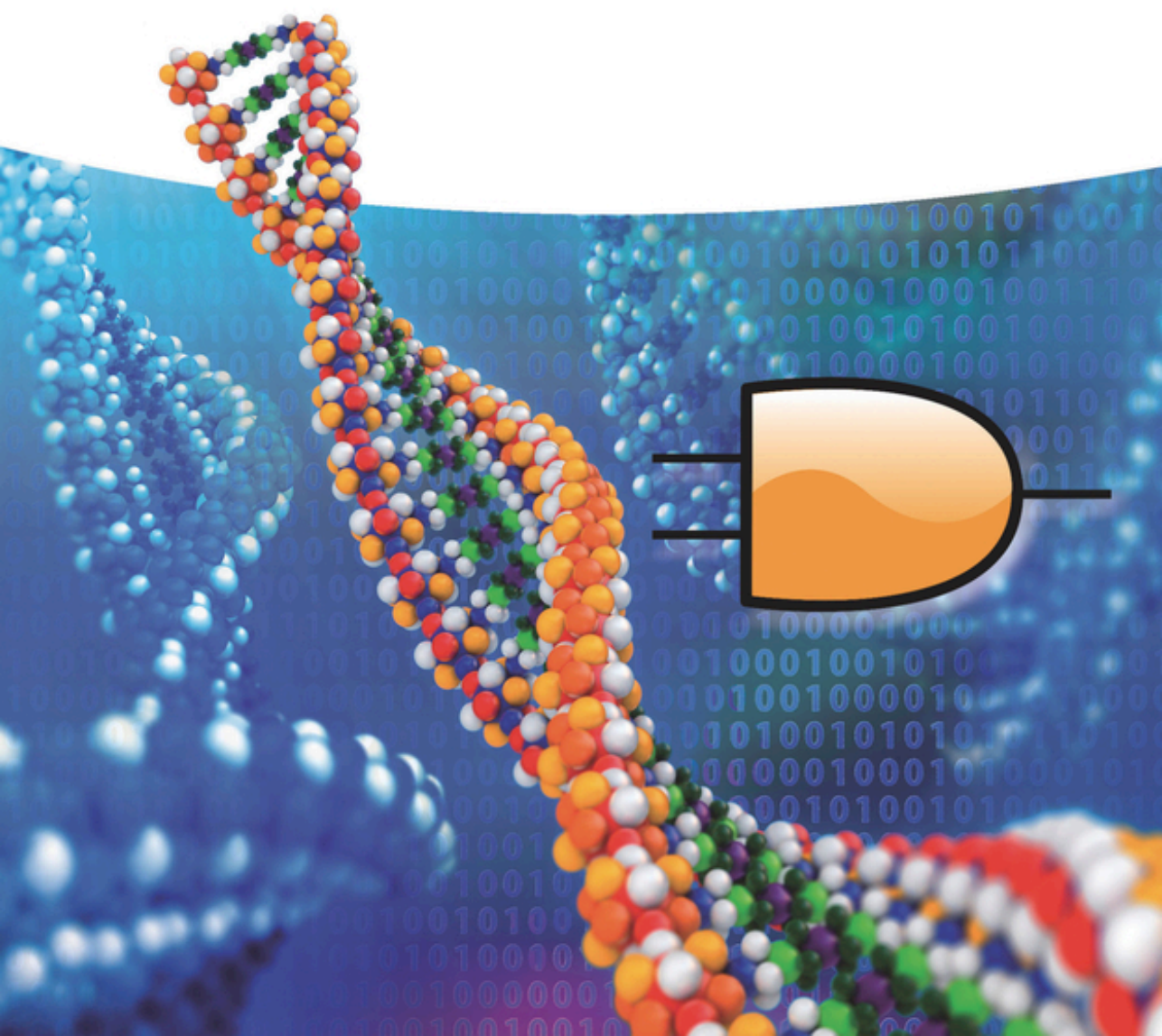


Edited by
Evgeny Katz

DNA- and RNA- Based Computing Systems



DNA- and RNA-Based Computing Systems

DNA- and RNA-Based Computing Systems

Edited by

Evgeny Katz

WILEY-VCH

Editor**Dr. Evgeny Katz**

Clarkson University
Department of Chemistry
Clarkson Avenue 8
NY
United States

■ All books published by **Wiley-VCH** are carefully produced. Nevertheless, authors, editors, and publisher do not warrant the information contained in these books, including this book, to be free of errors. Readers are advised to keep in mind that statements, data, illustrations, procedural details or other items may inadvertently be inaccurate.

Library of Congress Card No.:
applied for

British Library Cataloguing-in-Publication Data

A catalogue record for this book is available from the British Library.

