
Article:
Dunn, Katherine Elizabeth orcid.org/0000-0002-5068-4354, Trefzer, Martin Albrecht orcid.org/0000-0002-6196-6832, Johnson, Steven David orcid.org/0000-0002-1786-3182 et al. (1 more author) (2016) Assessing the potential of surface-immobilized molecular logic machines for integration with solid state technology. Biosystems. pp. 3-9. ISSN: 0303-2647

https://doi.org/10.1016/j.biosystems.2016.05.006

Assessing the potential of surface-immobilized molecular logic machines for integration with solid state technology

Katherine E. Dunn, Martin A. Trefzer, Steven Johnson, Andy M. Tyrrell

Department of Electronics, University of York, Heslington, York, YO10 5DD, U.K.

Abstract

Molecular computation with DNA has great potential for low power, highly parallel information processing in a biological or biochemical context. However, significant challenges remain for the field of DNA computation. New technology is needed to allow multiplexed label-free readout and to enable regulation of molecular state without addition of new DNA strands. These capabilities could be provided by hybrid bioelectronic systems in which biomolecular computing is integrated with conventional electronics through immobilization of DNA machines on the surface of electronic circuitry. Here we present a quantitative experimental analysis of a surface-immobilized OR gate made from DNA and driven by strand displacement. The purpose of our work is to examine the performance of a simple representative surface-immobilized DNA logic machine, to provide valuable information for future work on hybrid bioelectronic systems involving DNA devices. We used a quartz crystal microbalance to examine a DNA monolayer containing approximately 5×10^{11} gates cm^{-2}, with an inter-gate separation of approximately 14 nm, and we found that the ensemble of gates took approximately 6 minutes to switch. The gates could be switched repeatedly, but the switching efficiency was significantly degraded on the second and subsequent cycles when the binding site for the input was near to the surface. Otherwise, the switching efficiency could be 80% or better, and the power dissipated by the ensemble of gates during switching was approximately $0.1\,\mathrm{nW\,cm}^{-2}$, which is orders of magnitude less than the power dissipated during switching of an equivalent array of transistors. We propose an architecture for hybrid

*Corresponding author.

Email address: katherine.dunn@york.ac.uk (Katherine E. Dunn)

DNA-electronic systems in which information can be stored and processed, either in series or in parallel, by a combination of molecular machines and conventional electronics. In this architecture, information can flow freely and in both directions between the solution-phase and the underlying electronics via surface-immobilized DNA machines that provide the interface between the molecular and electronic domains.

Keywords: Molecular computation; DNA nanotechnology; molecular machine; bioelectronics; biochemical information processing

1. Introduction

The use of biological molecules to perform computation was pioneered by Adleman [1], who demonstrated that it was possible to solve an instance of the Hamiltonian path problem with DNA strands. DNA is particularly attractive for molecular computation because the interactions between oligonucleotides are highly predictable and programmable (through the base sequence), and the raw materials are relatively cheap and easy to acquire. Various types of logic gates have been assembled using DNA [25, 28], and they have also been implemented in biocompatible nanorobots that could in principle be used for smart drug delivery [7]. It has also been shown that simple gates can be combined to form adders [15] and subtractors [17]. Recently, DNA circuits have been proven to be capable of computing a square root [20] and a DNA-based neural network has been demonstrated [21]. Mixtures of DNA and DNA-manipulating enzymes have been used for computation, and enabled the implementation of a finite automaton [3]. Later, Costa Santini *et al.* constructed a DNA finite state machine in which transitions from one state to another were triggered by a clock signal [5].

Most of these previous achievements involved solution-phase reactions, and it is the use of freely diffusing molecules that underpins many of the challenges that currently limit DNA computation. At present, most techniques used to read out the results of computations involve the use of probes or reporter complexes which carry fluorescent labels. This severely limits the potential for multiplexing and complicates the measurement process. Furthermore, it is difficult to control the state of the ensemble of biological machines without adding new molecular species. We suggest that it would be possible to address these challenges through development of hybrid DNA-electronic technology, in which computational machines made from DNA are

immobilized directly onto underlying electronic circuitry.

In our proposed hybrid approach, information would be able to pass freely in both directions across the molecular-electronic interface. The underlying electronic components could interrogate the state of the surface-immobilized DNA machines, apply logic processing to the data received and act accordingly to regulate the molecular state. This would enable multiplexed readout of the DNA machines, eliminating the limitations imposed by the use of fluorescent labels, and could provide a new approach for the construction of cascades of multiple DNA machines, where the underlying electronics would be used to form connections between individual devices.

Importantly, biomolecular processes in solution offer huge potential for parallelization of computation, and this could be harnessed in a hybrid system, while it would also be possible to interface the technology with biological matter for applications in biology and medicine. Conventional semiconductor devices could be used to perform high-speed operations, complementing the slow bio-compatible molecular elements.

Surface-immobilization of DNA machines is integral to our proposal for hybrid computation, but to date, surface-based DNA computation has received comparatively little attention. However, with a combination of DNA molecules, an endonuclease, and a ligase, it is possible to implement a finite automaton on a surface, as described by Soreni *et al* [26], while Frezza *et al.* proved that DNA strand displacement can be used to operate and cascade logic gates immobilized on the surface of a gold nanoparticle [10].

Another strategy presented for surface-based DNA computing is as follows. DNA molecules encoding 'words' are immobilized on a surface, and each word contains both a representation of variable values and a label for directing word selection through controlled hybridization. Application of specific oligonucleotides and DNA-processing enzymes to the surface enables a computation to be performed, where some strands are eliminated and others are retained. This approach was used by Liu *et al* to solve a four-variable four-clause 3-SAT (satisfiability) problem [18], and Wang *et al* extended the technique to accommodate DNA strands containing multiple words [32]. Subsequently, Su & Smith showed that such a model could be used to construct a universal computer based on NOR and OR gates [29].

In this paper we present a quantitative assessment of the performance of a surface-immobilized OR gate made from DNA and driven by DNA strand displacement. It is not directly linked to electronic circuitry, but by studying the behaviour and properties of the gate when it is immobilized on the surface

of a gold electrode we can obtain valuable insights into the advantages and constraints of hybrid computation. We have designed it to serve as a useful case study to inform future development of systems of the kind described above, and to be a representative example of a surface-immobilized molecular logic machine.

The implementation of our logic gate is very similar to that of Frezza *et al.* [10], but our design is simpler and our choice of experimental technique enables us to use a reporter-free method to read the output and to observe the kinetics of switching. We also examine the behaviour of the gate when it is restored to its original condition and switched repeatedly. Based on our results, we present a quantitative assessment of the performance and attributes of the device, in terms of speed, error frequency, feature size, potential for parallel processing, and power dissipation, in comparison with conventional silicon technology. We proceed to suggest a possible architecture for hybrid systems.

