCTCCTCT	19
8H22	AGCTAAACAGGAGGCCGATTAAAGGGATTTTAGACAGG TTTAGAGGAGGTTTTCCTCCTCT	22
8H25	AATTAACCGTTGTAGCAATACTTCTTTGATTAGTAATA TTTAGAGGAGGTTTTCCTCCTCT	25
8H28	TTGCAACAGGAAAAACGCTCATGGAAATACCTACATTT TTTAGAGGAGGTTTTCCTCCTCT	28
8H31	TCTGGCCAACAGAGATAGAACCCTTCTGACCTGAAAGC TTTAGAGGAGGTTTTCCTCCTCT	31
8H34	CGCCATTAAAAATACCGAACGAACCACCAGCAGAAGAT TTTAGAGGAGGTTTTCCTCCTCT	34
8H61	AACTTTTCAAATATATTTTAGTTAATTTCATCTTCTG TTTAGAGGAGGTTTTCCTCCTCT	61
8H67	TGTAATTTAGGCAGAGGCATTTTCGAGCCAGTAATAAG TTTAGAGGAGGTTTTCCTCCTCT	67
8H70	GAACGCGCCTGTTTATCAACAATAGATAAGTCCTGAAC TTTAGAGGAGGTTTTCCTCCTCT	70
8H73	CAAGAACGGGTATTAAACCAAGTACCGCACTCATCGAG TTTAGAGGAGGTTTTCCTCCTCT	73
8H79	TGAATCTTACCAACGCTAACGAGCGTCTTTCCAGAGCC TTTAGAGGAGGTTTTCCTCCTCT	79
8H82	AAAAATGAAAATAGCAGCCTTTACAGAGAGAATAACAT TTTAGAGGAGGTTTTCCTCCTCT	82
8H109	GTTTTAACGGGGTCAGTGCCTTGAGTAACAGTGCCCGT TTTAGAGGAGGTTTTCCTCCTCT	109
8H118	TCGTCACCAGTACAAACTACAACGCCTGTAGCATTCCA TTTAGAGGAGGTTTTCCTCCTCT	118
8H133	CTCATCTTTGACCCCCAGCGATTATACCAAGCGCGAAA TTTAGAGGAGGTTTTCCTCCTCT	133
8H136	GGAACGAGGCGCAGACGGTCAATCATAAGGGAACCGAA TTTAGAGGAGGTTTTCCTCCTCT	136
8H139	ATCTTGACAAGAACCGGATATTCATTACCCAAATCAAC TTTAGAGGAGGTTTTCCTCCTCT	139
8H142	GATGGTTTAATTTCAACTTTAATCATTGTGAATTACCT TTTAGAGGAGGTTTTCCTCCTCT	142
8H145	ACGGAACAACATTATTACAGGTAGAAAGATTCATCAGT TTTAGAGGAGGTTTTCCTCCTCT	145
8H148	CACTATCATAACCCTCGTTTACCAGACGACGATAAAAA TTTAGAGGAGGTTTTCCTCCTCT	148
8H151	CCAATACTGCGGAATCGTCATAAATATTCATTGAATCC TTTAGAGGAGGTTTTCCTCCTCT	151
8H154	AGTCAGAAGCAAAGCGGATTGCATCAAAAAGATTAAGA TTTAGAGGAGGTTTTCCTCCTCT	154

Table S-3. Oligos with 8 bp hairpins for the 56-bit DNA carrier.

The 112-bit DNA carrier was formed by linking two types of 56-bit DNA carriers. For the two types, everything is the same as shown above except that the first oligos were replaced with the linkers (Linkers A and B) shown below. Then the link was achieved by mixing two samples and keeping at room temperature for 1 h.

Linker A

AAGTCCTCGTCGGTCGCG CGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATC

Linker B

CGCGACCGACGAGGACTT CGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATC

Section 2. Nanopore characterization and measurement

The fabrication parameters, scanning electron microscopy (SEM) images, model and calculations and example data are shown below.

2.1 Nanopore fabrication parameters

We used one-line programs for the fabrication of the nanopores. The parameters are shown below. Please refer to the P-2000 manual[2] for the explanation of the parameters.

HEAT=575, FIL=0, VEL=25, DEL=170, PUL=225

2.2 SEM images

An example nanopore scanning electron microscope (SEM) image is shown in Figure S-1. The inner diameter was estimated by the outer diameter of the quartz capillary considering that the ratio of the inner diameter to outer diameter kept constant during fabrication.

