Molecular logic computing model based on DNA self-assembly strand branch migration

ARTICLE *in* CHINESE SCIENCE BULLETIN · JANUARY 2013

Impact Factor: 1.58 · DOI: 10.1007/s11434-012-5498-z

5 AUTHORS, INCLUDING:

Cheng Zhang
Peking University
25 PUBLICATIONS **74** CITATIONS
SEE PROFILE

Yafei Dong
Shaanxi Normal University
18 PUBLICATIONS **55** CITATIONS
SEE PROFILE

Jing Yang
North China Electric Power University
26 PUBLICATIONS **74** CITATIONS
SEE PROFILE

Article

Chinese Science Bulletin

Progress of Project Supported by NSFC
Computer Science & Technology

January 2013 Vol.58 No.1: 32–38
doi: 10.1007/s11434-012-5498-z

Molecular logic computing model based on DNA self-assembly strand branch migration

ZHANG Cheng[1†], MA LiNa[2†], DONG YaFei[2], YANG Jing[3*] & XU Jin[1*]

[1] *Institute of Software, School of Electronics Engineering and Computer Science, Key Laboratory of High Confidence Software Technologies of Ministry of Education, Peking University, Beijing 100871, China;*
[2] *College of Life Science, Shaanxi Normal University, Xi'an 710062, China;*
[3] *School of Control and Computer Engineering, North China Electric Power University, Beijing 102206, China*

Received June 1, 2012; accepted September 4, 2012

In this study, the DNA logic computing model is established based on the methods of DNA self-assembly and strand branch migration. By adding the signal strands, the preprogrammed signals are released with the disintegrating of initial assembly structures. Then, the computing results are able to be detected by gel electrophoresis. The whole process is controlled automatically and parallely, even triggered by the mixture of input signals. In addition, the conception of single polar and bipolar is introduced into system designing, which leads to synchronization and modularization. Recognizing the specific signal DNA strands, the computing model gives all correct results by gel experiment.

DNA strand branch migration, DNA self-assembly, molecular logic computing, molecular intelligence, modularized design

Citation: Zhang C, Ma L N, Dong Y F, et al. Molecular logic computing model based on DNA self-assembly strand branch migration. Chin Sci Bull, 2013, 58: 32–38, doi: 10.1007/s11434-012-5498-z

In recent years, with the approaching of the Moore's Law, the pressure of traditional electronic computers increased greatly for handling mass information. The interests of researchers have been attracted to the area of novel computing. Taking advantage of some new methods of quantum and molecular computing, researchers attempt to use nano-materials and technologies to implement super large scale information processing. In particular, DNA computing has become a research focus in molecular computing, combined with information science, biology and nanotechnology [1–5]. Because the DNA molecules have lots of natural advantages in huge parallelism and microscopic, the mass parallel information processing could be achieved by using DNA computing. Thus, DNA computing may become a flourishing route in future computing. In fact, since the end of last century, it has made a great progress, both in theoretical models and experimental operations of DNA computing. In addition, it has developed greatly in the interdisciplinary fields of information processing, molecular encoding, nanomachines and so on [6–11].

In the development of DNA computing, a variety of molecular operations have been utilized such as polymerase chain reaction (PCR), fluorescence techniques, strand branch migration and self assembly techniques. In these methods, DNA strand branch migration with fluorescent labeling develops rapidly for constructing various molecular computing models [12–15]. Importantly, Professor Winfree used DNA strand branch migration to implement simple square-root computing and neural networks computing in 2011. These works were reported in *Science* and *Nature*, and attracted lots of attentions from researchers in the field of information computing [4,5]. DNA strand branch migration is able to be combined with not only fluorescent detecting and DNA self-assembly, but also nano-particles, quantum dots and proteins. Moreover, it has promoted the development of research fields as parallel computing, cryptography and nanoelectronics.

†These authors contributed equally to the work.
*Corresponding authors (email: yangjing369@gmail.com; jxu@pku.edu.cn)

In this paper, a DNA logic computing model was constructed using strand branch migration, combined with DNA self-assembly technology. In general, this model was divided into two parts: DNA self-assembly module and controller based on strand branch migration. In the process of computing, basic computing units were constructed using DNA hybridization to form some molecule self-assembly structures. Through the input of specific DNA signals, basic computing units recognized signals, and then released specific output signal molecules by strand branch migration. In this model, there were several characteristics such as parallel information processing and modular assembly. In particular, using DNA molecule branch migration, precise computing results were able to be obtained.