2. Experimental setup and methods

Our DNA OR gate is shown in Fig. 1. To study the operation of our surface-immobilized DNA logic gate, we used the technique of Quartz Crystal Microbalance with Dissipation monitoring (QCM-D), which involves the measurement of the resonant frequencies of oscillation and associated Q-values for an acoustic wave generated by a driven piezoelectric sensor crystal, as described in [6]. The crystal has a gold electrode on either side, across which the driving voltage is applied. One of the electrodes is exposed to a solution, such that molecules from the solution can bind to the surface or molecules from the surface can dissociate into solution. In this paper we consider changes in resonant frequency, which are closely correlated with changes in surface-immobilized mass. In general, an increase in frequency implies that mass has been lost from the surface, while a decrease in frequency implies that mass has been added. Measurements can be made at multiple resonant frequencies (overtones), and here we focus on the 13^{th} overtone, which has the shortest penetration depth - equal to about 70 nm in pure water and is thus less sensitive to changes in the bulk solution. The apparatus also allows us to measure the dissipation of energy by the acoustic wave as it propagates from the sensor through the molecular layer and into solution, where the dissipation is defined as the inverse of the quality factor.

The experimental procedure we used is as follows. All experiments were performed in 1×TE buffer in the presence of 1M NaCl and DNA sequences are provided in Table 1.

The logic gate consists of a partially double-stranded DNA molecule (the 'capture complex') and a 'set' strand G that binds to the capture complex through hybridization, as shown in Fig. 1. The capture complex was formed in solution by hybridization of thiolated strand CS with unmodified strand X, where the sequences of strands CS and X are given in Table 1 and the concentration of each strand was 300 nM. The hybridization reaction occurred in a tube placed on the laboratory bench, where the room temperature was approximately 20°C. The mixture was left for an extended period of time (over one hour) before the tube was transferred to the QCM-D system. Freshly cleaned gold-coated sensors were placed in the QCM-D flow module, frequency and dissipation baseline signals were established, and the capture complex was immobilized on the surface of all sensors by flowing the pre-formed construct through the flow module at a rate of 20 μL/min (Fig. 1 (A)). Next, mercaptohexanol (MCH) was supplied to the sensor at a concentration of 1 mM (Fig. 1 (B)). MCH acts as a backfilling agent, filling in the spaces between immobilized DNA molecules. The gate was then 'set' by application of strand G (Fig. 1 (C)), and operated by application of one or both of the 'input' strands (Fig. 1 (D)). Between application of gate and input strands, fresh buffer solution was used to rinse the flow module for a few minutes. The concentration of the input strands was 600 nM and the sequences are given in Table 1. Throughout the experiment the flow module temperature was maintained at 16°C.

The gate is operated by the well-known process of toehold-mediated DNA strand displacement [35, 36, 27]. When an input strand is supplied to the DNA-functionalized surface, the recognition domain (a* or c*) of the input can bind to the toehold of an immobilized capture strand. Branch migration occurs and the incumbent strand G is displaced, dissociating into solution and leaving a capture complex bound to the surface.

All chemicals were supplied by Sigma Aldrich and all DNA was acquired from Integrated DNA Technologies (IDT), with standard desalting purification for unmodified strands and HPLC purificiation for the thiol-modified oligonucleotide. Oligos were stored at 4°C in 1×TE.

QCM-D experiments were performed using a Q-Sense E4 system with gold-coated QSX 301 quartz sensors (fundamental frequency 5 MHz, best sensitivity 0.5 ng cm^{-2}), both supplied by Biolin Scientific. We provided

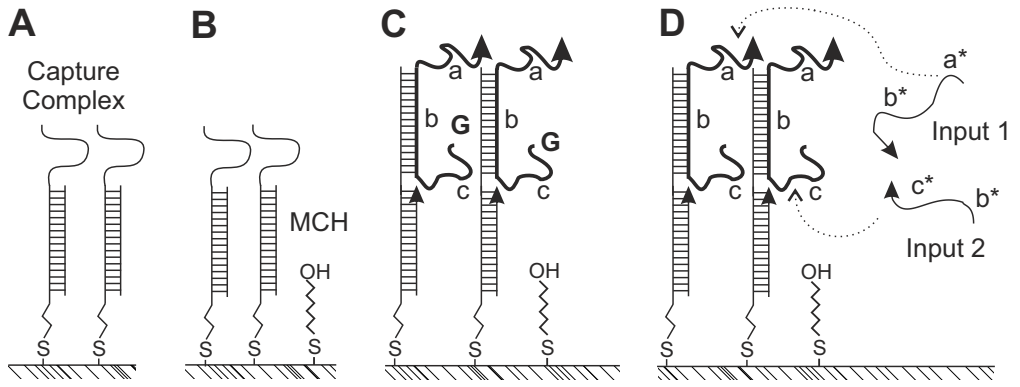


Figure 1: Testing a surface-immobilized OR gate made from DNA. (A) The pre-formed capture complex is immobilized on the surface to form a monolayer of DNA gates. (B) The backfilling agent mercaptohexanol (MCH) is used to fill in spaces between the DNA molecules in the monolayer. (C) Strand G (bold line) is hybridized with the capture complex (the gate is 'set'). Strand G consists of three domains, indicated by 'a', 'b' and 'c'. (D) The DNA-functionalized surface is exposed to input DNA strands. The domains labelled a* and c* on the input strands bind to the recognition domains of G, as indicated by the dotted arrows. When G has been displaced by either input, the system reverts to the state shown in (B) and must be reset by re-hybridization with G before it can respond to a new input. The displacement of G corresponds to the OR operation.

details of our sensor cleaning procedure in [8].

3. Results

The results of our QCM-D measurements are presented in Fig. 2 for the first cycle. During immobilization of the capture complex forming the core of the logic gate, the mass of the layer on the surface increases, which causes the measured resonant frequency to decrease. The same effect is observed upon addition of the backfilling agent and the set strand G. In the absence of any inputs, the signal does not change, as evidenced by the plateau observed in all three traces before addition of the inputs. When either or both of the inputs are supplied, the frequency increases, which indicates that the surface-immobilized mass decreases. This is attributable to the dissociation of the waste product formed after displacement of strand G from the capture complex where the waste product is a duplex comprising strand G and one of the inputs. The observation of a significant mass loss and an associated

Table 1: Sequences of DNA strands used in this study, written 5' to 3'. The thiol modification was in oxidized form (as a disulphide), with a short hydrocarbon chain on both sides. We used NUPACK [34] to assess the degree of complementarity of sequences. For a solution reaction at 16°C between the gate strand G and either one of the input strands or the control strand, where all strands are at a concentration of 600 nM, the package predicts that 100% of the input strands will hybridize with the gate but NONE of the control strands will do so. Even at temperatures as low as 4°C none of the control strands are predicted to bind to the gate.

