DNA computing using single-molecule hybridization detection

Kristiane A. Schmidt[1,2,*], Christiaan V. Henkel[1,2], Grzegorz Rozenberg[2,3] and Herman P. Spaink[1,2]

[1]Institute of Biology, Leiden University, Wassenaarseweg 64, 2333 AL Leiden, The Netherlands, [2]Leiden Center for Natural Computing and [3]Leiden Institute of Advanced Computer Science, Leiden University, Niels Bohrweg 1, 2333 CA Leiden, The Netherlands

ABSTRACT

DNA computing aims at using nucleic acids for computing. Since micromolar DNA solutions can act as billions of parallel nanoprocessors, DNA computers can in theory solve optimization problems that require vast search spaces. However, the actual parallelism currently being achieved is at least a hundred million-fold lower than the number of DNA molecules used. This is due to the quantity of DNA molecules of one species that is required to produce a detectable output to the computations. In order to miniaturize the computation and considerably reduce the amount of DNA needed, we have combined DNA computing with single-molecule detection. Reliable hybridization detection was achieved at the level of single DNA molecules with fluorescence cross-correlation spectroscopy. To illustrate the use of this approach, we implemented a DNA-based computation and solved a 4-variable 4-clause instance of the computationally hard Satisfiability (SAT) problem.

INTRODUCTION

Biomolecular computing studies the use of DNA or other biomolecules for solving various computational problems (1–3). Owing to the inherent logic of DNA hybridization and the massive parallelism intrinsic to molecules, DNA computers have the potential to extend the range of solvability for computationally hard problems (2,3). Parallel search algorithms have been employed in a number of experiments for solving small-scale instances of such problems, e.g. the Hamiltonian Path problem (1) and the Satisfiability (SAT) problem (4–7). A second branch of DNA computing research investigates the development of DNA-based nanodevices. Examples are the DNA finite automata (8,9) and the realization of logic gates in single deoxyribozymes (10).

The two areas of research are related, and while they may both yield important applications, future molecular computers that combine both approaches also hold considerable promise.

For example, instead of utilizing huge amounts of electronic computer power to perform relatively simple analyses on vast quantities of biochemical information, it might be possible to construct a molecular computer, which efficiently processes these data at the molecular level.

For the successful implementation of DNA-based computations, the detection of output molecules is of prime importance. Many of the currently available techniques for the detection of DNA have been used in molecular computing: gel electrophoresis with fluorescent or radiometric visualization, fluorescent labelling and fluorescence resonance energy transfer (FRET), mass spectrometry or surface-based techniques. However, all these methods either detect DNA in bulk quantities or destroy the output molecules. This severely limits the size of the library to be searched: the largest parallel computation reported filtered 2^{20} different molecular species (7), which is less than the number of molecules of one variety necessary for the detection using gel electrophoresis (35 pg) (11). For molecular automata, this detection limit imposes an equally severe redundancy. Therefore, the application of more sensitive detection technology may significantly enhance the power of DNA computations.

Recent progress in optical detectors has enabled the efficient detection of single molecules by fluorescence microscopy (12). One of the most prominent single-molecule techniques for biological research is fluorescence-correlation spectroscopy (FCS) (13,14). FCS studies fluorescence fluctuations caused by a single molecule diffusing in the focal detection volume. Since binding of a small fluorescently labelled molecule to a larger ligand results in the change in diffusion time, FCS allows quantification of the interaction of biomolecules at extremely low concentrations. Extension of the method to dual-colour fluorescence cross-correlation spectroscopy (15) circumvents the need for a mass difference between the binding partners.

In this study, we report the detection of single molecules of DNA performing a computation. Our procedure for experimental implementation relies on the so-called blocking algorithm (16), a parallel search algorithm which involves direct inactivation of those molecules that are not a solution. Fluorescence cross-correlation spectroscopy was employed to detect hybridization between single DNA molecules. We have

*To whom correspondence should be addressed at present address: Philips Research, Personal Care Institute, Prof. Holstlaan 4 (WA11), 5656 AA Eindhoven, The Netherlands. Tel: +31 40 27 46773; Fax: +31 40 27 44288; Email: kristiane.schmidt@philips.com

benchmarked this technology on a small instance of the NP complete SAT problem.

MATERIALS AND METHODS

Sequence design

The library for a 4-variable SAT problem (2^4 possible solutions) was encoded by 16 different oligonucleotides of 36 bp each (Table 1). They have the general structure:

$5'$ start - *a b c d* - stop,

where start and stop are a leader and end sequence, CTT GCA and TTG CAC, respectively, and *a b c d* stand for the four different variables of the SAT problem. For each of these variables, identical subsequences were used to encode its value, ATC ACC for 0 (false), and GTC TGA for 1 (true). The variable which is encoded by one of these two different bit sequences is determined by its position on the DNA strand. The use of such an identical bit encoding for all variables enables us to read the bit sequence directly from the DNA sequence without the need for translation tables. For the design of the two bit sequences, the following constraints were applied: GC-content <50%, identical GC-content of both sequences, thermodynamic uniform behaviour [calculated according to the nearest-neighbour model (17)] of both sequences, <3 bp repeat within the two sequences, <3 bp complementary subsequence and no self complementarity.

The blocker encoding is designed to be complementary to the library oligonucleotides:

start - *a b c d* - stop$(3'-5')$

using the complementary subsequences (Table 1).