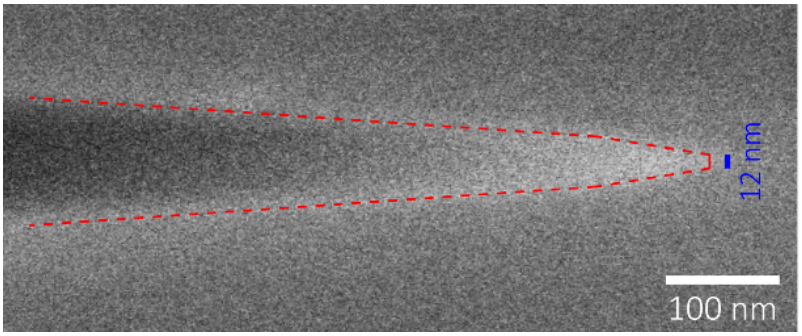

Figure S-1. An example SEM image of the nanopore. Scale bar = 100 nm. The blue bars show the outer diameters.

2.3 Model and calculations

We model the geometry of the pores as two parts with two different conical angles – the part close to the tip with length L and the other with infinite length (Figure S-2).

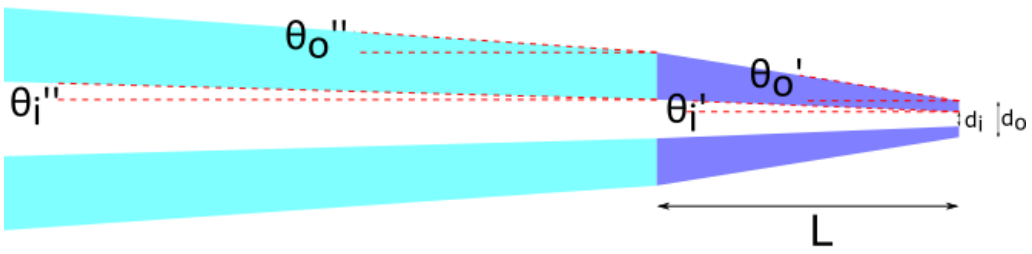

Figure S-2. Model of the geometry of the glass nanopores. The two parts with different conical angles are marked in different colours.

The inner conical angles are θ_i' and θ_i'' and outer conical angles are θ_o' and θ_o'' respectively. d_i and d_o are the inner and outer diameters.

We then calculate the ionic current through the pore with and without DNA blocking it.

The resistance of the pore without DNA is calculated as[3]

$$R_{Nanopore} = \frac{4L}{\sigma\pi d_i\left(d_i + 2Ltan\theta_i^{'}\right)} + \frac{2}{\sigma\pi\left(d_i + 2Ltan\theta_i^{'}\right)tan\theta_i^{''}} + \frac{1}{2\sigma d_i}$$

where σ is the conductivity of the solution.

The inner diameter and conical angles are inferred from the outer values considering the ratio of the inner diameter to outer diameter is constant, and are calculated as

$$d_i = 0.4d_o$$

$$tan\theta_i^{'} = 0.4tan\theta_o^{'}$$

$$tan\theta_i^{''} = 0.4 tan\theta_o^{''}$$

where 0.4 (0.2/0.5) is the ratio.

When the DNA is in the pore, the resistance is then (a is the DNA radius)

$$R_{Nanopore\ with\ DNA} = \frac{1}{2\sigma\pi a tan\theta_i^{'}}\left(\ln\frac{2Ltan\theta_i^{'} + d_i - 2a}{2Ltan\theta_i^{'} + d_i + 2a} - ln\frac{d_i - 2a}{d_i + 2a}\right)$$

$$-\frac{1}{2\sigma\pi a tan\theta_i^{''}}\ln\frac{2Ltan\theta_i^{'} + d_i - 2a}{2Ltan\theta_i^{'} + d_i + 2a} + \frac{1}{2\sigma d_i}$$

Substituting in the values we estimate from the SEM images, the current values can be calculated. The conductivity of 4 M LiCl of σ=18.4 S/m and the DNA radius of a=1.25 nm are used.

For the ~5 nm diameter nanopores

d_i=5 nm, L=100 nm, θ_i'=0.064 rad, θ_i''=0.027 rad

The current values at 400 mV without DNA and with DNA are calculated as 2.51 nA and 2.32 nA, with the difference of 0.18 nA. The calculated current values agree with our experimental data shown below.