Strand	**Sequence**
CS	ACA CGC ATA CAC CCA T-thiol
X	ATG GGT GTA TGC GTG TTT AAA GAC CCT AAG CT
G	TCC CGA CCA GCT TAG GGT CTT TAA GCG TGA AG
Input 1	CTT CAC GCT TAA AGA CCC TAA GCT
Input 2	TTA AAG ACC CTA AGC TGG TCG GGA
Control strand	GTC ATT TCT CTA AGT A

frequency shift corresponds to an output of 'true', and consequently our data confirms that the device yields a true result in the presence of input 1 or input 2 or both inputs, exactly as expected for an OR gate. No change is observed upon addition of a control strand of DNA with a sequence unrelated to those of the other strands used in the experiment (Fig. 4(A), last segment of data), which confirms that the logic gate is triggered specifically and selectively by the designed inputs.

The switching efficiency is defined as the frequency change induced by the addition of an input (for input 1, this is F_i in Fig. 2(A)), divided by the frequency change induced by addition of strand G (for input 1, this is F_g in Fig. 2(A)). For complete switching, if all of the gates were triggered, every loaded strand 'G' would be removed and the efficiency would be 100%. The measured values are shown in Fig. 3, which shows that the efficiency is lower for input 2 than for input 1, and the highest efficiency ($\approx$100%) is obtained when both inputs are present. The efficiency calculated for the case of the control is illustrated, and it is seen to be negligible.

When the system is cycled, the efficiency of switching decreases significantly for input 2, but decreases slowly when input 1 is present either alone or in conjunction with input 2. This is shown explicitly in Fig. 4(A). As an example, for input 1 the switching efficiency for iteration 1 is F_i/F_g, for iteration 2 it is F_i'/F_g and for the third iteration it is F_i''/F_g, where these quantities are defined in Fig. 4E. Upon the third application of input 2, the

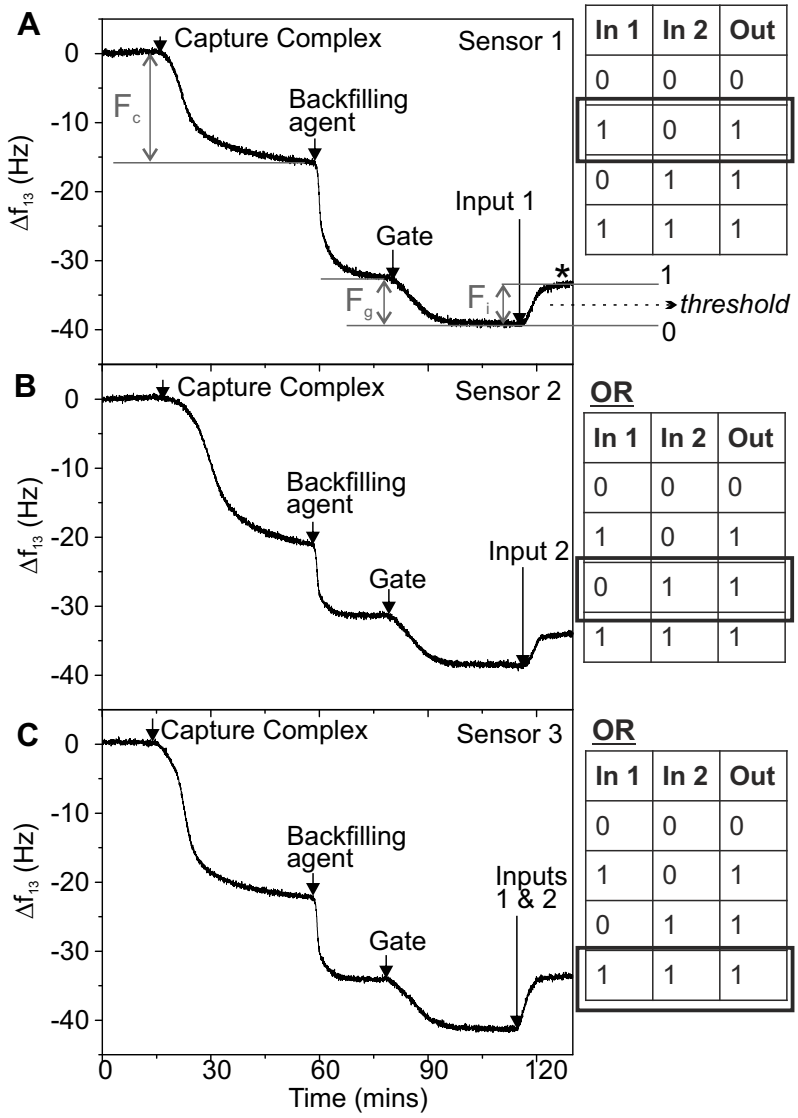


Figure 2: Frequency shifts observed with QCM-D as a function of time as the indicated molecules are applied to the surface. Data illustrates operation of the logic gate for different inputs. (A) Input 1. F_c, F_g, F_i are the frequency shifts measured upon immobilization of the capture construct, the initalization of the gate with G (iteration 0), and the first application of input 1 (iteration 1), respectively. The asterisk denotes the plateau corresponding to the output after the switching event. (B) Input 2. (C) Both inputs, simultaneously. Arrows indicate the times at which the indicated molecules were supplied. Truth tables indicate values of inputs and output for each case.

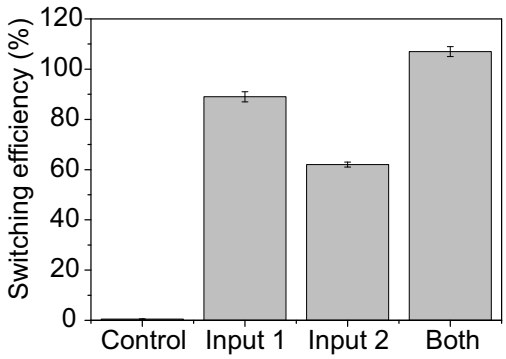


Figure 3: Efficiency of switching (F_i/F_g) as a percentage, for the first application of the indicated strands.

induced frequency change is practically zero, but subsequent application of input 1 does result in switching, with an efficiency of approximately 80%.

The efficiency of setting the gate is defined as the frequency shift caused by application of the strand G divided by the frequency shift resulting from immobilization of the capture construct - for initialization (iteration 0) of the gate in the experiment with input 1, this is F_g/F_c (Fig. 2(A)). The mass of a single molecule of strand G is less than the mass of a single capture construct, and hence the setting efficiency observed for 1:1 binding should be scaled by the ratio of the masses. The normalized setting efficiency is shown explicitly for each iteration in Fig. 4(D). For iteration 1, 2, and 3 the frequency shift caused by application of the strand G is defined as F_g', F_g'' and F_g''' respectively (shown on Fig. 4(A), for the experiment with input 1). The setting efficiency is approximately independent of iteration number.