Table 1. Encoding scheme for library and blocker oligonucleotides

	Sequence ($5'$–$3'$)	*a b c d*	Hybridizes to
Start	CTTGCA		
Stop	TTGCAC		
Bit = 0	ATCACC		
Bit = 1	GTCTGA		
Library 00	ATCACCATCACCATCACCATCACC	0000	
Library 01	ATCACCATCACCATCACCGTCTGA	0001	Blocker D0
Library 02	ATCACCATCACCGTCTGAATCACC	0010	
Library 03	ATCACCATCACCGTCTGAGTCTGA	0011	
Library 04	ATCACCGTCTGAATCACCATCACC	0100	Blocker B0
Library 05	ATCACCGTCTGAATCACCGTCTGA	0101	
Library 06	ATCACCGTCTGAGTCTGAATCACC	0110	Blocker B1
Library 07	ATCACCGTCTGAGTCTGAGTCTGA	0111	
Library 08	GTCTGAATCACCATCACCATCACC	1000	
Library 09	GTCTGAATCACCATCACCGTCTGA	1001	Blocker C0
Library 10	GTCTGAATCACCGTCTGAATCACC	1010	Blocker A0
Library 11	GTCTGAATCACCGTCTGAGTCTGA	1011	Blocker A1
Library 12	GTCTGAGTCTGAATCACCATCACC	1100	
Library 13	GTCTGAGTCTGAATCACCGTCTGA	1101	Blocker C1
Library 14	GTCTGAGTCTGAGTCTGAATCACC	1110	
Library 15	GTCTGAGTCTGAGTCTGAGTCTGA	1111	
Blocker A0	GGTGATTCAGACGGTGATTCAGAC	1010	Library 10
Blocker A1	TCAGACTCAGACGGTGATTCAGAC	1011	Library 11
Blocker B0	GGTGATGGTGATTCAGACGGTGAT	0100	Library 04
Blocker B1	GGTGATTCAGACTCAGACGGTGAT	0110	Library 06
Blocker C0	TCAGACGGTGATGGTGATTCAGAC	1001	Library 09
Blocker C1	TCAGACGGTGATTCAGACTCAGAC	1101	Library 13
Blocker D0	TCAGACGGTGATGGTGATGGTGAT	0001	Library 01

Oligonucleotide synthesis and fluorescent labelling

Custom synthesized DNA oligonucleotides were purchased from Isogen Bioscience (Maarsen, The Netherlands) and IBA (Göttingen, Germany). Library oligonucleotides were covalently labelled with Cy5 (Amersham-Biosciences, Piscataway, NJ) at their $5'$ end (Isogen Bioscience), blocker oligonucleotides with Rhodamine Green (Molecular Probes, Leiden, The Netherlands) at their $5'$ end (IBA). Each oligonucleotide was purified by denaturing gel electrophoresis (12% polyacrylamide) to remove unbound dye as well as failure sequences (18). The purified oligonucleotides were diluted in water and their concentration was estimated from the absorption spectra, calculating the molar absorption coefficients at 260 nm according to Sambrook and Russell (18). For estimating the dye concentration, absorption coefficients of $\varepsilon_{647\,nm} = 250\,000$ cm^{-1} M^{-1} (Cy5) and $\varepsilon_{500\,nm} = 54\,000$ cm^{-1} M^{-1} (Rhodamine Green) were used.

Hybridization assay

For hybridization experiments, equal amounts of oligonucleotides were mixed at a concentration of 10 nM each. The mixture was heated to 90°C for 2 min and cooled to room temperature. For the short oligonucleotides utilized in this study, the rate of cooling was found to be irrelevant for the amount of hybridization. All experiments were performed in 1× SSC buffer (150 mM NaCl and 15 mM sodium citrate, pH 7.0) at room temperature (20°C). Sodium hydroxide was added as indicated in the text to prevent mismatch hybridization by lowering the melting temperature. Since leader and end sequences are identical for all library and blocker oligonucleotides, we assume that any effect of the fluorescent dyes on DNA duplex stability will be the same for all oligonucleotide combinations.

Theoretical concept of fluorescence autocorrelation and cross-correlation spectroscopy

FCS (13,14) was used to analyse the fluorescence intensity fluctuations originating from single fluorescent labelled DNA molecules diffusing in a confocal detection volume of <0.5 fl. Correlation of the intensity fluctuations over time yields the so-called autocorrelation function, $G(t)$. Using a one component model the experimental autocorrelation curves were fitted by:

$$G(t) = 1 + \frac{1}{N} \cdot \frac{1 - T + T \cdot e^{-t/\tau_t}}{\left(1 + \frac{t}{\tau_{diff}}\right) \cdot \sqrt{1 + \frac{t}{SP^2 \cdot \tau_{diff}}}},$$

where N denotes the number of fluorescent particles in the detection volume, T the fraction of fluorophores in triplet state, τ_t the triplet lifetime, τ_{diff} the diffusion time and SP the structural parameter describing the confocal volume.