2.4 Nanopore measurement

The nanopores were integrated into Polydimethylsiloxane (PDMS) chips (Figure S-3).

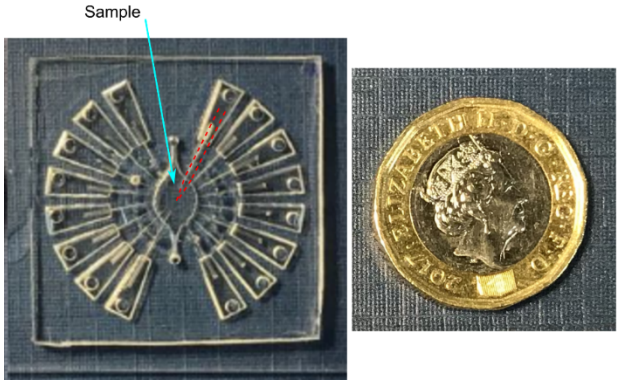

Figure S-3. A nanopore chip integrated with 16 nanopores. The coin (one British Pound) shows the scale. An example channel is marked by red dashed lines.

IV curves of the pores are measured in 4 M LiCl solution and shown in Figure S-4.

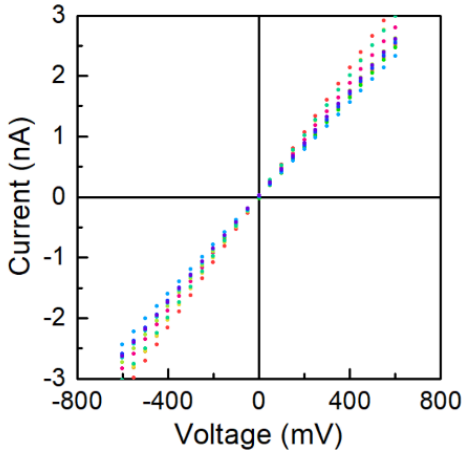

Figure S-4. IV curves of the ~5 nm diameter nanopores in 4 M LiCl solution. 12 pores were measured.

Examples of a current trace and typical events are shown in Figure S-5.

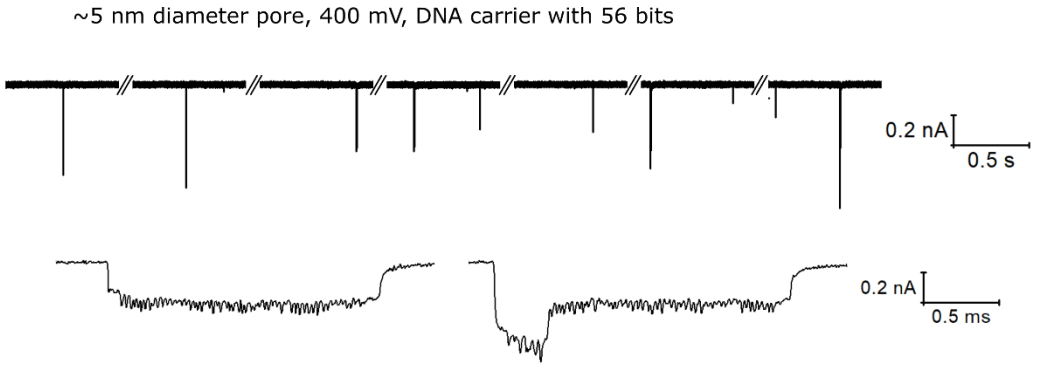

Figure S-5. An example current trace and example events. An example current trace and example events caused by the 56-bit carriers at 400 mV measured with a ~5 nm diameter nanopore.

Section 3. Data analysis for the DNA carrier with two hairpins

3.1 Data analysis for the DNA carrier with two hairpins, one 16 bp and one 8 bp

The data was analysed with a peak finding algorithm[4]. Single unfolded events were picked to extract the amplitudes of the secondary current drop caused by the 16 and 8 bp hairpins. Example events are shown in Figure S-6.