For all cases presented here, efficiency was calculated by averaging data over the observed plateau regions. The error bars on the graph correspond to the standard deviation of the accumulated data and are not visible in Fig. 4.

4. Discussion

4.1. Operation of the device

The data presented in the previous section confirms the operation of the immobilized OR gate. The system returns a 'true' output in the presence of

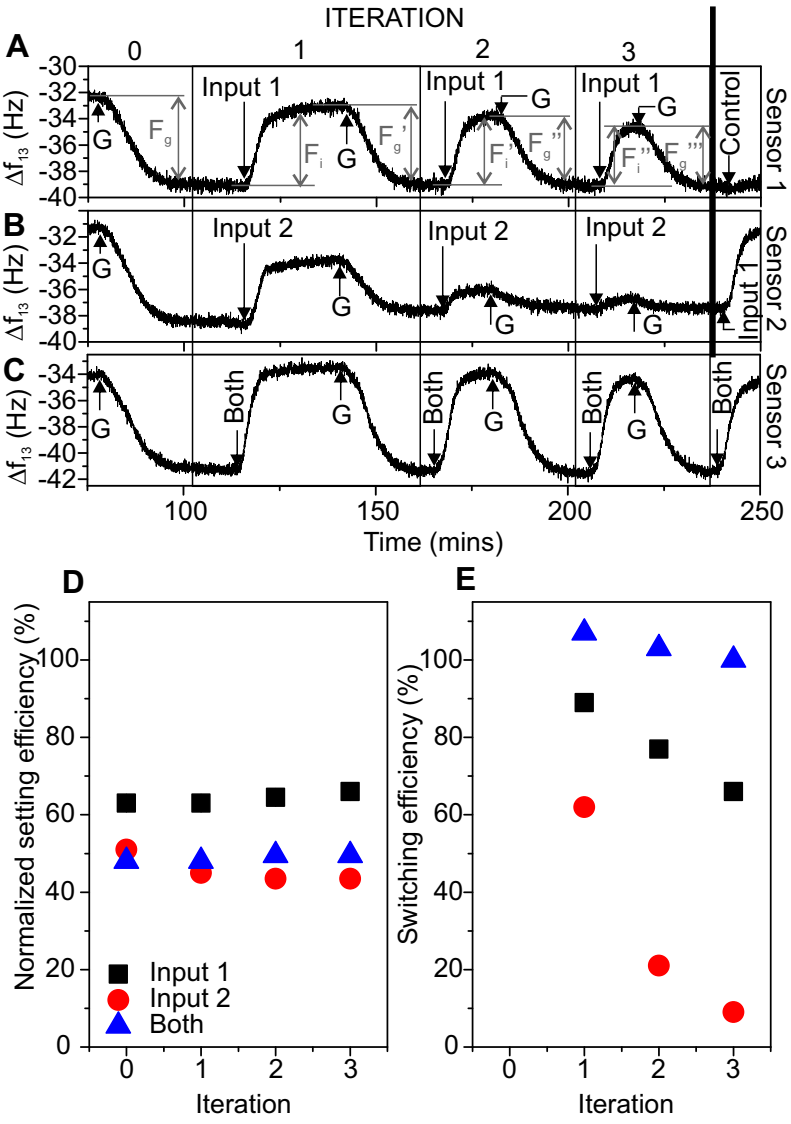


Figure 4: (A)-(C) Frequency shifts observed with QCM-D for repeated cycling of the gate, for (A) input 1 or a control (B) input 2 or input 1 (C) both inputs simultaneously. A true output results in a noticeable frequency shift after addition of the appropriate input. Arrows show the approximate times at which the indicated molecules reached the sensor surface. (D) Efficiency of setting or (E) switching the gate for initial and subsequent cycles of operation. Efficiency definitions are given in the text. For (D), the efficiency is normalized because the capture complex and the gate strand G do not have the same mass.

either input or both. It returns no output (i.e. a 'false' result) when there is no input or when a non-complementary DNA strand is supplied as an input.

Input 2 binds to a toehold at the lower end of the gate construct, close to the surface-molecular interface, and we found that the switching efficiency is lower for input 2 than for input 1. This implies that the toehold target of input 2 is less accessible than that of input 1 and that binding and/or displacement is impaired when one of the strands is buried within the layer of molecules on the surface. For input 2, the efficiency of switching decreases dramatically in the second and third cycles, much more than for input 1, and the number of gates loaded with G does not decrease significantly upon cycling. However, the gate exposed repeatedly to input 2 exhibits a strong response when input 1 is applied subsequently. This suggests that when input 2 is used, the system accumulates a population of gates that are loaded with strand G and can respond to input 1 but are unable to respond to input 2. The cumulative growth in the number of gates which do not respond to input 2 indicates that some of the binding sites for input 2 become blocked, which may occur if input 2 molecules hybridize to the gate without initiating strand displacement. It is therefore possible that both branch migration and toehold binding are suppressed as a result of interactions within the layer, where non-specific molecular interactions and steric effects inhibit the formation of the displacement complex.

The efficiency of switching is higher when both inputs are present, which may be due to the fact that the total concentration of inputs is higher in this case. This result could indicate that the selected concentration of input strands does not fully saturate the machines on the surface.

We find that the efficiency of loading the gate with strand G is only around 50%, but it does not deteriorate significantly when the gates are cycled. This is an important result because it demonstrates that the gates can be reused, a necessary feature of a component of a practical bioelectronic computer. However, it is important to note here that we assume the frequency change to be directly proportional to the mass change, which corresponds to the assumption that the Sauerbrey equation [24] is valid. This is an approximation, and it is generally true only when the layer on the sensor surface is rigid and very little energy is dissipated by the propagating acoustic wave. This is unlikely to be strictly correct in these experiments due to the viscoelasticity of the immobilized molecular layer, and our measurements of dissipation (not shown) confirm that deviations from the Sauerbrey equation should be expected. This implies that the computed 'normalized setting

efficiency' is not necessarily quantitatively equivalent to the percentage of logic gates that have been set. However, these two quantities are expected to be very strongly correlated, and it is therefore correct to infer from the data that the majority of gates are loaded and that the setting efficiency does not decrease significantly as the cycle is repeated.

When we consider the operation of dynamic surface-immobilized DNA devices, we must take care to distinguish between the behaviour of single molecules and the observed response of a molecular ensemble. The population is far from homogeneous due to the probabilistic nature of the process by which molecules associate. It will be essential to accommodate this in the design of any practical computing system based on such components.