Bibliographic information published by the Deutsche Nationalbibliothek

The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available on the Internet at <<http://dnb.d-nb.de>>.

© 2021 WILEY-VCH GmbH, Boschstr.
12, 69469 Weinheim, Germany

All rights reserved (including those of translation into other languages). No part of this book may be reproduced in any form – by photoprinting, microfilm, or any other means – nor transmitted or translated into a machine language without written permission from the publishers. Registered names, trademarks, etc. used in this book, even when not specifically marked as such, are not to be considered unprotected by law.

Print ISBN: 978-3-527-34720-9

ePDF ISBN: 978-3-527-82540-0

ePub ISBN: 978-3-527-82541-7

oBook ISBN: 978-3-527-82542-4

Cover Design Adam-Design,
Weinheim, Germany

Typesetting SPi Global, Chennai, India

Printing and Binding

Printed on acid-free paper

10 9 8 7 6 5 4 3 2 1

Contents

Preface *xiii*

1	DNA Computing: Origination, Motivation, and Goals – Illustrated Introduction	1
	<i>Evgeny Katz</i>	
1.1	Motivation and Applications	1
1.2	DNA- and RNA-Based Biocomputing Systems in Progress	3
1.3	DNA-Based Information Storage Systems	8
1.4	Short Conclusions and Comments on the Book	10
	References	11
2	DNA Computing: Methodologies and Challenges	15
	<i>Deepak Sharma and Manojkumar Ramteke</i>	
2.1	Introduction to DNA Computing Methodologies	15
2.2	Key Developments in DNA Computing	16
2.2.1	Adleman Model	16
2.2.2	Lipton's Model	18
2.2.3	Smith's Model	19
2.2.4	Sakamoto's Model	21
2.2.5	Ouyang's Model	22
2.2.6	Chao's Model	24
2.2.7	DNA Origami	24
2.2.8	DNA-Based Data Storage	26
2.3	Challenges	26
	Acknowledgment	27
	References	28
3	DNA Computing and Circuits	31
	<i>Chuan Zhang</i>	
3.1	From Theory to DNA Implementations	31
3.2	Application-Specific DNA Circuits	35
	Acknowledgments	41
	References	41

4	Connecting DNA Logic Gates in Computational Circuits	45
	<i>Dmitry M. Kolpashchikov and Aresenij J. Kalnin</i>	
4.1	DNA Logic Gates in the Context of Molecular Computation	45
4.2	Connecting Deoxyribozyme Logic Gates	46
4.3	Connecting Gates Based on DNA Strand Displacement	47
4.4	Logic Gates Connected Via DNA Four-Way Junction (4WJ)	50
4.5	Conclusion	53
	References	53
5	Development of Logic Gate Nanodevices from Fluorogenic RNA Aptamers	57
	<i>Trinity Jackson, Rachel Fitzgerald, Daniel K. Miller, and Emil F. Khisamutdinov</i>	
5.1	Nucleic Acid: The Material of Choice for Nanotechnology	57
5.2	RNA Aptamers are Modular and Programmable Biosensing Units	58
5.3	Construction of RNA Nanoparticles with Integrated Logic Gate Operations Using Light-Up Aptamers	64
5.3.1	Implementation of MG-Binding RNA Aptamer to Design Binary Logic Gates	65
5.3.2	Implementation of MG-Binding RNA Aptamer and Broccoli RNA Aptamer to Design Half-Adder Circuit	68
5.4	Conclusion	70
	Acknowledgments	70
	References	70
6	Programming Molecular Circuitry and Intracellular Computing with Framework Nucleic Acids	77
	<i>Jiang Li and Chunhai Fan</i>	
6.1	Framework Nucleic Acids	77
6.2	A Toolbox for Biomolecular Engineering of Living Systems	80
6.2.1	Biomolecular Scaffolds	80
6.2.2	Logic Units	81
6.2.3	Cell Entry Vehicles	82
6.2.4	Isothermal Construction	83
6.2.5	Targeting and Editing	83
6.2.6	Signal Readout	84
6.2.7	Triggers and Switches	84
6.2.8	Error Correction and Resilience	84
6.3	Targeted Applications	85
6.3.1	Drug Delivery	85
6.3.2	Cellular Imaging	85
6.3.3	Metabolic Engineering and Cellular Pathway Investigation	86
6.4	Nucleic Acid Nanotechnology-Enabled Computing Kernel	86
6.5	I/O and Human–Computer Interfacing	89
6.6	Information Storage	90
6.7	Perspectives	91
6.8	Conclusion	95
6.8.1	Terminology	96
	References	97