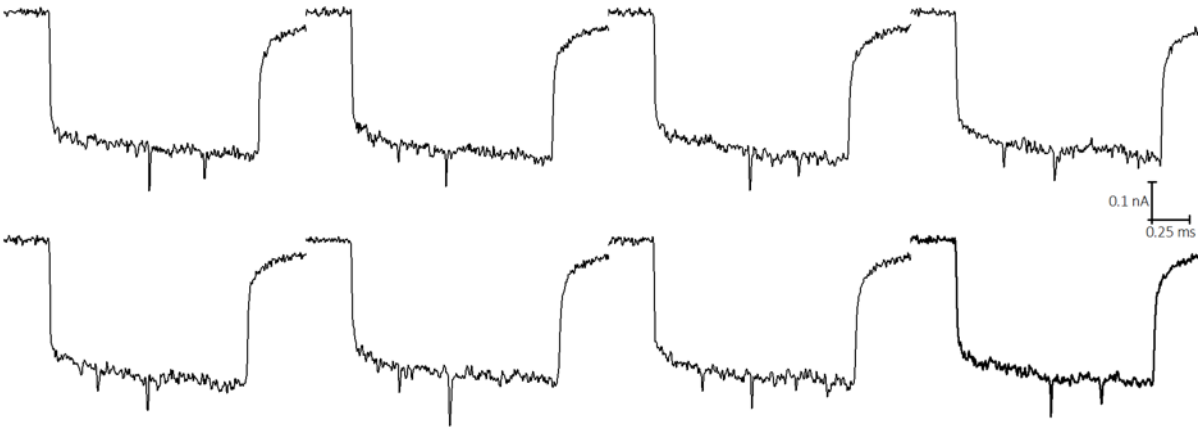

Figure S-6. Example events caused by the DNA carrier with one 16 bp and one 8 bp hairpins.

3.2 Data analysis for the DNA carrier with two close 16 hairpins

Three nanopores were used for the measurement with example events shown in Figure S-7 and statistics shown in Table S-4.

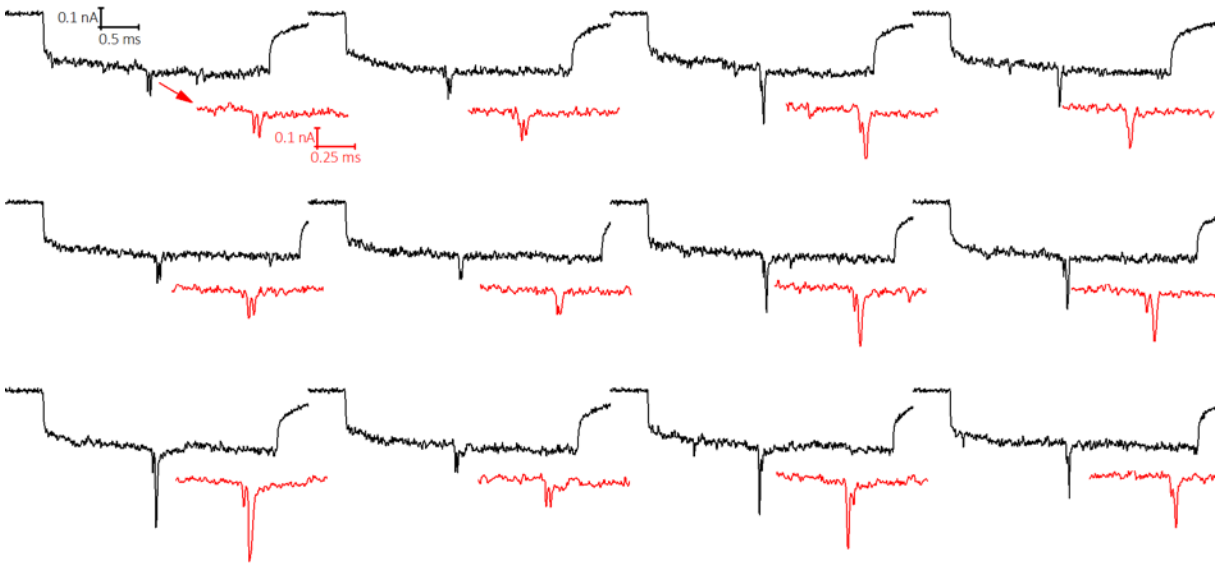

Figure S-7. Example events caused by the DNA carrier with two 16 bp hairpins with a 114 bp **distance.** Red plots show the zoom in details.

No.	Pore name	Unfolded event No.	Readable event No.	Resolved event No.	Percentage of the resolved event (%)
1	76_A	90	88	68	77.3
2	76_B	116	109	50	45.9
3	76_C	108	103	51	49.5

Table S-4. Statistics of the events from 3 pores measured with the DNA carrier with two 16 bp hairpins.