In our experiments, we are probing an ensemble containing approximately 5×10^{11} cm^{-2} surface-immobilized OR gates (calculation given in Supplementary Information) and consequently, when we observe a change in the resonant frequency of the QCM-D sensor, a very large number of molecules have undergone a transition. With the signal-to-noise ratio observed in the experiments described here, we would not be able to detect switching if the number of gates of one type were to be reduced by more than a factor of 5. However, electronic transducers have been demonstrated that are significantly more sensitive than QCM-D [13].

4.2. Error frequency

The switching efficiency we measure for the ensemble of OR gates is typically not 100%. This implies that a significant number of the gates may remain inactive. This is particularly likely if the concentration of inputs is low. Consequently, if these gates were to be incorporated into a hybrid computer, the underlying electronics would need to be able to accommodate a random percentage of failed switching events. This suggests that significant difficulties might be encountered in the development of any architecture featuring only one individual of each species of surface-immobilized bionanomachine. The likelihood is that a radically different approach to computation would be required for such single-molecule technology.

4.3. Speed and potential for parallel computation

In our experiments, we observe that the time for the whole population of surface-immobilized OR gates to switch is approximately 6 minutes (estimated time to plateau). This implies that around 1 billion identical switching

operations occur every second per cm^2 on the surface of the sensor. It is impossible to predict when any particular individual gate will switch within the 6-minute window. If every gate were different, the system could in principle perform 5×10^{11} different operations per cm^2 in parallel over the 6-minute time frame. It is not clear if it would be possible to force the DNA gates to pack more closely together; if so, the number of simultaneous operations could be increased but the performance might be degraded as a result of unwanted intermolecular interactions and steric hindrance.

Ultimately, in the hybrid systems we envisage, it is likely that DNA machines will be immobilized on an array of microelectrodes, where each electrode is associated with a different computational process. Here, the number of computations which can be performed in parallel will be limited by the number of electrodes available, where the minimum possible size of the electrode is determined by the sensitivity of the readout circuitry and the fabrication technique employed.

In principle, the potential for parallel processing in a DNA computer is also limited by the number of possible sequences of DNA, but this number is sufficiently large that it is unlikely to prove restrictive for most applications. As a first approximation, it is given by 4^N, where N is the number of bases and this expression defines the number of different sequences that can be constructed. However, many of these strands will interact with each other, introducing unwanted 'leak' reactions. For instance, within the set of all possible sequences every strand will find its own complement, which reduces the number of useful sequences by a factor of 2. Unwanted interactions between subsequences will reduce the number of useful sequences further, and the potential for secondary structure formation will have a similar effect. These factors will become more significant as the length of the strand increases, but a large number of useful sequences will remain.

In a system containing both solution-phase DNA machines and surface-immobilized devices, the different molecular species can perform different processes at the same time. Cascades of machines can also be devised, in which the output from one computation becomes the input of the next. As we noted above, the electrode area and readout sensitivity can limit the number of processes which can be performed in parallel by surface-immobilized molecules. This limitation does not apply to the solution phase, but the great advantage of surface-immobilization is the relative ease of connecting the biomolecular machines to underlying electronic circuits.

4.4. Feature size

The transistors used in present silicon circuitry can have features as small as a few nm [14], and in the future it is expected that nanowires with a diameter of 5 nm will be used. For comparison, the diameter of our OR gate is 2 nm, equal to the diameter of double-stranded DNA, and we estimate that the spacing between individual gate molecules is approximately 14 nm. If using an ensemble of molecules the effective feature size is larger. However, if z logic gates must be switched to produce a measurable response with a particular readout system, the effective feature size could be defined as a measure of the space occupied by all z machines (i.e. $14\sqrt{z}$ nm).

It is interesting to speculate that biomolecular self-assembly may provide a route for the fabrication of ultra-nanoscale devices, which will be needed to keep pace with Moore's Law. For example, it has recently been shown that DNA origami nanostructures can be used for placement of molecules with Bohr radius resolution [11].

4.5. Power

One of the major potential advantages of biomolecular computation is the low power dissipation. To quantify this, we may consider the change in free energy associated with the strand displacement reaction that drives our OR gate. According to the online analysis package NUPACK [34], the free energy of the complex formed by strands CS, X and G is -53.17 kcal/mol, and the combined free energy of the two duplexes CS-X and G-Input 1 is -68.38 kcal/mol. Thus the free energy change of hybridization is approximately 64 kJ/mol or 10^{-19} J per molecule. We are observing approximately 5×10^{11} molecules cm^{-2}, which switch over a period of 6 minutes with a switching efficiency of 80% or better, and this therefore corresponds to power dissipation of approximately 0.1 nW cm^{-2} for the ensemble of DNA OR gates. By comparison, an array of 5×10^{11} logic transistors in a system-on-chip might dissipate approximately 375 nW of dynamic power if they all switched once over a period of 6 minutes (see Supplementary Information for the calculation and [33] for discussion of power considerations in CMOS circuits). It is clear that the use of biomolecular components can significantly reduce power consumption in comparison with conventional circuitry, potentially by three orders of magnitude.

4.6. Implementation

In the case study we present here, all logic gates execute the same operation. To fully exploit the power of a hybrid bioelectronic system, it would be necessary to incorporate multiple types of gate that perform different functions. This could be achieved by using DNA nanostructures such as origami tiles [23] to arrange specific molecular machines in a precise configuration, potentially with Bohr radius resolution [11], as described above. Here, the output released from one gate could diffuse to the next, which is separated from it by a short distance (of order tens of nm). It has been shown that optimal kinetics are achieved in such a system when the gates are separated by approximately 20 nm [30], which is slightly larger than the estimated separation in our experiments. For smaller inter-gate spacing, leak reactions occur, and at larger distances molecules do not find their targets effectively. In this case, the flow of information through the cascade of DNA gates is determined by the spatial arrangement of the components. The supply of input molecules to a hybrid system could alternatively be controlled via the solution-phase using microfluidics, with no need for co-localization of particular machines on the surface.

In an alternative approach, different DNA computational machines would be immobilized on different electrodes, as mentioned above; it is already known to be possible to immobilize DNA selectively on individually addressable electrodes separated by less than 50 nm [31], which provides much greater spatial resolution than a standard DNA microarray. Coupled with the use of DNA machines that could be switched electronically or electrochemically [22], this would enable the different gates to be connected and controlled by the underlying electronics, allowing operations to be carried out in a more sophisticated manner than in a standard biomolecular system that relies solely on interactions with solution-phase molecules.

5. Conclusion

In this paper we have presented reporter-free measurements of the operation of a surface-immobilized DNA OR gate. The OR gate is a simple machine; a range of DNA-based implementations of such a gate have been reported in the literature (e.g. [25, 10]) and it should also be noted that the OR gate is not a universal gate. However, for our purposes it represents a proof of principle, and it illustrates many of the factors which will be significant in the development of future hybrid systems based on more

complex elements or utilizing higher levels of abstraction. We have therefore used our results to discuss the fundamental attributes of molecular logic machines, with reference to their specific advantages over conventional silicon devices and their potential for integration with solid state technology. We established that the OR-gate can be reset and recycled, but the efficiency of switching after resetting can be affected by the design of the gate because processes occurring within the surface-immobilized layer can be inhibited by non-specific interactions and steric hindrance.