In dual-colour cross-correlation spectroscopy (15), a sample containing two fluorescent species labelled with two different dyes is excited simultaneously with two laser lines, and the fluorescence signals from the two dyes are separately detected. In addition to the autocorrelation functions, the cross-correlation function, $G_{cc}(t)$, is calculated. The latter was fitted

according to:

$$G_{cc}(t) = 1 + \frac{1}{N_x} \cdot \frac{1}{\left(1 + \frac{t}{\tau_{diff,gr}}\right) \cdot \sqrt{1 + \frac{t}{SP^2 \cdot \tau_{diff,gr}}}}.$$

This equation uses an apparent particle number, N_x, which is the inverse amplitude of the cross-correlation function:

$$N_x = \frac{1}{G_{cc}(0) - 1}.$$

The diffusion behaviour of the doubly labelled molecules is described by $\tau_{diff,gr}$. Assuming that there is no cross-talk between the two detection channels (15), the number of doubly labelled particles, N_{cc}, may be determined from:

$$N_{cc} = \frac{N_{ac,r} \cdot N_{ac,g}}{N_x}.$$

Here, $N_{ac,r}$ and $N_{ac,g}$ are the particle numbers obtained from the autocorrelation functions detected in the red and the green detection channel.

Cross-correlation measurements

Dual-colour fluorescence cross-correlation measurements were performed on a ConfoCor2 (Zeiss, Jena, Germany), using the cross-correlation configuration (15,19). Dual-colour excitation was achieved with the 488 and 633 nm laser lines. The fluorescence emission was split by a dichroic mirror (secondary beam splitter 635 nm), passed through a 505–550 nm bandpass and a 650 nm longpass filter, respectively, and recorded in two separate channels. Detection of fluorescence fluctuations was achieved with two avalanche photodiodes. The signals were software-correlated to obtain the autocorrelation and cross-correlation functions. Calibration measurements with standard dyes (Rhodamine6G and Cy5) were performed to determine the geometry and the size of the detection volumes in both channels. Typically, volumes of 0.12 and 0.43 fl were found for the green and red channel, respectively. To maximize the overlap of the two detection volumes, pinhole alignment was performed as described by Bacia *et al*. (19). All measurements were carried out in 10 µl vol in 8-well glass bottom chambers (Nunc GmbH, Wiesbaden, Germany). For each cross-correlation curve, five individual measurements (30 s each) were averaged. The ConfoCor2 software was used for fitting autocorrelation and cross-correlation curves and to calculate the average number of doubly labelled molecules in the detection volume.

RESULTS

DNA computation

SAT problems have frequently been tackled by DNA-based computations and may be considered a benchmark for new algorithms (4–7). The specific SAT problem solved here is a 4-variable three SAT problem consisting of four clauses:

$$F = (b \lor c \lor \sim d) \,\&\, (\sim a \lor b \lor \sim c) \,\&\, (a \lor \sim b \lor d) \,\&\, (\sim a \lor c \lor \sim d),$$

where a, b, c and d are the four variables with values of true (1) or false (0). OR operations are denoted by '$\lor$', AND operations by '&', while '$\sim$' symbolizes the negation of a variable. Since the clauses are connected by AND, falsifying one clause is sufficient for falsifying the complete formula.

Our experimental algorithm (16) works as follows: first, all library molecules are synthesized, i.e. a mixture of single-stranded (ss) DNA molecules (oligonucleotides) encoding all candidate solutions for a given problem. Then, a set of so-called blocker oligonucleotides is created which encode the falsifying assignments for each of the clauses. These falsifiers or blockers are used to block those library molecules that are not a solution. Their sequence is chosen to be complementary to the corresponding library molecules. Addition of the blockers to the library molecules results in the hybridization of blockers to all 'wrong' assignments. Hence, the remaining ss DNA molecules are those that do represent a solution. Double-stranded (ds) DNA, i.e. the non-satisfying assignments, is recognized using hybridization detection as output. In order to obtain an addressed array for the output, the library molecules (2^n for a problem with n variables) may either be immobilized on a surface (DNA chip) or distributed in 2^n different tubes.

A number of different techniques were previously tested for the experimental implementation of the blocking algorithm (20,21). However, all these methods, a hybridization assay based on FRET, a heteroduplex migration assay and enzymatic mismatch recognition (all employing oligonucleotide concentrations in the micromolar range), did not allow clear discrimination of all satisfying solutions and all non-satisfying ones.

This study reports the utilization of single-molecule fluorescence spectroscopy for hybridization detection. First, the development of an assay for hybridization detection at the level of single molecules will be described. Optimization of the experimental conditions allowed hybridization detection with the addition of multiple blocker oligonucleotides in parallel, i.e. at the same time. This approach was used to solve the 4-variable 4-clause SAT problem described above.