Section 4. Data analysis for the 56-bit and 112-bit DNA carriers

4.1 Example events and statistics

We measured the 56-bit and 112-bit DNA carriers with two nanopores. The sample was made by mixing 56-bit DNA carriers with the linking ends. Not all the carriers were linked together so we had both 56-bit and 112-bit carriers in the solution. In the nanopore measurement, we had 103 unfolded events in total, with 66 of monomers (56-bit carriers) and 37 dimers (112-bit carriers) separated by the translocation durations. A preselected algorithm was used to assess the events with 58 monomers and 19 dimers retained for the bit analysis. Example events are shown in Figures S-8 and S-9.

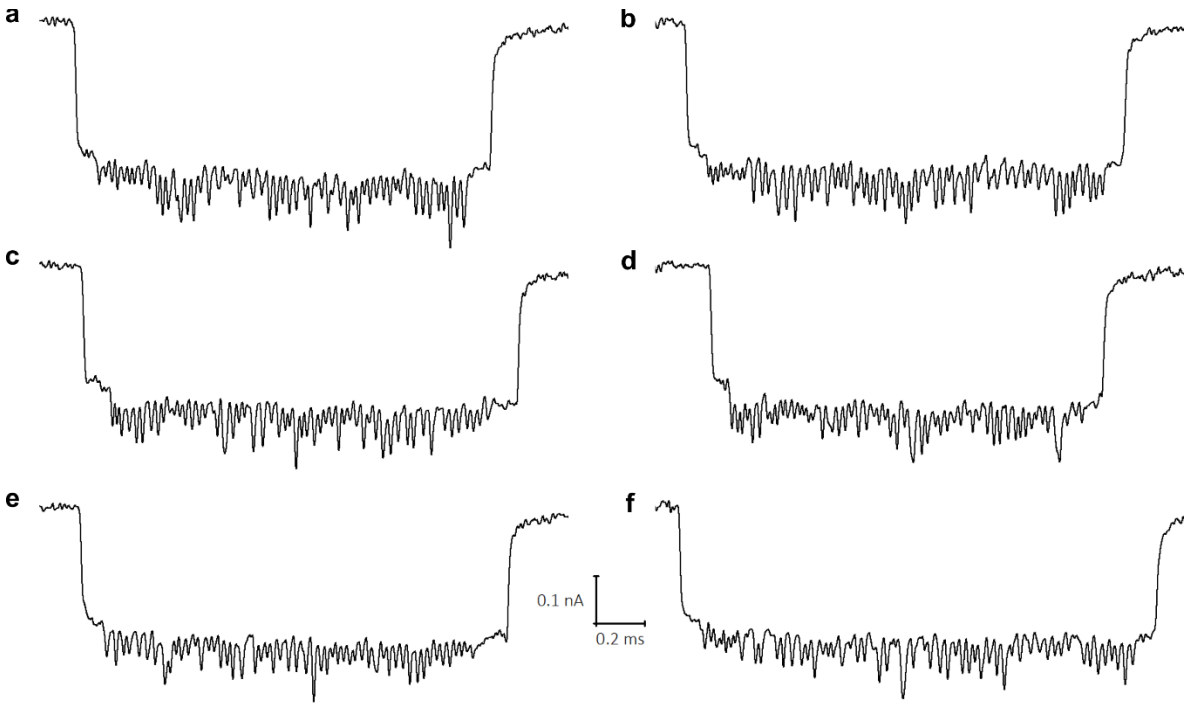

Figure S-8. Examples of events caused by the 56-bit DNA carriers.