The results of our experiments have profound implications for the development of hybrid systems containing surface-immobilized molecular machines. It is clear that the performance of these machines can be impaired when critical components are buried within the molecular layer on the surface, and it will be necessary to take this into account when designing the configuration of the biomolecular elements of the system. However, we have also succeeded in demonstrating that the machines can be reused under certain circumstances, and this is a necessary prerequisite for their use in practical computation. Improvements in design and optimization of the density of the immobilized DNA layer could enable surface-immobilized biomolecular devices to be cycled thousands of times, mirroring the behaviour of many naturally occurring molecular machines, which are able to operate continuously and repeatedly even under the crowded conditions that exist within a living cell.

Our analysis indicates that the potential for parallel computation with DNA is greater for solution-phase processes than those involving surface-immobilized devices because the number of surface-immobilized species that can be used is much smaller than the number that can operate in a three-dimensional suspension. However, by definition, it is not possible to make an electronic connection directly to molecules floating freely in solution. This leads us to propose the scheme shown in Fig. 5 as a possible architecture for a hybrid computer, in accordance with the framework we defined in [8].

DNA machines in solution and on surfaces can be used to perform computation either sequentially, by utilizing devices in series, or in parallel, by employing different molecular species to perform separate calculations simultaneously. This is represented in our proposed architecture. On the surface, the connections between different machines in a cascade can be established by co-localizing them in fixed positions or using the underlying electronics, while in solution the overall sequence is determined by the kinetics of the underlying reactions. As discussed in Section 4.3, operations in solution may

offer greater potential for parallel processing, because surface-phase reactions are limited by the number of electrodes used or their surface area, but the motivation for the use of surface-immobilized machines is the possibility to connect them to electronics for control and readout purposes, and the significance of this should not be underestimated.

One of the greatest advantages of biomolecular logic is the low power dissipation. Our calculations indicate that the power dissipated in a switching operation can be three orders of magnitude lower in a DNA-based system than in a comparable silicon device. Also of huge interest are the possibilities for interfacing molecular machinery with biological systems. For instance, nucleic acid based computers could be used for decision-making associated with recognition or processing of RNA, such as miRNAs, which are involved in gene expression and regulation, including the control of cell differentiation [9]. DNA-protein interactions may also be relevant, and it may be appropriate to design systems that take as inputs transcription factors or biochemically active compounds, including drugs that target DNA. However, biomolecular processes can be very slow, compared to the silicon equivalent, and processes may take several minutes. This implies that silicon circuitry may still be needed to perform the time-critical elements of a computation, where high speed is required.

Although we have focussed on the use of double-stranded DNA for molecular logic, other biomolecular components could be used instead [8], including alternative DNA structures such as G-quadruplexes [4], i-motifs [19, 16] or hairpins [12]. Other classes of molecule have also been shown to have potential in this area, including peptides, and it has been demonstrated that a network of synthetic peptides is capable of performing logic functions [2]. In terms of power consumption, speed and potential for parallel computation, similar considerations apply to other biomolecular machines as to the DNA OR gate we examined in this paper.

Ultimately, the construction of hybrid bioelectronic computers, in which biomolecular components are integrated with solid state devices, will harness the potential of bionanoscience to provide new functionality, while retaining the advantages and high speed of mature semiconductor technology.

Author contributions

All authors contributed to the conception of this study and development of the ideas described. The experiments were performed by K.E.D., who also

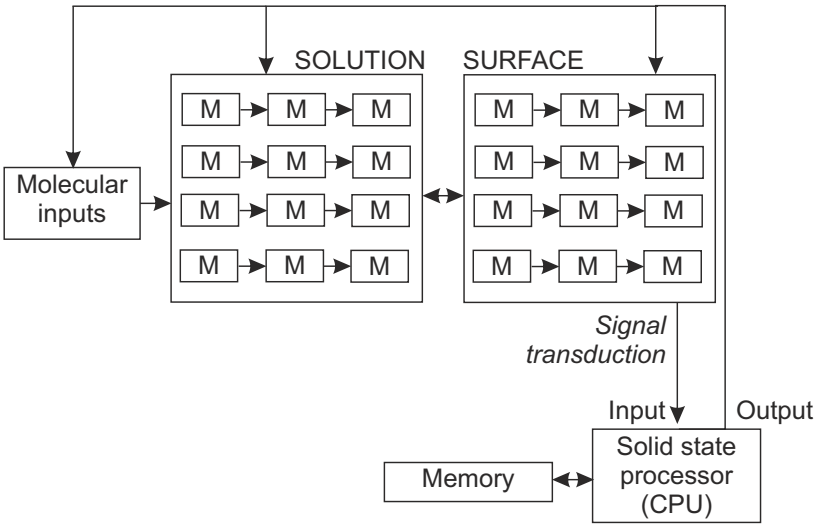


Figure 5: Possible architecture for a hybrid computer. Inputs are encoded as molecules. 'M' denotes molecular information processing - the number of processes indicated is not intended to be prescriptive and the number used is limited only by considerations of practicality. Computation can be performed either in series (by cascades of molecular machines) or in parallel (by different species of machines), either in the solution phase or on the surface. Transduction of the signal allows the output of the surface-phase computation to be communicated to a solid state processor that can communicate with a memory unit and generate outputs, which may be used to specify further molecular inputs.

analysed the data. K.E.D. wrote the paper, with contributions from all other authors. All authors have approved the final version of the manuscript.

Acknowledgements

The authors thank EPSRC for financial support (under Platform Grant EP/K040820/1), and we also thank The University of York for an Institutional Equipment Grant.

Data availability

Data generated during this research will be made available at the following DOI: 10.15124/aac407ae-9fd5-4003-b91e-416ea1df99da

Bibliography

[1] L.M. Adleman. Molecular computation of solutions to combinatorial problems. *Science*, 266(5187):1021–1024, 1994.

[2] G. Ashkenasy and M. R. Ghadiri. Boolean logic functions of a synthetic peptide network. *Journal of the American Chemical Society*, 126(36): 11140–11141, 2004.

[3] Y. Benenson, T. Paz-Elizur, R. Adar, E. Keinan, Z. Livneh, and E. Shapiro. Programmable and autonomous computing machine made of biomolecules. *Nature*, 414(6862):430–434, 2001.

[4] S. Burge, G.N. Parkinson, P. Hazel, A.K. Todd, and S Neidle. Quadruplex DNA: sequence, topology and structure. *Nucleic Acids Research*, 34(19):5402–5415, 2006.