Hybridization detection at the level of single DNA molecules

In order to detect hybridization of single DNA molecules, we applied dual-colour fluorescence cross-correlation spectroscopy. For this purpose, the library molecules were covalently 5' labelled with a red fluorescent dye (Cy5) and blockers with a green fluorescent dye (Rhodamine Green). Hybridization of a blocker molecule to a library molecule results in the formation of a dsDNA molecule which is labelled with both the red and the green dye. Since the cross-correlation function contains only dynamic information about doubly labelled molecules (15), a cross-correlation signal is not detected unless the formation of dsDNA, i.e. hybridization, occurred. Figure 1a depicts two different cross-correlation experiments using two different library molecules (04 and 05) to which the same blocker molecule (B0) was added. Blocker B0 and
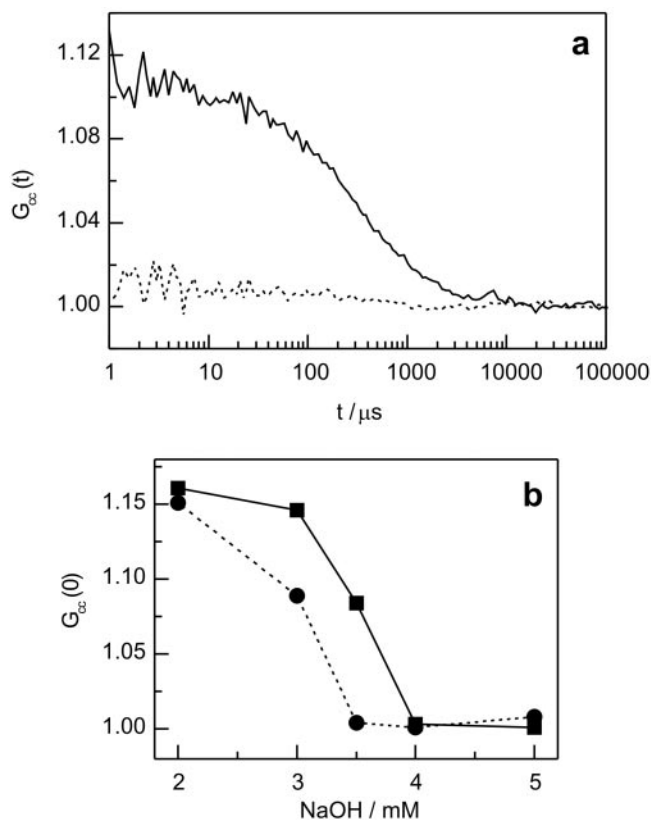
Figure 1. Hybridization detection with single molecules. (**a**) Cross-correlation curves of two different combinations of library and blocker oligonucleotides. Solid line: library 04 and blocker B0, perfect match. Dashed line: library 05 and blocker B0, mismatch of 4 bp. All oligonucleotide concentrations were 10 nM (in 1× SSC buffer). An aliquot of 3.5 mM NaOH was added for permissive conditions. Both traces are the average of five individual measurements. (**b**) Effect of NaOH concentration on the initial amplitude of the cross-correlation curve, $G_{cc}(0)$. Squares: library 04 and blocker B0, perfect match. Circles: library 05 and blocker B0, mismatch of 4 bp.

Figure 2. Addition of multiple blocker oligonucleotides in parallel. (**a**) Cross-correlation curves for hybridization with two blockers. Solid line: library 11 plus blockers A0 and A1 (perfect match). Dashed line: library 03 plus blockers A0 and A1. (**b**) Cross-correlation curves for hybridization with three blockers. Solid line: library 04 plus blockers B0 (perfect match), B1 and C1. Dashed line: library 05 plus blockers B0, B1 and C1. All oligonucleotide concentrations were 10 nM (in 1× SSC buffer with 3.5 mM NaOH).

library 04 form a perfect duplex, whereas hybridization with library 05 results in four mismatched basepairs. As expected, a cross-correlation signal was only observed for the perfect duplex (blocking combination); the amplitude detected for the 4 bp mismatch combination is nearly 0, $G_{cc}(t) = 1$. However, this clear difference was only observed under stringent conditions. At room temperature, without the addition of denaturing chemicals (sodium hydroxide, urea and formamide), the amplitude of the signal for the mismatch combination was considerably higher, indicating the occurrence of mismatch hybridization. Sodium hydroxide (or more general alkaline pH) lowers the thermal melting point of DNA (22). The effect of NaOH on hybridization is shown for both combinations in Figure 1b. The largest difference between the amplitudes for the perfect duplex and the 4 bp mismatch was observed after the addition of 3.5 mM NaOH, indicating that this concentration lowers the melting temperature of the mismatch combinations to <20°C. Therefore, in all subsequent experiments 3.5 mM NaOH was added. It is worth mentioning that the optimal amount of NaOH needed for the discrimination between perfect duplex and mismatch hybridization depends on the buffer composition, the temperature as well
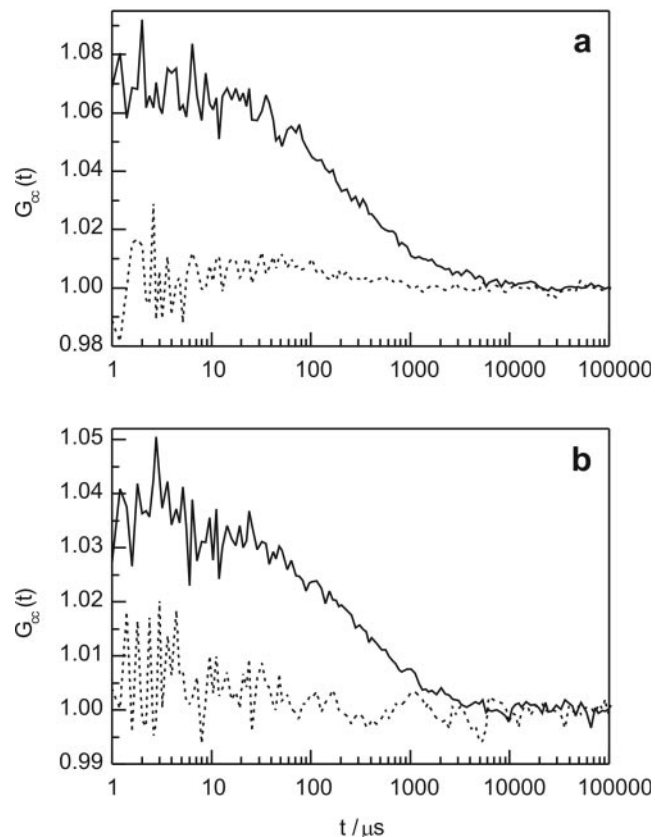
as the length and sequence of the oligonucleotides (data not shown).

