

01 Jan 2004

Interactive DNA Sequence and Structure Design for DNA Nanotechnology and DNA Computation

Mingjun Zhang

Chaman Sabharwal

Missouri University of Science and Technology, chaman@mst.edu

Weimin Tao

Tzyh-Jong Tarn

et. al. For a complete list of authors, see https://scholarsmine.mst.edu/comsci_facwork/265

Follow this and additional works at: https://scholarsmine.mst.edu/comsci_facwork



Part of the [Computer Sciences Commons](#)

Recommended Citation

M. Zhang et al., "Interactive DNA Sequence and Structure Design for DNA Nanotechnology and DNA Computation," *Proceedings of the 4th IEEE Conference on Nanotechnology, 2004*, Institute of Electrical and Electronics Engineers (IEEE), Jan 2004.

The definitive version is available at <https://doi.org/10.1109/NANO.2004.1392331>

This Article - Conference proceedings is brought to you for free and open access by Scholars' Mine. It has been accepted for inclusion in Computer Science Faculty Research & Creative Works by an authorized administrator of Scholars' Mine. This work is protected by U. S. Copyright Law. Unauthorized use including reproduction for redistribution requires the permission of the copyright holder. For more information, please contact scholarsmine@mst.edu.

Interactive DNA Sequence and Structure Design for DNA Nanotechnology and DNA Computation

Mingjun Zhang*, Chaman L. Sabharwal†, Weimin Tao‡, Tzyh-Jong Tam§, Ning Xi¶ and Guangyong Li¶

*Life Sciences and Chemical Analysis Division, Agilent Technologies, CA. Email: mingjunzhang@ieee.org

†Department of Computer Science, University of Missouri-Rolla, MO. Email: chaman@umr.edu

‡Brooks Automation, CA. Email: david_tao@brooks-pri.com

§Department of Electrical and Systems Engineering, Washington University in St. Louis, MO. Email: tarn@wuauto.wustl.edu

¶Department of Electrical Engineering, Michigan State University, MI. Email: {xin,liguangy}@egr.msu.edu

Abstract—DNA sequence and structure design is very important for DNA nanotechnology and DNA computation. A computer aided design tool is needed for exploring DNA sequence and structure of interests before experimental synthesis, which is a very time and labor consuming process. In this paper, an interactive DNA sequence and structure design software tool called DNA shop is proposed and implemented.

I. INTRODUCTION

As the carrier of genetic information for all living species, DNA has been well-known for its unique properties of information encoding, structure self-recognition and self-assembly. DNA computation and DNA nanotechnology are two emerging fields aiming to use these properties.

DNA nanotechnology takes advantage of the fact that the intermolecular interactions of DNA are highly specific and readily programmed through Watson-Crick complementary property. Possible applications include scaffolds for molecular electronic devices and nanometer-scale robots construction. Seeman [7] proposed DNA as scaffolds to organize structures of other molecules. Mao et al. [6] showed that self-assembly structure of branched DNA motifs can provide basis for dynamic assembling switchable molecular machines. Yurke et al. [9] reported that the construction of a DNA machine in which the DNA is used not only as a structural material, but also as fuel.

In addition, the combination of DNA information-encoding and recognition properties, and the enzymatic machinery capability for DNA manipulation facilitate the emergence field of DNA computation. The feasibility of DNA computation was first demonstrated by Adleman in 1994 [1]. Benenson [2] discussed a programmable finite automation comprising DNA and DNA-manipulating enzymes that solves computational problems autonomously. The automation's hardware consists of a restriction nuclease and ligase, the software and inputs are encoded by double-stranded DNA, and programming amounts to choose appropriate software molecules. Upon mixing solutions containing these components, the input molecules are processed via a cascade of restriction, hybridization and ligation cycles, producing a detectable output molecules that encode the computational results.

For above DNA nano-applications, DNA sequence and structure design is a critical step. Currently, one must obtain proper conditions, refine designs and determine experimental windows for DNA structure design through tedious and often expensive processes of trials and errors [8]. A visualization tool is needed for DNA sequence and structure design. The purpose of this paper is to propose a software tool for DNA sequence and structure design.

II. DNA PROPERTIES

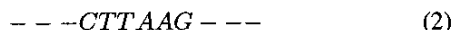
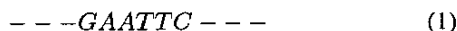
DNA's unique biological properties include the specificity of the base pairing that holds two strands of double helix together: Adenine (A) pairs with Thymine (T) and Guanine (G) pairs with Cytosine (C). Usually, DNA double helix generated from the complementary interactions is a linear molecule. Its axis is not branched in the biological sense. However, branched DNA molecules do occur as key intermediate state in DNA metabolism, particularly in the processes of replication, recombination.

Properties of DNA double helix are unlike those of any other natural or synthetic polymers. The molecule's characteristic base stacking and braided architecture lend it unusual stiffness: it takes about 50 times more energy to bend a double-stranded DNA molecule into a circle than to perform the same operation on single-stranded DNA. Furthermore, the double-stranded DNA molecule is very stable. These features make the double-stranded DNA molecule a great candidate of scaffolds for other molecules. DNA electrical properties make DNA one of the most interesting bio-molecules for molecular electronics [10].