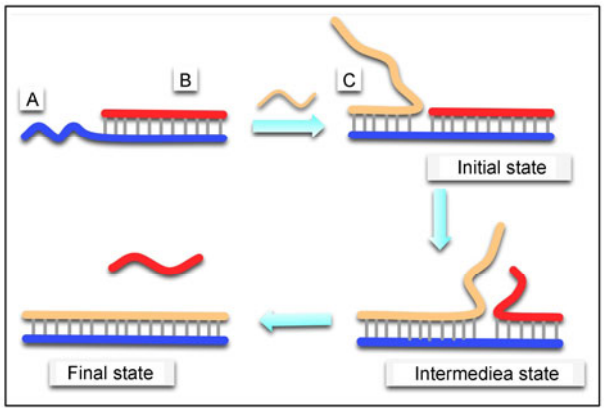

Figure 1 The principle of DNA strand branch migration.

1 The backgrounds

1.1 DNA self-assembly

DNA self-assembly structure is formed by hydrogen bond and van der waals force between base pairs. According to this principle, simple DNA motifs are assembled to complex and hierarchical structures. The whole process is an automatic course from simple to complex, from disorder to well-regulated. In 1960s, the idea of using DNA single strands to assemble tile structures was proposed. Until 1990s, the idea of DNA self-assembly had been realized by professor Seeman, which opened a new door for this area. His group constructed many kinds of complex assembly structures as quadrangle, loop, and knot [16–18]. Because of these researches, Seeman owned Kavli award of nanotechnology in 2010. Another method of DNA self-assembly is using a long DNA single strand (M13 bacteriophage) to form structures. In this way, the yield of self-assembly structures is very high, only with rather lower cost. In 2006, Paul Rothemund folded M13 DNA strand and short oligonucleotide staple strands into many kinds of two dimensional shapes [19]. With the development of self-assembly techniques, various complex self-assembly structures have been established in recent years [20–24].

1.2 DNA strand branch migration

DNA strand branch migration is a mechanism of molecules self-driving hybridization that makes free energy from active to stable states. In the course of migration, the factors like the length and sequences of input signal strands play an important role, which will control the downstream reaction of strand displacement [4,5]. In other words, it is a process in which the longer DNA strands have more favorable trends to hybridize than that of short strands. The principle of DNA strand displacement is illustrated in Figure 1. At first, two strands A and B are hybridized with partial complementary regions (the length of strand A is longer than strand B). After adding the strand C (totally complementary with strand A), the specific recognition regions of strand C will react with toe-hold regions of strand A (unhybridized regions). In order to make system free energy stable, the strand C will gradually occupy the binding site of strand A, and displace strand B at the same time. By highly specific recognition, this process can be triggered in parallel and in multiple levels. Therefore, with the rapid development of DNA nanotechnology, DNA branch migration has become an attractive focus in the DNA computing research [12–14].

2 Experimental methods

2.1 Computing model structure and DNA sequence design

In the last century, Boolean operation was utilized to implement logic computing, and the logical calculus was also successfully constructed. In 1930s, logic computing is applied to the fields of circuit and system. In traditional logical circuits, data is processed by a series of binary system electronic information: a lower and a higher voltage. With the development of microelectronic technology and computer, a variety of complicated super-large systems have been established. However, until now, the basic computing process is still accomplished by logic computing. In 2011, *Nature* reports the special introduction of DNA logical operator. Therefore, no matter in computer science or other related fields, it is significant to use natural molecules to simulate logic computing.

In this experiment, two basic logic computing model (“AND” and “OR”) are implemented by DNA strand branch migration. The “AND” and “OR” logic modules are constructed by single polar and bipolar strand displacements respectively, which are carried out by a series of experiments. The specific methods are as follows.

(1) The design of logic computing model based on strand branch migration. In the course of establishing logic computing models, there are two main aspects: DNA self-assembly structure and DNA sequence design. Firstly, the DNA self-

assembly structure is designed according to the computing principle, which includes the design of recognition sites and hidden non-signal regions. Secondly, the appropriate DNA sequences are carefully designed based on lots of constraints. In the process of DNA sequence designing, false polymers and mismatches should be avoided. In addition, according to the differences of computing methods and principles, related DNA self-assembly structure should be constructed. Particularly, in the course of designing, the methods of computer aided encodings are also important.