Addition of multiple blockers in parallel

An important goal for the experimental implementation of the blocking algorithm is the addition of multiple blockers in parallel. Figure 2a shows the hybridization experiments with the addition of two different blocker oligonucleotides at the same time. Again, two different experiments using two different library molecules are compared. Library 11 forms a perfect duplex with blocker A1 and a 4 bp mismatch with blocker A0, whereas library 03 has a mismatch of 4 and 8 bp with blocker A0 and A1, respectively. Similar to the experiments with just one blocker added, a cross-correlation signal was only observed for the perfect duplex but not for the mismatch combination. Similar experiments using three blockers in parallel are depicted in Figure 2b. Library 04 perfectly matches blocker B0 and has mismatches of 4 and 8 bp with blocker B1 and C1, respectively. Library 05 has a mismatch of 4 bp with blockers B0 and C1, and of 8 bp with blocker B1. Again, a cross-correlation signal is only detected for the blocking combination (i.e. the perfect duplex) and the amplitude

observed for the mismatch combination is close to 0. These experiments demonstrate that the cross-correlation technique is suited for hybridization detection with the addition of multiple blocker molecules at the same time, i.e. a high background of competing oligonucleotides. However, our data indicate that the amplitude of the signal for the perfect match decreases with an increasing number of blocker oligonucleotides added. In agreement with this observation, the amplitudes observed in experiments using four different blocker molecules in parallel were too low to achieve reliable hybridization detection (data not shown).

Complete computation

To test the applicability of the single-molecule approach for a complete computation, the first clause of the 4-clause SAT problem described above was used: $(b \lor c \lor \sim d)$, which is falsified by $[a\,b\,c\,d] = 0001 \lor 1001$, encoded by blockers C0 and D0. In order to compare the amount of hybridization, the average number of doubly labelled molecules in the detection volume, N_{cc}, was determined from the experimental cross-correlation curves. Figure 3 shows the results for the 16 different library molecules. Significant numbers of doubly labelled molecules, meaning hybridization with blockers, were only detected for library molecules 01 and 09. Library 01 forms a perfect duplex with blocker D0, while library 09 perfectly matches to blocker C0. Independent experiments with the same blocker and library molecules indicate a rather good reproducibility (see error bars in Figure 3) and reliable distinction of blocking versus non-blocking combinations.

The second clause of formula F, $(\sim a \lor b \lor \sim c)$, is falsified by blockers A0 and A1. The corresponding experiment is depicted in Figure 4a. Again, a significant amount of doubly labelled molecules was only observed for two combinations. Library 10 and 11 form perfect matches with blockers A0 and A1, respectively. Figure 4b demonstrates the experiment for
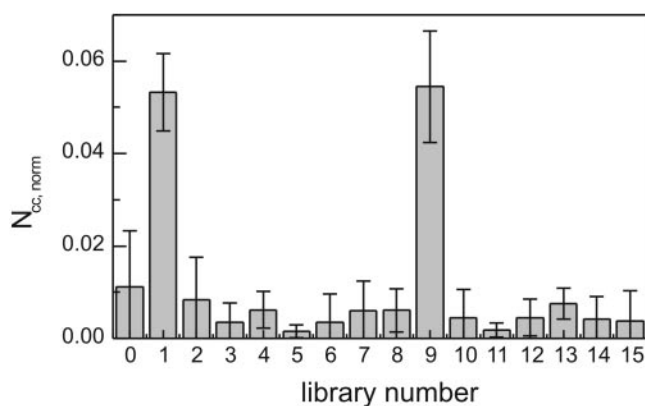
solving the third and fourth clause of the problem, $(a \lor \sim b \lor d)$ & $(\sim a \lor c \lor \sim d)$. This conjunction is falsified by blockers B0, B1, C0 and C1. Since blocker C0 was already used in the first experiment, i.e. the corresponding library molecule was already sorted out, it may be omitted from the experiment. Experiments with the three other blockers resulted in high cross-correlation signals, i.e. considerable number of doubly labelled molecules, for library molecules 04, 06 and 13, each of which forms a perfect duplex with one of the blockers.