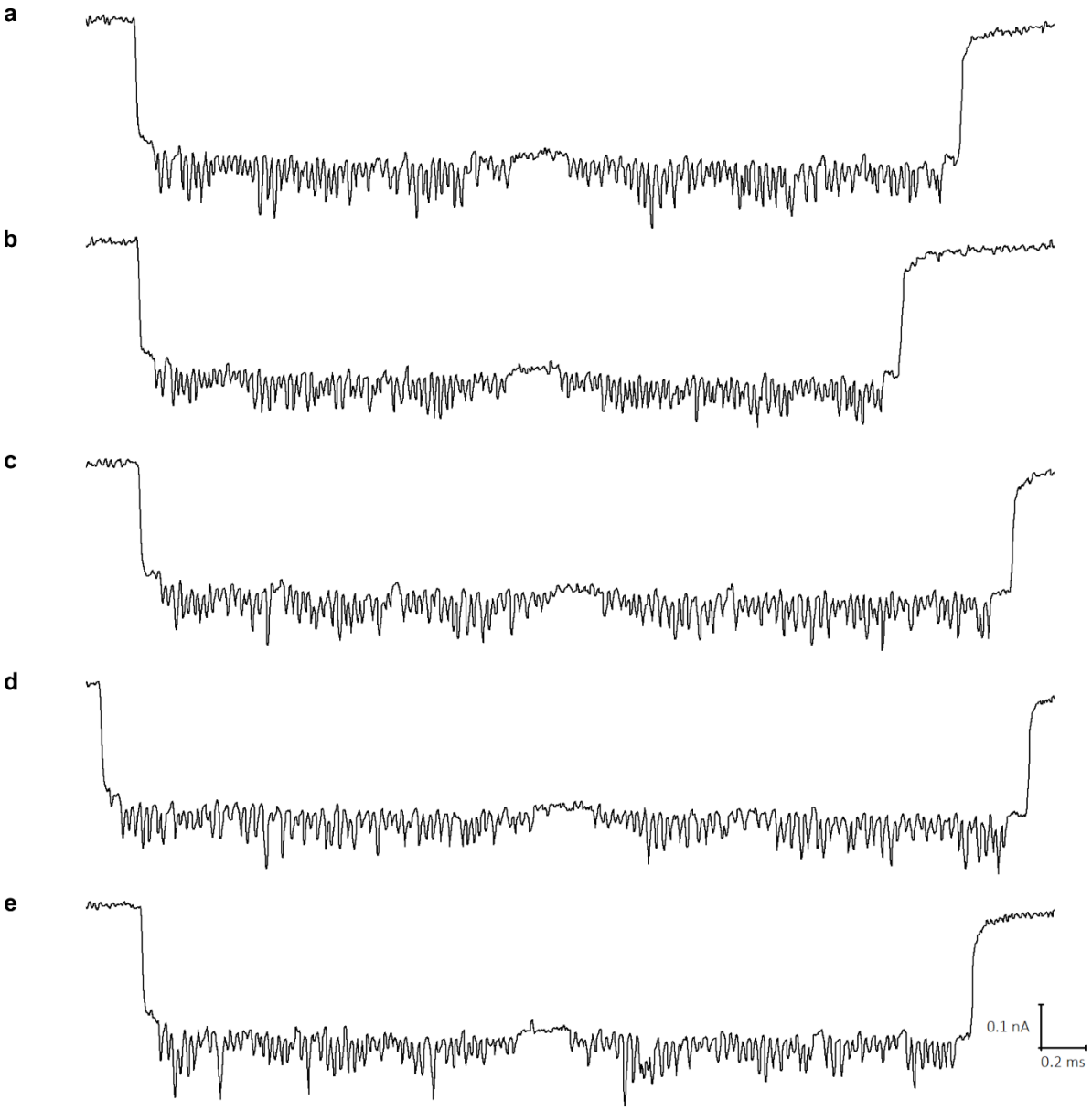

Figure S-9. Examples of events caused by the 112-bit DNA carriers.

4.2 Data analysis with the Bayesian inference method

Bayesian techniques emerge from a fundamental statement about conditional probabilities called Bayes' rule.[5] The Bayesian inference method for nanopore data analysis will be discussed in an upcoming paper from our group.

References

1. Bell, N.A.W. & Keyser, U.F. Digitally encoded DNA nanostructures for multiplexed, single-molecule protein sensing with nanopores. *Nat. Nanotechnol.* **11**, 645-651 (2016).
2. P-2000 laser-based micropipette puller system operation manual REV. 2.4 (20120105).
3. Bell, N.A. & Keyser, U.F. Specific protein detection using designed DNA carriers and nanopores. *J. Am. Chem. Soc.* **137**, 2035-2041 (2015).
4. Chen, K. et al. Ionic Current-Based Mapping of Short Sequence Motifs in Single DNA Molecules Using Solid-State Nanopores. *Nano Lett.* **17**, 5199-5205 (2017).
5. Gelman, A., Carlin, J.B., Stern, H.S. & Rubin, D.B. Bayesian Data Analysis Second edition Corrected version (30 Jan 2008). (1995).