In the "AND" self-assembly computing system, to carry out the modular design and the reaction of strand displacement, the main structure is constituted by two kinds of strands, which are a long structural strand A and two signal recognition strands (ab and cd) (Figure 2). Using the method of single-polar strand displacement, the input signal strand ab-1 or cd-1 can only hybridize with a single signal recognition strand ab or cd respectively. The long structural strand A is made of 20 bp (two sections, labeled by yellow and red), binding two signal recognition strands ab and cd together. Additionally, strands ab and cd consist of two sections, which are the structural region (b and c) and the recognition region (a and b). In this experiment, the sequences of two regions are designed respectively. The structural region, a frame of the computing system, is not allowed to disassemble randomly, and its length is set to be 10 bp. On the other hand, the length of the recognition region is set about 6 to 7 bp. This region, the key site of strand displacement, plays an important role in specific hybridization and recognition. As shown in Figure 2, the complementary sequences are represented by the same colors.

Similarly, the "OR" self-assembly system is established by strand displacement to implement computing. Different from the "AND" logic system, the method of bipolar strand displacement is introduced in the "OR" logic system (Figure 3). In other words, two different recognition regions (e and h) are able to be simultaneously activated. In this system, the structural strand B and recognition regions (e, f, g, h) are employed. In addition, the concept of mixture signals is also introduced into the experiment. That is to say, an input signal consists of various mixed signal strands, such as signal 3 (strands ef-1 and eg-1) and signal 4 (strands fh-1 and gh-1). After displacing, the reaction results are made of some complementary double-stranded DNA and single-stranded DNA.

(2) The strand displacement process triggered by inputting signal strands. The related input signal strands are added in different logic computing modules. Then, different output strands or structures are released according to the recognition between specific DNA strands (Figures 2 and 3).

(3) The detection of computing results. By detecting PAGE gel, the computing results are obtained from the constructed logic computing models.

2.2 The methods and materials

(1) The experimental methods. Logic computing models are formed by adding various DNA strands under slow annealing. The specific methods are shown as follows. In the course of constructing DNA logic computing models, various equimolar DNA strands are added to a final concentration of 1 pmol/ L, in TAE/Mg^{2+} buffer. The mixture is hybridized under the reaction condition of 95°C for 4 min, 65°C for 30 min, 50°C for 30 min, 37°C for 30 min, and 22°C for 30 min. Then, the obtained products are detected by PAGE. During the course of strand branch migration, all reactions are produced in 1×PBS buffer. In addition, the input DNA strands and DNA logic modules are mixed and react for over 6 h at room temperature. At the end of the reactions, the products are preserved at 4°C.

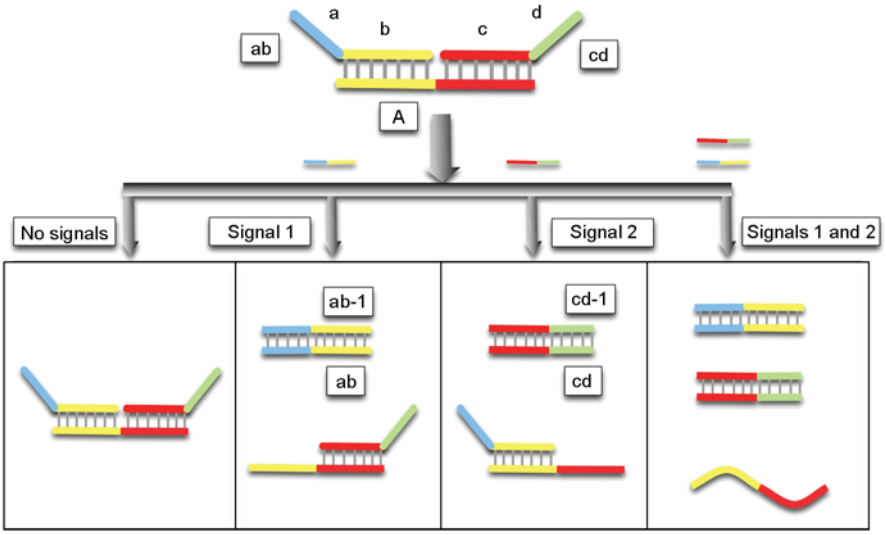

Figure 2 The principle of "AND" computing system.