To solve the complete 4-clause SAT problem, we just have to combine the results from the three previous experiments. Every library molecule that shows a cross-correlation signal in one of the experiments, meaning that it hybridizes to one of the blockers, is identified as wrong assignment. Figure 5 shows that these are library molecules 01, 04, 06, 09, 10, 11 and 13. Our SAT problem is therefore satisfied by the remaining library molecules 00, 02, 03, 05, 07, 08, 12, 14 and 15 (see Table 1 for encoding).
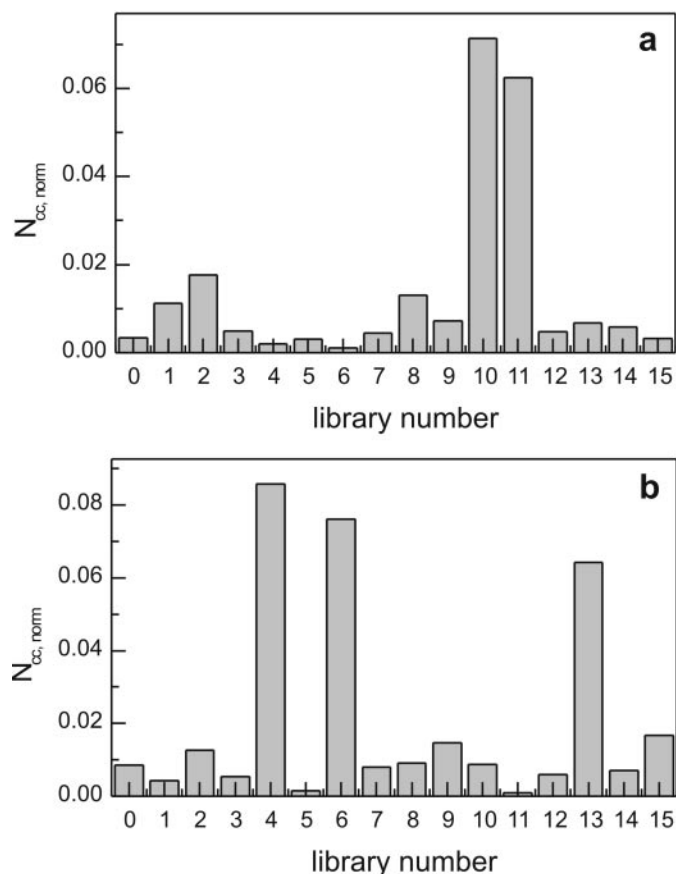
Figure 3. Solution of the first clause of a 4-clause SAT problem: $F = (b \lor c \lor \sim d)$, falsified by $[a\,b\,c\,d] = 0001 \lor 1001$. Cross-correlation curves were measured for the 16 different library oligonucleotides after hybridization with blocker oligonucleotides C0 and D0. The average number of doubly labelled molecules, N_{cc}, was determined by fitting the average of five cross-correlation curves for each library/blocker combination. In order to compare experiments from different measurement series, all values were normalized for the same number of blocker molecules, yielding $N_{cc,norm}$. To test the reproducibility of the approach, the results from three to four individual experiments were averaged; the error bars give the SD.

Figure 4. Solution of the second, third and fourth clause of a 4-clause SAT problem. (**a**) Second clause, $F = (\sim a \lor b \lor \sim c)$, falsified by $1010 \lor 1011$. Cross-correlation curves were measured after hybridization with blockers A0 and A1. (**b**) Third and fourth clause of the problem, $F = (a \lor \sim b \lor d)$ & $(\sim a \lor c \lor \sim d)$, falsified by $0100 \lor 0110 \lor 1001 \lor 1101$. Cross-correlation curves were measured after hybridization with blockers B0, B1 and C1. Blocker C0 was omitted from the experiment since it was already used for falsifying the first clause. Again, the results of three to four individual experiments were averaged (error bars not shown). The average number of doubly labelled molecules was determined as described for Figure 3.
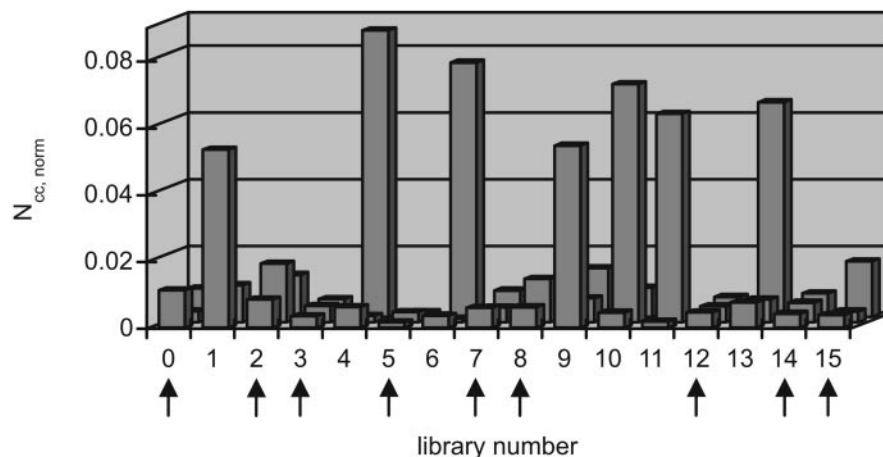
Figure 5. Solution of a 4-clause 4-variable SAT problem, $F = (b \vee c \vee \sim d)$ & $(\sim a \vee b \vee \sim c)$ & $(a \vee \sim b \vee d)$ & $(\sim a \vee c \vee \sim d)$. Summary of the three experiments depicted in Figures 3 and 4. The arrows indicate those library molecules encoding those assignments which satisfy the problem.

DISCUSSION

The potential of single-molecule techniques was already recognized immediately following the first experiments in DNA computing (23,24). Adleman (23) proposed that the use of single-molecule fluorescence spectroscopy could enable DNA detection without previous amplification. Our study presents the first example for the utilization of single-molecule techniques for DNA computing. Fluorescence cross-correlation spectroscopy was employed to monitor a DNA computation at the level of single molecules. Given that even single basepair mismatch discrimination can be achieved (data not shown), the technique may also prove to be useful for medical applications, e.g. mismatch detection assays or detection of gene expression in living cells (25).

The single-molecule hybridization assay proved to be a considerable improvement compared to the ensemble measurement by the FRET assay tested before (20). Apparently, fluorescence cross-correlation spectroscopy is less prone to experimental errors, most probably because the signal amplitude does not depend on the distance between the two dyes. In this view, hybridization detection would be less susceptible to small differences in DNA structure [which may be caused for example by interaction between nucleobases and the dye molecules (26)]. In addition, the method can also be applied for oligonucleotides longer than the Förster distance ($\sim$70 Å) of the two dyes.

