Effective algorithm to encrypt information based on self-assembly of DNA tiles

Miki Hirabayashi¹, Hiroaki Kojima and Kazuhiro Oiwa
Kobe Advanced ICT Research Center (KARC), National Institute of Information and Communications Technology (NICT), 588-2 Iwaoka, Nishi-ku, Kobe 651-2492, JAPAN

ABSTRACT

We present an error-tolerance scheme to encrypt information in DNA structures based on a one-time-pad (OTP) cryptosystem that provides theoretically unbreakable security. The problem of the DNA-based OTP encryption is the loss of synchronization between the message and the encryption key due to the DNA property of accepting mismatched base pairs. We propose a new implementation idea of encrypting algorithm with the fourfold fault tolerance against mismatches than the ordinary DNA XOR operation. Although there are several problems to be solved toward the practical use at this moment, it is expected that the molecular computation using DNA tiles will enlarge the application possibility of the OTP cryptosystem.

INTRODUCTION

The one-time pad (OTP) is an unbreakable perfect encryption algorithm if a random key or pad is truly random, never reused, and kept secret [1]. Although the generation of truly random numbers costs much, the huge amount of randomly assembled keys of the same size as the input message is required for one-time use. Therefore OTP encryption has been limited to vital intelligence secrets. Since the idea of DNA OTP cryptosystem using DNA XOR logical computation was proposed [2,3], DNA has received considerable attention as an effective tool to realize the practical OTP. Several algorithms are proposed for cryptography and steganography using DNA computation [3-5]. To implement OTP encryption algorithm with DNA, a plaintext message is encoded as binary data in DNA sequence. Then a ciphertext is created by DNA XOR system using a random key [3,4]. The n bits of plaintext binary input $I$, is XOR’ed with the random key $K$, to produce the encrypted bit $C_i = I_i \oplus K_i$ for $i = 1, \ldots, n$.

It was suggested that the problem of this system is the loss of synchronization between the input and the key, which would occur when a bit is spuriously introduced or deleted from either sequence [3].

In this work we propose new idea to realize fourfold fault tolerance against misintroduction of a bit than the standard DNA XOR OTP operation and discuss the implementation difficulty of DNA computation.

RESULTS AND DISCUSSION

Figure 1a shows a typical triple-helix tile for the XOR operation based on Ref. [3]. Each tile consists of four oligonucleotide strands shown as different line types with arrowheads on their 3’ ends. The three double-helical domains are represented by horizontal helix axes. The upper and lower helices have bare strand ends which can connect to the neighbour tiles. The central helix is capped at both ends with hairpins. Four paired vertical lines in Fig. 1b represent crossover points where two strands exchange between two helices. The tile shown in Fig. 1 has three uncomplemented sticky ends and one blunt end. Sticky ends are designed for assembly of large structure including computational functions as shown in an example of XOR computation in Fig. 2.

For DNA OTP encryption, calculation tiles binds into appropriate binding slots between input tiles and key tiles by the matching of sticky ends (Fig. 3). We can get a reporter strand that runs through the entire assembly as an output by adding ligase enzyme. The reporter strand is shown as a bold black line in all figures. This strand is extracted by first melting apart the hydrogen bonding between strands and then purifying the polyacrylamide gel electrophoresis [3]. Although XOR calculation tiles in Fig.

Fig. 1 A typical DNA tile. (a) An individual triple-crossover (TX) tile, with three helical domains through Watson-Crick base paring which consist of four backbone strands shown as different line types with arrowheads on their 3’ ends [2]. Sticky end “1” has the sequence that represents the tile value. Sticky end “2” represents the specific sequence to connect input tiles according to requirements. (b) An abstract expression of the tile shown in (a). Three helical domains are drawn as rectangles, flanked by sticky ends shown as geometrical shapes with arrows. The bold black line represents the reporter strand to associate the calculated output with the input.
Fig. 2 XOR computation by self-assembly of DNA tiles. (a) An abstract expression of cumulative XOR computation using DNA tiles based on Ref. [2]. (b) Component tiles. Black strands are reporter strands. At the end of the assembly, the reporter strand is ligated to associate the calculation tiles with the input tiles.

2b consist of specific sticky ends only, the OTP calculation tiles in Fig. 3b include non-specific sticky ends. As a result, input tiles and key tiles lose synchronization and calculation tiles fail to form a correct reporter strand.

To solve this problem, we propose to use blunt ends in calculation tiles to form a reporter strand. Usually sticky ends are used for assembly and calculation in DNA-tile operations. In the OTP operation, the system can perform the calculation without the connection of the reporter strand through the sticky ends as shown in Fig. 3. When blunt ends are used to form a reporter strand, the efficiency of ligation becomes low. On the other hand, the probability of mishybridization of calculation tiles becomes low, too. Thus the fault tolerance is provided. To estimate the probability of error correction, we assume that the energy to open the double strand at each mismatched sticky end equals zero and it equals 8G at each matched sticky end to replace an error tile with a correct one. If the probability of error correction is inversely proportional to total recovery energy from possible states, the calculation tile with a blunt end has quadruplex error tolerance than that with sticky-ends for a reporter strand.

CONCLUSION

We proposed an idea of the effective OTP cryptosystem using DNA tiles to provide fault tolerance. The molecular computation can be operated in parallel and DNA works as high-performance information storage media. Thus OTP cryptosystem is expected as a promising application of DNA computation. Although several methods have been presented, the perfect implementation of DNA XOR OTP is not realized due to operational difficulty. For example, in...

Fig. 3 XOR computation for OTP operation. (a) An possible example of DNA OTP encryption based on Ref. [3]. (b) An example of a set of component tiles. Each calculation tile has two ends with the specific sequence for computation and two ends with the common sequence to connect the reporter strand.

our system, the extraction of a reporter strand often becomes difficult if the ligation efficiency of blunt ends for a reporter strand is very low. DNA assembly can be extended to three dimensions. Utilizing the three dimensional folding in a bio-inspired manner, we expect that we can solve trade-off problems caused by the introduction of fault tolerance system and realize a practical DNA XOR OTP cryptosystem.

REFERENCES


*Corresponding author. E-mail: m_hirabayashi@office.so-net.ne.jp (miki.hirabayashi@nict.go.jp)