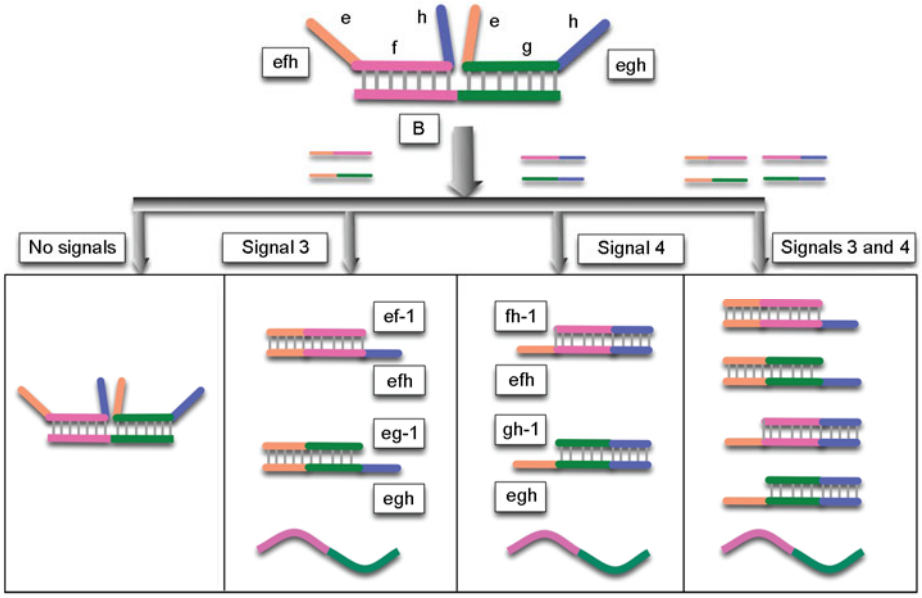

Figure 3 The principle of "OR" computing system.

(2) The experimental reagents. All DNA strands were synthesized from Shanghai Sangon and purified by PAGE. DNA sequences in the experiment are shown in Table 1.

The main reagents are as follows: Tris, Na_2EDTA, ethidium bromide (EB), acrylamide, N,N′-Methylenebisacrylamide, ammonium persulfate, TEMED and Stain all were bought from Sigma-Aldrich Co. LLC. DNA marker (20 bp ladder and 50 bp ladder), and bromophenol blue were from TaKaRa biotechnology Co. LTD. 10×PBS buffer was from solarbio bioscience and technology Co. LTD. TAE/ Mg^{2+} buffer (0.04 mol L^{-1} Tris acetate, 1 mmol L^{-1} EDTA, 12.5 mmol L^{-1} Mg acetate, pH 8.3), and the 500 mL mother liquor of acrylamide at the concentration of 45% (217 g acrylamide and 8 g N,N′-Methylenebisacrylamide).

3 Results

In this work, the computing results were detected by PAGE electrophoresis. Because the speeds of assembly polymers in gel changes with the various DNA self-assembly structures, it was easier to obtain the computing results from gels by directly analyzing DNA bands.

In the "AND" computing module, there were four conditions should be considered, as no signal, signal 1, signal 2 and signals (1, 2). (1) When there was no signal added, the structures of DNA self-assembly did not have any change. (2) When signal 1 (ab-1) was added, the DNA strands ab were displaced by specific recognition of strand regions a. (3) When signal 2 (cd-1) was added, the DNA strands cd were completely displaced through specific recognition of strand regions d. (4) Both signals (1, 2) were added. The specific recognitions occurred at the same time in regions a and d. Then, strand A was fully released from the self-assembly structures. In the structure design principle, only when both signals 1 and 2 were added together, the initial DNA self-assembly structures would disintegrate. From PAGE gel results (Figure 4), lanes 1 to 5 displayed various DNA bands of different assembly polymers, and lanes 6 to 10 indicated the computing results. The gel place marked by blue arrow was the gel band of initial self-assembly polymers. The computing results was able to be analyzed from

Table 1 DNA sequences

Name	Sequence (5′ to 3′)	Name	Sequence (5′ to 3′)
A	TCTGGCACTATGACAAGCGA	efh	CTAATCACCCTGCTTCGGAACTGG
ab	TAGTGCCAGAGGTATCC	ef-1	CCAGTTCCGAAGCAGGG
ab-1	GGATACCTCTGGCACTA	eg-1	CCAGTTCAATGCGTATG
cd	ATTCATCTCGCTTGTCA	egh	CTAATCACATACGCATTGAACTGG
cd-1	TGACAAGCGAGATGAAT	fh-1	CGAAGCAGGGTGATTAG
B	CGAAGCAGGGAATGCGTATG	gh-1	AATGCGTATGTGATTAG