The DNA computation described here compares the hybridization behaviour of 112 different oligonucleotide combinations (Table 1). For oligonucleotide sequence design nearest-neighbour thermodynamic parameters were employed (17), which were derived from the ultraviolet absorption measurements using micromoles of oligonucleotides (27). Nonetheless, we could remarkably well predict the hybridization behaviour of all 112 different oligonucleotide combinations in the single-molecule hybridization assay. Moreover, our experiments demonstrate that chemical melting with alkaline pH corresponds very well to thermal melting.

The largest DNA computation reported up to now is the solution of a 20-variable SAT problem (7) using gel electrophoresis techniques. Our 4-variable test problem is rather small, but in contrast to all previous approaches, DNA computing using single-molecule detection requires neither large quantities of DNA for detection (4–7) nor does it destroy the output molecules (4–7). Therefore, the method has a considerable promise for extending the size of the libraries to be searched and thereby enhance the performance of parallel search algorithms.

Our approach for solving SAT is the first experimental implementation of the blocking algorithm described by Rozenberg and Spaink (16). Opposed to other parallel search algorithms, this methodology sorts out those molecules that encode wrong assignments. In theory, only three experimental steps are required: DNA synthesis, hybridization and detection of ssDNA. Our actual computation involved seven steps, DNA synthesis plus three times hybridization and detection of ssDNA. Currently, we are investigating different algorithms and technologies for up scaling the computation. Single-molecule fluorescence spectroscopy will be combined with lab-on-a-chip technology to enhance the parallelism of the computation, implying that a large number of hybridization experiments can be performed at the same time. Sophisticated temperature controlled microfluidics systems may be used to implement high-throughput hybridization detection and possibly will allow the addition of more than three blockers in parallel. In addition, the utilization of universal nucleotides may decrease the number of blocker molecules needed for the computation (28).

For practical reasons, the experiments described in this study were performed in a volume of 10 μl. Since the detection volume of our setup is <0.5 fl, the amount of DNA still may be very much reduced. State-of-the-art spotting techniques used for lab-on-a-chip technologies enable dispensing of 60 pl in nanolitre wells (29), which would decrease the amount of DNA needed for our computation to as little as 0.6 attomol of each species. Compared with gel electrophoresis (11) which is the most commonly applied detection method for DNA computing, a reduction by four orders of magnitude can be achieved. Thus, the search space, that is the size of the library to be searched, may be increased 10 000 times. Hybridization detection by single-molecule fluorescence spectroscopy is also much faster than gel electrophoresis. All experiments

for solving our 4-clause SAT problem can be performed in the laboratory in just one single day. FCS can easily be combined with high-throughput screening, meaning that the speed of the computation can be further increased by automation of the procedure and even larger problems may be tackled with affordable computing time.

The combination of single-molecule fluorescence spectroscopy with lab-on-a-chip technology appears to be especially attractive for evolutionary DNA computing, i.e. the implementation of evolutionary algorithms with DNA (30,31). So far, no successful experimental realization of this approach has been reported. In particular, the selection step is difficult to implement: selection of the fittest molecules requires very accurate detection (and even sorting) of low quantities of molecules with desirable properties, while a very high background of 'wrong' molecules is present at the same time. One can envisage that single-molecule technology will enable selection and manipulation of single desirable molecules for this purpose. Furthermore, the non-destructive character of single-molecule detection allows for iterated selection cycles.

In summary, the application of single-molecule techniques has the potential to overcome some of the current limitations of DNA computing and to extend the scale of computations from proof-of-principle towards real applications.