[5] C. Costa Santini, J. Bath, A.M. Tyrrell, and A.J. Turberfield. A clocked finite state machine built from DNA. *Chemical Communications*, 49(3): 237–239, 2013.

[6] M.C. Dixon. Quartz crystal microbalance with dissipation monitoring: enabling real-time characterization of biological materials and their interactions. *Journal of Biomolecular Techniques*, 19(3):151–158, 2008.

[7] S.M. Douglas, I. Bachelet, and Church G.M. A logic-gated nanorobot for targeted transport of molecular payloads. *Science*, 335(6070):831–834, 2012.

[8] K.E. Dunn, T.L. Morgan, M.A. Trefzer, S.D. Johnson, and A.M Tyrrell. Surface-Immobilised DNA Molecular Machines for Information Processing. *Lecture Notes in Computer Science*, 9303:3–12, 2015.

[9] G.B. Fogel, T. Tallon, A.S. Wong, A.D. Lopez, and C.C. King. miRNA Regulation of Human Embryonic Stem Cell Differentiation. *Lecture Notes in Computer Science*, 9303:93–102, 2015.

[10] B. M. Frezza, S. L. Cockroft, and M. R. Ghadiri. Modular multi-level circuits from immobilized DNA-based logic gates. *Journal of the American Chemical Society*, 129(48):14875–14879, 2007.

[11] J.J. Funke and H. Dietz. Placing molecules with Bohr radius resolution using DNA origami. *Nature Nanotechnology*, 11(1):47–52, 2016.

[12] S.J. Green, D. Lubrich, and A.J. Turberfield. DNA hairpins: fuel for autonomous DNA devices. *Biophysical Journal*, 91(8):2966–2975, 2006.

[13] J. Hahm and C.M. Lieber. Direct ultrasensitive electrical detection of DNA and DNA sequence variations using nanowire nanosensors. *Nano Letters*, 4(1):51–54, 2004.

[14] International Technology Roadmap for Semiconductors. More Moore Roadmap - ITRS 2.0 White Paper, 2014. Page 7.

[15] H. Lederman, J. Macdonald, D. Stefanovic, and M.N. Stojanovic. Deoxyribozyme-based three-input logic gates and construction of a molecular full adder. *Biochemistry*, 45(4):1194–1199, 2006.

[16] T. Liedl, M. Olapinski, and F. C. Simmel. A surface-bound DNA switch driven by a chemical oscillator. *Angewandte Chemie International Edition*, 45(30):5007–10, 2006.

[17] H. Y. Lin, J. Z. Chen, H. Y. Li, and C. N. Yang. A simple three-input DNA-based system works as a full-subtractor. *Scientific Reports*, 5(10686), 2015.

[18] Q. Liu, L. Wang, A. G. Frutos, A. E. Condon, R. M. Corn, and L. M. Smith. DNA computing on surfaces. *Nature*, 403(6766):175–9, 2000.

[19] M. Guéron and J-L. Leroy. The i-motif in nucleic acids. *Current Opinion in Structural Biology*, 10(3):326–331, 2000.

[20] L. Qian and E. Winfree. Scaling Up Digital Circuit Computation with DNA Strand Displacement Cascades. *Science*, 332(6034):1196–1201, 2011.

[21] L. Qian, E. Winfree, and J Bruck. Neural network computation with DNA strand displacement cascades. *Nature*, 475(7356):368–372, 2011.

[22] S. Ranallo, A. Amodio, A. Idili, A. Porchetta, and F. Ricci. Electronic control of DNA-based nanoswitches and nanodevices. *Chemical Science*, 7(1):66–71, 2016.

[23] P.W.K Rothemund. Folding DNA to create nanoscale shapes and patterns. *Nature*, 440(7082):297–302, 2006.

[24] G. Sauerbrey. Verwendung von Schwingquarzen zur Wägung dünner Schichten und zur Mikrowägung (in German). *Zeitschrift für Physik*, 155(2):206–222, 1959.

[25] G. Seelig, D. Soloveichik, D. Y. Zhang, and E. Winfree. Enzyme-free nucleic acid logic circuits. *Science*, 314(5805):1585–1588, 2006.

[26] M. Soreni, S. Yogev, E. Kossoy, Y. Shoham, and E. Keinan. Parallel biomolecular computation on surfaces with advanced finite automata. *Journal of the American Chemical Society*, 127(11):3935–3943, 2005.

[27] N. Srinivas, T.E. Ouldridge, P. Šulc, J.M. Schaeffer, B. Yurke, A.A. Louis, J.P. K. Doye, and E. Winfree. On the biophysics and kinetics of toehold-mediated DNA strand displacement. *Nucleic Acids Research*, 41(22):10641–10658, 2013.

[28] M. N. Stojanovic, T. E. Mitchell, and D. Stefanovic. Deoxyribozyme-based logic gates. *Journal of the American Chemical Society*, 124(14): 3555–3561, 2002.

[29] X. Su and L.M. Smith. Demonstration of a universal surface DNA computer. *Nucleic Acids Research*, 32(10):3115–3123, 2004.

[30] M. Teichmann, E. Kopperger, and F.C. Simmel. Robustness of localized DNA strand displacement cascades. *ACS Nano*, 8(8):8487–8496, 2014.

[31] C. Wälti, R. Wirtz, W.A. Germishuizen, D.M.D. Bailey, M. Pepper, A.P.J. Middelberg, and A.G. Davies. Direct selective functionalization of nanometer-separated gold electrodes with DNA oligonucleotides. *Langmuir*, 19(4):981–984, 2003.

[32] L. M. Wang, Q. H. Liu, R. M. Corn, A. E. Condon, and L. M. Smith. Multiple word DNA computing on surfaces. *Journal of the American Chemical Society*, 122(31):7435–7440, 2000.

[33] N. Weste and D. Harris. *CMOS VLSI Design: A Circuits and Systems Perspective*. Addison-Wesley Publishing Company, USA, 4th edition, 2010. ISBN 0321547748, 9780321547743.

[34] J.N. Zadeh, C.D. Steenberg, J.S. Bois, B.R. Wolfe, M.B. Pierce, A.R. Khan, R.M. Dirks, and N.A. Pierce. NUPACK: Analysis and Design of Nucleic Acid Systems. *Journal of Computational Chemistry*, 32(1): 170–173, 2010.

[35] D.Y. Zhang and G. Seelig. Dynamic DNA nanotechnology using strand-displacement reactions. *Nature Chemistry*, 3(2):103–113, 2011.

[36] D.Y. Zhang and E. Winfree. Control of DNA Strand Displacement Kinetics Using Toehold Exchange. *Journal of the American Chemical Society*, 131(47):17303–17314, 2009.