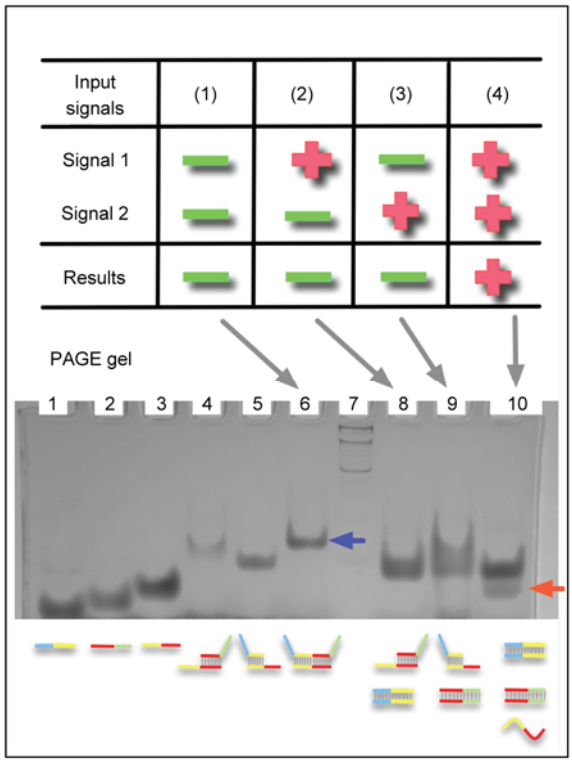

Figure 4 The gel results of "AND" logic computing (DNA markers in lane 7).

the place indicated by the reddish orange arrow, where was the position of strand A gel band. By detecting, only lane 10 produced the gel band of strand A (truth-value, both signals 1 and 2 were added). Compared with lanes 6, 8 and 9 in the same place on gel picture, there were no gel band released (false-value).

In Figure 4, the initial self-assembly polymer was displayed in lane 6. At first, by comparing strand A in Lane 3, there was no other strands existing in lanes 8 and 9. However, the strands of DNA dimmers were produced in lane 3, which had slower speeds than strand A. On the other hand, after adding signals 1 and 2, the new strand was generated in lane 10, and had almost the same speed with strand A. Thus, it turned out that this new strand was the released signal.

Similarly, in the module of "OR" computing, there were also four input conditions: no signal, signal 3, signal 4 and signal (3 and 4). Different from "AND" computing, the signals here were mixed DNA strands, which would displace with specific recognition regions individually. (1) No inputs, no changes happened in initial assembly structures. (2) When signal 3 (mixtures of strands ef-1 and eg-1) was added, the displacements happened in the recognition regions e both of strands efh and egh. (3) When signal 4 (mixtures of strands fh-1 and gh-1) was added, the strands of efh and egh were also completely displaced. (4) When both signals 3 and 4 were added, all specific recognition regions e and h were activated with the initial self-assembly structure disintegrating. In principle, adding any of signals 3 or 4 would lead to disassembling of initial DNA polymers. In gel results (Figure 5), lanes 5 to 8 indicated results of "OR" computing. When no signal was input, the assembly structures were kept as initial conditions (blue arrow in lane 5). However, when any of signals 3 or 4 were added, the displacements happened with the whole initial structures disintegrated. After disassembling, for those produced assembly polymers, the gel speeds increased greatly (indicated by the reddish orange arrow). By analyzing gel results, the initial assembly products in lanes 3 to 5 were disintegrated (any of input signals 6 or 8 leads to truth-value). After detecting, all experimental results were in accordance with expectation.