ACKNOWLEDGEMENTS

REFERENCES

1. Adleman,L.M. (1994) Molecular computation of solutions to combinatorial problems. *Science*, **266**, 1021–1024.
2. Paun,G., Rozenberg,G. and Salomaa,A. (1998) *DNA Computing: New Computing Paradigms*. Springer-Verlag, Heidelberg, Germany.
3. Hagiya,M. and Ohuchi,A. (eds) (2003) *DNA Computing*. In *8th International Workshop on DNA-based Computers, DNA8*, Sapporo, Japan. Springer-Verlag, Heidelberg, Germany.
4. Faulhammer,D., Cukras,A.R., Lipton,R.J. and Landweber,L.F. (2000) Molecular computation: RNA solutions to chess problems. *Proc. Natl Acad. Sci. USA*, **97**, 1385–1389.
5. Liu,Q.H., Wang,L.M., Frutos,A.G., Condon,A.E., Corn,R.M. and Smith,L.M. (2000) DNA computing on surfaces. *Nature*, **403**, 175–179.
6. Sakamoto,K., Gouzu,H., Komiya,K., Kiga,D., Yokoyama,S., Yokomori,T. and Hagiya,M. (2000) Molecular computation by DNA hairpin formation. *Science*, **288**, 1223–1226.
7. Braich,R.S., Chelyapov,N., Johnson,C., Rothemund,P.W.K. and Adleman,L. (2002) Solution of a 20-variable 3-SAT problem on a DNA computer. *Science*, **296**, 499–502.
8. Benenson,Y., Paz,E.T., Adar,R., Keinan,E., Livneh,Z. and Shapiro,E. (2001) Programmable and autonomous computing machine made of biomolecules. *Nature*, **414**, 430–434.
9. Benenson,Y., Gil,B., Ben-Dor,U., Adar,R. and Shapiro,E. (2004) An autonomous molecular computer for logical control of gene expression. *Nature*, **429**, 423–429.
10. Stojanovic,M.N. and Stefanovic,D. (2003) A deoxyribozyme-based molecular automaton. *Nat. Biotechnol.*, **21**, 1069–1074.
11. Tuma,R.S., Beaudet,M.P., Jin,X., Jones,L.J., Cheung,C.Y., Yue,S. and Singer,V.L. (1999) Characterization of SYBR gold nucleic acid gel stain: a dye optimized for use with 300-nm ultraviolet transilluminators. *Anal. Biochem.*, **268**, 278–288.
12. Weiss,S. (1999) Fluorescence spectroscopy of single biomolecules. *Science*, **283**, 1676–1683.
13. Madge,D., Webb,W.W. and Elson,E. (1972) Thermodynamic fluctuations in a reacting system – measurement by fluorescence correlation spectroscopy. *Phys. Rev. Lett.*, **29**, 705–708.
14. Eigen,M. and Rigler,R. (1994) Sorting single molecules: application to diagnostics and evolutionary biotechnology. *Proc. Natl Acad. Sci. USA*, **91**, 5740–5747.
15. Schwille,P., Meyer-Almes,F.-J. and Rigler,R. (1997) Dual-color fluorescence cross-correlation spectroscopy for multicomponent diffusional analysis in solution. *Biophys. J.*, **72**, 1878–1886.
16. Rozenberg,G. and Spaink,H. (2003) DNA computing by blocking. *Theor. Comp. Sci.*, **292**, 653–665.
17. SantaLucia,J.,Jr (1998) A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. *Proc. Natl Acad. Sci. USA*, **95**, 1460–1465.
18. Sambrook,J. and Russell,D.W. (2001) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
19. Bacia,K., Majoul,I.V. and Schwille,P. (2002) Probing the endocytic pathway in live cells using dual-color fluorescence cross-correlation analysis. *Biophys. J.*, **83**, 1184–1193.
20. Schmidt,K.A., Henkel,C.V., Rozenberg,G. and Spaink,H.P. (2002) Experimental aspects of DNA computing by blocking: use of fluorescence techniques for detection. In Kraayenhof,R., Visser,A.J.W.G. and Gerritsen,H.C. (eds), *Fluorescence Spectroscopy, Imaging and Probes—New Tools in Chemical, Physical and Life Science*. Springer-Verlag, Heidelberg, Germany, pp. 123–128.
21. Henkel,C.V., Rozenberg,G. and Spaink,H.P. (2004) Application of mismatch detection methods in DNA computing. In Ferreti,C., Mauri,G. and Zandron,C. (eds), *DNA10, 10th International Meeting on DNA Computing, Preliminary Proceedings*. Università di Milano-Bicocca, pp. 183–192.
22. Williams,M.C., Wenner,J.R., Rouzina,I. and Bloomfield,V.A. (2001) Effect of pH on the overstretching transition of double-stranded DNA: evidence of force-induced DNA melting. *Biophys. J.*, **80**, 874–881.
23. Adleman,L.M. (1996) On constructing a molecular computer. In Lipton,R.J. and Baum,E. (eds), *DNA Based Computers. DIMACS 27*. American Mathematical Society, pp. 1–21.
24. Reif,J.H. (1998) Paradigms for biomolecular computation. In Calude,C.S., Casti,J. and Dinneen,M.J. (eds), *Unconventional Models of Computation*. Springer-Verlag, Heidelberg, Germany, pp. 72–93.
25. Tsuji,A., Koshimoto,H., Sato,Y., Hirano,M., Sei-lida,Y., Kondo,S. and Ishibashi,K. (2000) Direct observation of specific messenger RNA in a single living cell under a fluorescence microscope. *Biophys. J.*, **78**, 3260–3274.
26. Marras,S.A.E., Kramer,F.R. and Tyagi,S. (2002) Efficiencies of fluorescence resonance energy transfer and contact-mediated quenching in oligonucleotide probes. *Nucleic Acids Res.*, **30**, e122.
27. SantaLucia,J.,Jr, Allawi,H.T. and Seneviratne,P.A. (1996) Improved nearest-neighbor parameters for predicting DNA duplex stability. *Biochemistry*, **35**, 3555–3562.
28. Loakes,D. (2001) Survey and summary: the applications of universal DNA base analogues. *Nucleic Acids Res.*, **29**, 2437–2447.
29. Young,I.T., Moerman,R., van den Doel,L.R., Iordanoc,V., Kroon,A., Dietrich,H.R.C., Dedem,G.W.K., Bossche,A., Gray,B.L., Sarro,L. *et al.* (2003) Monitoring enzymatic reactions in nanoliter wells. *J. Microsc.*, **212**, 254–267.
30. Bäck,T., Kok,J.N. and Rozenberg,G. (2003) Evolutionary computation as a paradigm for DNA-based computing. In Landweber,L.F. and Winfree,E. (eds), *Evolution as Computation. DIMACS workshop*, Princeton, January 1999. Springer-Verlag, Heidelberg, Germany, pp. 15–40.
31. Chen,J.H. and Wood,D.H. (2000) Computation with biomolecules. *Proc. Natl Acad. Sci. USA*, **97**, 1328–1330.