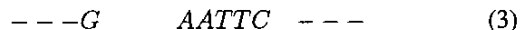
DNA molecules are in the nanometer scale and encode binary information. A DNA double helix is about 2 nm in diameter with helical repeat of about 10 base pairs, which produces a pitch of 3.4-3.6 nm. A small volume of DNA contains a vast number of molecules. DNA in weak solution of one liter of water can encode 10^7 to 10^8 tera-bytes information, which makes DNA molecule a potential material for information storing. DNA sequence may be used for encoding information that can be read externally by proteins and nucleic acids. Most current data storing media has a life around 100 years. DNA molecules may have much longer, stable and larger capacity for information storage.

A. DNA sticky ends

DNA can be cut at precise locations by using restriction enzymes. Another enzyme—DNA ligase—can then be used to reassemble the pieces into any desired order. Together, these two enzymes allow researchers to assemble customized DNA structures. For example, restriction enzymes typically recognize a symmetrical sequence of DNA, such as the site of EcoRI:



For the DNA sequences, the top strand is the reverse of the bottom strand. Using restriction enzymes to react at this site and break hydrogen bonds holding the overlapping single stranded complementary strands, which are the strands between G and A, overhanging chains as follows can be obtained:



The segments *AATT* in (3) and *TTAA* in (4) are called sticky ends, which are complementary each other. DNA sticky ends provide a predictable, diverse, reliable and programmable set of intermolecular interactions. DNA molecules can be manipulated by commercially available enzymes – they can be joined by DNA ligases, cleaved at specific sites by restriction enzymes, phosphorylated by kinases and have their topology altered by DNA topoisomerases.

B. DNA self-assembly

The capability of fabricating individual molecules and atoms is the key for nanotechnology. Self-assembly, which is a method for constructing structures by spontaneously self-ordering of substructures, is an attractive approach for nano-structure fabrication. The technique works by simulating the way biological systems build molecules, viruses, and cells. DNA self-assembly presents a bottom-up approach to fabricate nano-scale objects. DNA self-assembly uses artificially synthesized single strand DNA to self-assemble into different DNA crossover molecules (tiles), which have sticky ends that preferentially match the sticky ends of certain other DNA tiles, facilitating the further assembly into tiling lattices – DNA machines.

III. A SOFTWARE TOOL FOR DNA SEQUENCE AND STRUCTURE DESIGN: DNA SHOP

Two steps usually are involved in the DNA sequence and structure design. The first step is called sequence selection and the second step can be regarded as interactive DNA structure design. Sequence selection is the critical part of DNA structure design. Sequence needs to be properly defined so that the Watson-Crick complementarity and sticky ends can generate desired structure. Branched target molecules correspond to an excited state must be taken to ensure that the excited product obtained is the one that is sought. During the structure design, users need to have a flexibility to move DNA segments

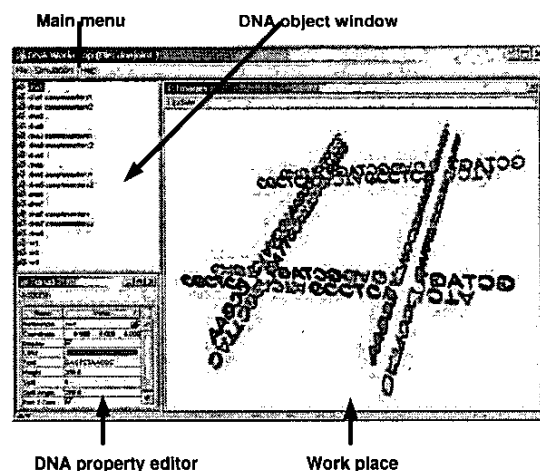


Fig. 1. User interface of the DNA shop

around interactively. Users may specify the length, start and ending bases of the sequence. Complementary sequences will be generated automatically following the Watson-Crick complementarity. Sticky ends are specified by users interactively.

To facilitate DNA sequence and structure design, a software tool called DNA shop is developed. DNA shop is a Java visualization tool for interactive DNA sequence and structure design. The software tool contains object management module, interactive user input module, graphics display module, coordinate transformation module, and DNA object module. The system is based on Java 3D class and Java J2SE SDK. Fig. 1 shows the user interface of the tool. It contains a “main menu”, a “work place”, a “DNA object window” and a “DNA property editor”. Components and functions of each windows are as follows:

- The main menu contains “File”, “Simulation” and “Help” three submenus. The “File” submenu is used to open and save a file, or exit the program. The “simulation” submenu is used to activate the work space.
- The work space shows the designed DNA structure. Users may select bases and change them. Users may move any part of a sequence to different locations by changing the coordinates.
- The DNA object window shows all DNA objects in the work space. For each DNA object, users can make copy, generate complementary or change the name.
- The DNA property editor shows each DNA’s object-oriented properties including “Reference”, “Coordinate”, “Display”, “Color”, “Text”, “Height”, “Left split”, “Split angle”, and “Left complementary”.
 - **Reference** represents origin of the DNA sequence. It could be any pre-defined world coordinate or another DNA object. In Fig. 1, W1 is the world coordinate.
 - **Coordinate** is the world coordinate of the DNA with respect to the “Reference” point.
 - **Display** property controls whether displays the ob-