7	Engineering DNA Switches for DNA Computing Applications	105
	<i>Dominic Lauzon, Guichi Zhu, and Alexis Vallée-Bélisle</i>	
7.1	Introduction	105
7.2	Selecting Recognition Element Based on Input	107
7.3	Engineering Switching Mechanisms	108
7.4	Engineering Logic Output Function Response	116
7.5	Optimizing Switch Response	117
7.6	Perspective	120
	Acknowledgments	120
	References	121
8	Fluorescent Signal Design in DNA Logic Circuits	125
	<i>Dan Huang, Shu Yang, and Qianfan Yang</i>	
8.1	Basic Signal Generation Strategies Based on DNA Structures	126
8.1.1	Strategies Based on Watson–Crick Hydrogen Bond	127
8.1.1.1	Signal Derived from Hairpin Structure/Molecular Beacon	127
8.1.1.2	Signal Derived from DNzyme Activity	128
8.1.1.3	Signal Derived from Strand Displacement Reaction	129
8.1.2	Strategies Based on Hoogsteen Hydrogen Bond	132
8.1.2.1	Signal Derived from G-Quadruplex	132
8.1.2.2	Signal with the Help of i-Motif	135
8.1.3	Signal Derived from Aptamer–Ligand Interaction	138
8.2	Designs for Constructing Multi-output Signals	138
8.2.1	Selecting Individual Signal Transducers	138
8.2.2	Designing Multifunctional Probes	141
8.3	Summary and Outlook	147
	References	149
9	Nontraditional Luminescent and Quenching Materials for Nucleic Acid-Based Molecular Photonic Logic	155
	<i>Rehan Higgins, Melissa Massey, and W. Russ Algar</i>	
9.1	Introduction	155
9.2	DNA Molecular Photonic Logic Gates	156
9.3	Nontraditional Luminescent Materials	158
9.4	Semiconductor “Quantum Dot” Nanocrystals	159
9.4.1	Quantum Dots	159
9.4.2	Logic Gates with QDs	160
9.5	Lanthanide-Based Materials	161
9.5.1	Luminescent Lanthanide Complexes	161
9.5.2	Coupling Lanthanide Complexes with Energy Transfer	163
9.5.3	Logic Gates with LLCs and Lanthanide Ions	163
9.5.4	Upconversion Nanoparticles	165
9.5.5	Logic Gates with UCNPs	165
9.6	Gold Nanoparticles	166
9.6.1	Gold Nanoparticles	166

9.6.2	Logic Gates with AuNPs and Colorimetric Output	166
9.6.3	Logic Gates with AuNPs and PL Quenching	168
9.7	Metal Nanoclusters	169
9.7.1	Metal Nanoclusters	169
9.7.2	Logic Gates with Metal Nanoclusters	170
9.8	Carbon Nanomaterials	171
9.8.1	Graphene and Graphene Oxide	171
9.8.2	Logic Gates with Graphene and GO	172
9.8.3	Carbon Dots	174
9.8.4	Logic Gates with CDs	175
9.9	Conjugated Polymers	175
9.9.1	Conjugated Polymers	175
9.9.2	Logic Gates with CPs	176
9.10	Conclusions and Perspective	177
	References	178
10	Programming Spatiotemporal Patterns with DNA-Based Circuits	185
	<i>Marc Van Der Hofstadt, Guillaume Gines, Jean-Christophe Galas, and André Estevez-Torres</i>	
10.1	Introduction	185
10.1.1	What is Spatial Computing?	185
10.1.2	Digital vs. Analog Computing	186
10.1.3	Computing Consumes Energy	186
10.1.4	Molecules Compute in Space Through Reaction–Diffusion Primitives	187
10.2	Experimental Implementation of DNA Analog Circuits	188
10.2.1	DNA Strand Displacement Oscillators	189
10.2.2	DNA/Enzyme Oscillators	189
10.2.2.1	Genelets	190
10.2.2.2	PEN Reactions	191
10.3	Time-Dependent Spatial Patterns	193
10.3.1	Edge Detection	194
10.3.2	Traveling Patterns	195
10.3.2.1	Fronts	196
10.3.2.2	Go-Fetch Fronts	197
10.3.2.3	Waves and Spirals	198
10.3.3	Controlling Spatio-Temporal Patterns	199
10.3.3.1	Controlling Diffusion Coefficients	199
10.3.3.2	Initial and Boundary Conditions	200
10.4	Steady-State Spatial Patterns	202
10.4.1	Colony Formation	202
10.4.2	Patterns with Positional Information	203
10.5	Conclusion and Perspectives	206
	Acknowledgments	207
	References	208