4 Discussion

In this paper, a DNA logic computing model was constructed using strand branch migration, combined with DNA self-assembly technology. In general, this model consisted of two parts: DNA self-assembly module and triggering module based on strand displacement. By adding specific DNA signal molecules, initial self-assembly structures were disintegrated via DNA strand displacement, and other specific output signal molecules were produced accordingly. Compared with previous computing models, there are several main

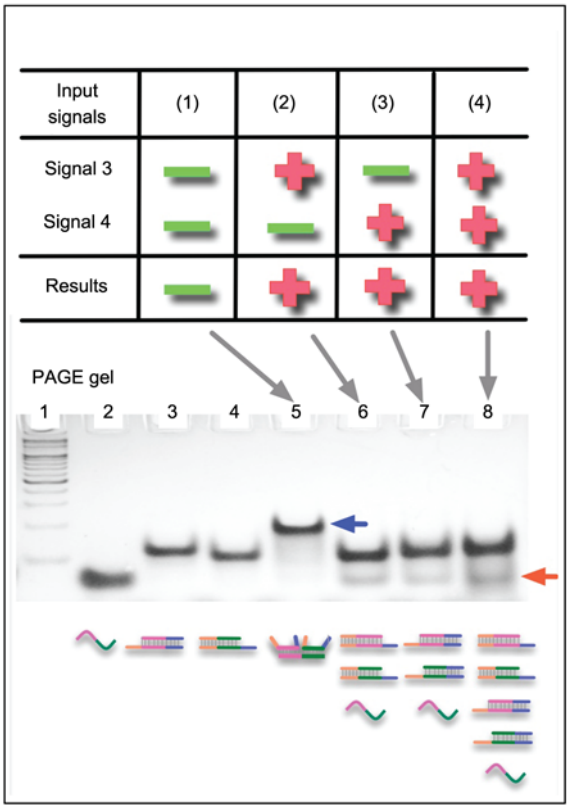

Figure 5 The gel results of "OR" logic computing (DNA markers in lane 1).

characteristics in this study as follows: (1) Spontaneous reacting of DNA strand displacement at room temperature. The whole reaction of DNA strand displacement was carried out at room temperature, so it avoided having denaturing influence on other self-assembly structures. In this way, the accuracy of experimental results was guaranteed. (2) The applications of single-polar and bipolar strand displacement. Since the method of strand displacement was utilized, information processing could be implemented by modular self-assembly in parallel. Therefore, there were several advantages for such designing. First, modularization of signal strands design. In this work, the different recognizing regions were designed just by changing a few of short DNA ends, while the main structural frame regions did not change at all. It was very convenient for large scale designing. Second, different computing modules were able to be triggered by the same signal strands. It would be beneficial for constructing various computing modules with just a few signal strands. Finally, the input signal information was constituted not only of a single kind of strands, but also some kinds of mixed strands. (3) The high accuracy of detecting experimental results. In this model, the disintegrating of self-assembly structures would lead to the varying of electrophoresis mobility of DNA polymers. Thus, the computing results were able to be detected by the positions of different electrophoresis bands. After a series of experimental operations, this logic computing model was implemented precisely by the input of specific signal DNA strands.

However, there were still some difficulties in the course of constructing model. Firstly, DNA coding design in the complex system was a major problem in this study. Because the method of bipolar strand displacement was utilized, how to design highly efficient DNA sequence was a key problem to be considered. In previous designs, the main target was to decrease the similarities between DNA sequences. However, in this work, two DNA displacement strands (such as strands ef-1 and eg-1) had the same specific recognition sites. Therefore, it was more difficult to design the qualified DNA code with little false base pairs than before. Secondly, it is difficult for those mixed signal strands to recognize with each other. Because many kinds signal strands reacted at the same time, it is very important to find suitable control conditions to improve specific signal recognitions. Thirdly, the detection method of PAGE gel was rather simple to some extent. Finally, the brightness of released strand was lower. Two possible reasons are as follows: (1) double strand DNA has stronger capability of coloring than single strand DNA. (2) the lower efficiency of strand branch migration. However, all initial self-assembly polymers were transferred into other smaller polymers, which demonstrated high efficiency of strand branch migration. Therefore, the answer for lower brightness of released strand may be reason (1). In the future, multiple detection methods [26–28] should be introduced, as combining with fluorescent labeling, atomic force microscope, and nanoparticles etc.

Although there are some insufficiencies in this study, the successful implementation of molecular computing model provides the evidences that: this model has a potential ability to solve complicated problems. In addition, it provides a way to other related fields such as computer science, nanotechnology and cryptography. With the advancing of molecular computing theory and experiment, it can be expected that DNA computing will play significant roles in solving difficult computing problems.

This work was supported by the National Natural Science Foundation of China (61127005, 61133010, 61033003, 60910002 and 61143003), Ph.D. Programs Foundation of the Ministry of Education of China (20110001130016), and the Postdoctoral Special and Normal Science Foundation of China (201104018, 20100480128, 2011M500197).