Supplementary Information

Calculation of the number of OR gates on the QCM-D sensor

The average frequency shift observed upon immobilization of the capture complex is -19.7 Hz, calculated using the measurements quoted in Table 2. According to the Sauerbrey equation, the mass density is $\Delta m = -17.7\Delta f_n/n = -17.7 \times 19.7/13 \approx 26.8\,\mathrm{ng\,cm^{-2}}$.

One DNA base pair has a mass of 650 Da (i.e. 650 g $\mathrm{mol^{-1}}$), and one capture complex contains 16 base pairs and 16 unpaired nucleotides, which means that it has a mass of 15600 g $\mathrm{mol^{-1}}$. Consequently there are $26.8 \times 10^{-9}/15600 = 1.7 \times 10^{-12}$ moles $\mathrm{cm^{-2}}$ of capture complexes on the surface,

Table 2: Frequency shifts for immobilization of capture complex.

Sensor	Δf_{13} (Hz)
1	-15.84
2	-21.06
3	-22.33

which is equivalent to 1×10^{12} molecules cm^{-2}. Typically approximately half the capture complexes are loaded with gates, so the number of logic gates on the surface is approximately 5×10^{11} cm^{-2}.

Calculation of the dynamic power dissipation in a single logic transistor within a system-on-chip

This calculation is based on equations and data given in Ref. [33].

We consider a CMOS inverter driving a load capacitance C. When the pMOS transistor is switched on, the load capacitance is charged to V_{DD}. After charging, the energy stored in the capacitance is $\frac{1}{2}CV_{DD}^2$. The energy delivered by the supply is:

$$E = \int_0^\infty I(t)V_{DD}dt = \int_0^\infty C\frac{dV}{dt}V_{DD}dt = CV_{DD}\int_0^{V_{DD}} dV = CV_{DD}^2 \qquad (1)$$

Consequently, the energy dissipated in a single switching event is $\frac{1}{2}CV_{DD}^2$. For a typical digital system on chip, the total contributing capacitance of the transistor is 1.8 fF/μm, the average width of a transistor is 300 nm, and the operating voltage is 1 V. Hence, the energy dissipated is $1.8\times10^{-15}\times0.3\times1^2 = 0.27 \times 10^{-15}$ J. In the case where the transistor switches once in 6 minutes. the power dissipated is 7.5×10^{-19} W. If 5×10^{11} transistors switch, the power dissipated is therefore 375 nW.

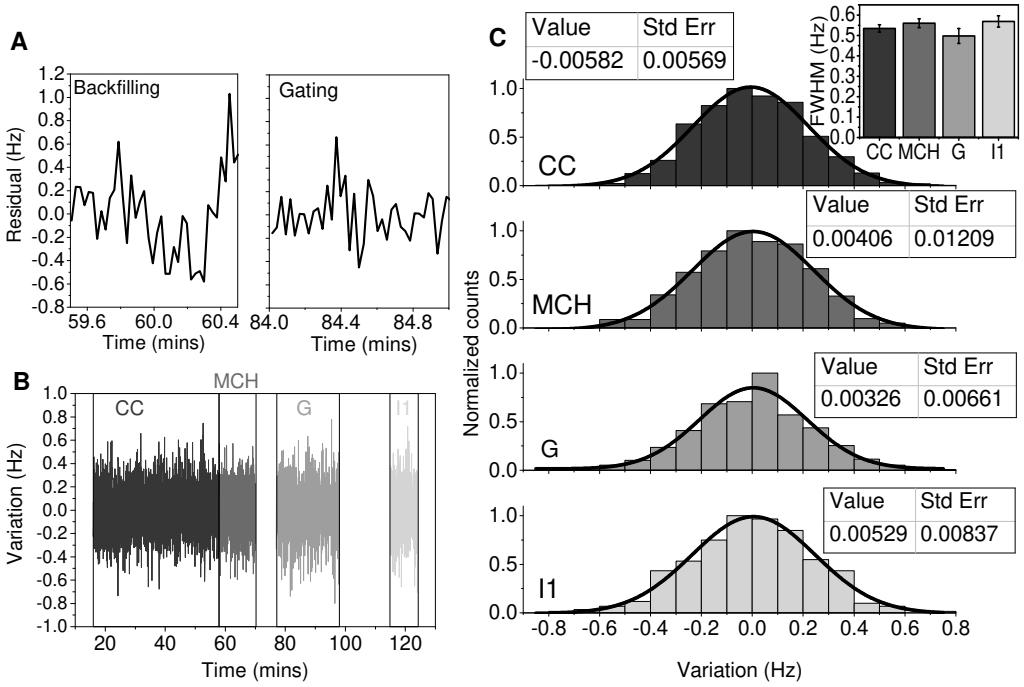
Analysis of noise in the measurement

In order to establish whether it was possible to extract any information from the noise fluctuations in the frequency measurements, we performed a detailed analysis of the noise present in the trace shown in Fig. 2A. Our results are presented in the Supplementary Figure below.

We extracted a one-minute long section of the data from the transitions corresponding to backfilling and gating. To ascertain the effect of removing

the underlying signal, we performed a linear fit and examined the residuals (the differences between the data and the fit), which are plotted in the Supplementary Figure (panel A). The backfilling step corresponds to the formation of strong bonds between the thiols of the mercaptohexanol molecules and the gold surface, while the gating step involves a very different phenomenon - hybridization of an incoming DNA molecule to a complementary region of a surface-immobilized construct. Nonetheless, the two processes exhibit very similar noise patterns, which indicates that the fluctuations are not influenced by systematic effects.

Confirmation of this is provided by the results provided in panels B and C of the Supplementary Figure. The variation is defined as the difference between the datapoints and a function derived by smoothing the data with a 10-point adjacent average filter. Panel B shows the variation as a function of time for extended sections of the data corresponding to immobilization of the capture complex, backfilling, gating and response to input. Datapoints to be included were selected manually, by visual inspection, and plateau regions are not shown. The variation is similar for all parts of the experiment, and for each part the values are plotted as a histogram in panel C, with a Gaussian fit. The histograms are practically indistinguishable - according to the fit, all are centred approximately on zero and the values of the full-width-at-half-maximum exhibit no significant differences (inset to figure, top right). Noise levels are therefore essentially constant throughout the experiment and do not provide any additional information.



Supplementary Figure: A. Residuals left after subtraction of linear fit from one-minute timespan of data from Fig. 2A for backfilling and gating transitions. B. Variation as a function of time (plateau regions not shown) for indicated transitions in the Fig. 2A dataset. Variation is defined as the difference between the data and values obtained by application of a 10-point adjacent average smoothing filter. CC: capture complex, MCH: backfilling, G: gate, I1: Input 1. C. Histograms of variation, with Gaussian fits shown as black lines. Values in boxes are fitted centres of the Gaussian functions, with associated standard errors. Inset: FWHM values obtained from fit, with associated errors.