1 Adleman L. Molecular computation of solution to combinatorial problems. Science, 1994, 66: 1021–1024
2 Ouyang Q, Kaplan P D, Liu S, et al. DNA solution of the maximal clique problem. Science, 1997, 278: 446–449
3 Mao C D, LaBean T H, Reif J H. Logical computation using algorithmic self-assembly of DNA triple-crossover molecules. Nature, 2000, 28: 493–496
4 Qian L L, Winfree E. Scaling up digital circuit computation with DNA strand displacement cascades. Science, 2011, 332: 1196–1201
5 Qian L L, Winfree E, Bruck J. Neural network computation with DNA strand displacement cascades. Nature, 2011, 475: 368–372
6 Pei R J, Matamoros E, Liu M H, et al. Training a molecular automaton to play a game. Nat Nanotechnol, 2010, 5: 773–777
7 Gu H Z, Chao J, Xiao S J, et al. Molecular robots guided by prescriptive landscapes. Nature, 2010, 465: 202–205
8 Rinaudo K, Bleris L, Maddamsetti R. A universal RNAi-based logic evaluator that operates in mammalian cells. Nat Biotechnol, 2007, 25: 795–801
9 Maune H T, Han S P, Barish R D, et al. Self-assembly of carbon nanotubes into two-dimensional geometries using DNA origami templates. Nat Nanotechnol, 2010, 5: 61–66
10 Morrow T J, Li M W, Kim J, et al. Programmed assembly of DNA-coated nanowire devices. Science, 2009, 323: 352–353
11 Clelland C T, Risca V, Bancroft C. Hiding messages in DNA microdots. Nature, 1999, 399: 533–534
12 Seelig G, Soloveichik D, Zhang Y D, et al. Enzyme-free nucleic acid logic circuits. Science, 2006, 314: 1585–1588
13 Shin J S, Pierce N A. Rewritable memory by controllable nanopatterning of DNA. Nano Lett, 2004, 4: 905–909
14 Zhang D Y, Turberfield A J, Yurke B, et al. Engineering entropy-driven reactions and networks catalyzed by DNA. Science, 2007, 318: 1121–1125
15 Frezza B M, Cockroft S L, Ghadiri M R. Modular multi-level circuits from immobilized DNA-based logic gates. J Am Chem Soc, 2007, 129: 14875–14879
16 Winfree E. Design and self-assembly of two-dimensional DNA crystals. Nature, 1998, 394: 539–544
17 LaBean T H, Reif J H, Seeman N C, et al. Construction, analysis, ligation, and self assembly of DNA triple crossover complexes. J Am Chem Soc, 2000, 122: 1848–1860
18 Carbone A, Seeman N C. Circuits and programmable self-assembling DNA structures. Proc Natl Acad Sci USA, 2002, 99: 12577–12582
19 Rothemund P W. Folding DNA to create nanoscale shapes and patterns. Nature, 2006, 440: 297–302
20 Goodman R P, Schaap I A T, Tardin C F, et al. Rapid chiral assembly of rigid DNA building blocks for molecular nanofabrication. Science, 2005, 310: 1661–1665

21 Zhang C, Su M, He Y, et al. Conformational flexibility facilitates self-assembly of complex DNA nanostructures. Proc Natl Acad Sci USA, 2008, 105: 10665–10669
22 Aldaye F A, Lo P K, Karam P, et al. Modular construction of DNA nanotubes of tunable geometry and single- or double-stranded character. Nat Nanotechnol, 2009, 4: 349–352
23 Han D R, Pal S, Nangreave J, et al. DNA origami with complex curvatures in three-dimensional space. Science, 2011, 332: 342–346
24 Ackermann D, Schmidt T L, Hannam J S, et al. A double-stranded DNA rotaxane. Nat Nanotechnol, 2010, 5: 436–442
25 Thomas C. DNA as a logic operator. Nature, 2011, 469: 45–46
26 Liu Q H, Wang L M, Frutos A G, et al. DNA computing on surfaces. Nature, 2000, 403: 175–179
27 Benenson Y, Elizur T P, Adar R, et al. Programmable and autonomous computing machine made of biomolecules. Nature, 2001, 411: 430–434
28 Benenson Y, Gil B, Dor U B, et al. An autonomous molecular computer for logical control of gene expression. Nature, 2004, 429: 423–429