11	Computing Without Computing: DNA Version	213
	<i>Vladik Kreinovich and Julio C. Urenda</i>	
11.1	Introduction	213
11.2	Computing Without Computing – Quantum Version: A Brief Reminder	214
11.3	Computing Without Computing – Version Involving Acausal Processes: A Reminder	215
11.4	Computing Without Computing: – DNA Version	217
11.4.1	Main Idea	217
11.4.2	It Is Not Easy to Stop Biological Processes	218
11.4.3	Towards Describing Ligation Prevention in Precise Terms	218
11.4.4	What Is Given	219
11.4.5	What We Want to Find	219
11.4.6	Let Us Prove that the Ligation Prevention Problem Is NP-Hard	219
11.4.7	How NP-Hardness Is Usually Proved	219
11.4.8	How We Will Prove NP-Hardness	220
11.4.9	The Actual Proof by Reduction	220
11.5	DNA Computing Without Computing Is Somewhat Less Powerful than Traditional DNA Computing: A Proof	222
11.5.1	Which of the Two DNA Computing Schemes is More Powerful?	222
11.5.2	W-hierarchy: A Brief Reminder	222
11.5.3	Conclusion	224
11.6	First Related Result: Security Is More Difficult to Achieve than Privacy	224
11.6.1	What We Plan to do in this Section	224
11.6.2	How to Describe Privacy in Graph Terms	224
11.6.3	How to Describe Security in Graph Terms	225
11.6.4	Conclusion: Security Is More Difficult to Maintain than Privacy	226
11.7	Second Related Result: Data Storage Is More Difficult than Data Transmission	226
11.7.1	Application to Information Science	226
11.7.2	Data Storage	226
11.7.3	Data Transmission	227
11.7.4	Conclusion: Data Storage Is More Difficult than Data Transmission	228
	Acknowledgments	228
	References	228
12	DNA Computing: Versatile Logic Circuits and Innovative Bio-applications	231
	<i>Daoqing Fan, Erkang Wang, and Shaojun Dong</i>	
12.1	Definition, Logical Principle, and Classification of DNA Computing	231
12.2	Advanced Arithmetic DNA Logic Devices	232
12.2.1	Half-Adder, Half-Subtractor	232
12.2.2	Full-Adder, Full-Subtractor	234
12.3	Advanced Non-arithmetic DNA Logic Devices	235

12.3.1	Data Conversion: Encoder/Decoder, Multiplexer/Demultiplexer	235
12.3.2	Distinguishing Even/Odd Natural Numbers: The Parity Checker	236
12.3.3	DNA Voter and Keypad Lock	236
12.3.4	Parity Generator/Checker (pG/pC) for Error Detection During Data Transmission	237
12.3.5	Non-Boolean Ternary Logic Gates	239
12.4	Concatenated Logic Circuits	239
12.5	Innovative Multifunctional DNA Logic Library	241
12.6	Intelligent Bio-applications	241
12.7	Prospects	244
	Acknowledgment	244
	References	244
13	Nucleic Acid-Based Computing in Living Cells Using Strand Displacement Processes	247
	<i>Lukas Oesinghaus and Friedrich C. Simmel</i>	
13.1	Nucleic Acid Strand Displacement	247
13.1.1	Basics	247
13.1.2	Computing with Strand Displacement Processes	248
13.1.3	Computing with Nucleic Acid Strand Displacement <i>In Vivo</i>	250
13.2	Synthetic Riboregulators	251
13.2.1	First-Generation Riboregulators	251
13.2.2	Toehold Switch Riboregulators	252
13.2.3	Other Transcriptional and Translational Regulators	254
13.3	Combining Strand Displacement and CRISPR Mechanisms	255
13.3.1	A Brief Introduction to CRISPR	255
13.4	Computing Via Nucleic Acid Strand Displacement in Mammalian Cells	258
13.5	Outlook	260
13.5.1	Interfacing Nucleic Acid Computing with Synthetic Biology	260
	References	262
14	Strand Displacement in DNA-Based Nanodevices and Logic	265
	<i>Antoine Bader and Scott L. Cockcroft</i>	
14.1	An Introduction to Strand Displacement Reactions	265
14.1.1	External Control of Strand Displacement Reactions	265
14.1.2	The Toehold Exchange Mechanism	268
14.2	Dynamic Reconfiguration of Structural Devices	268
14.3	Stepped and Autonomous DNA Walkers	271
14.4	Early Breakthroughs in DNA Computing	274
14.4.1	Hamiltonian Paths	275
14.4.2	Satisfiability (SAT) Problem	277
14.5	DNA-Based Molecular Logic	279
14.5.1	Computing with Boolean Logic	279
14.5.2	Deoxyribozyme Logic Gates	280

14.5.3	Autonomous DNA Translators	282
14.5.4	Catalytic Systems for Signal Amplification	285
14.6	Future Prospects for Strand Displacement-Based Devices	286
14.6.1	DNA Chemical Reaction Networks	286
14.6.2	DNA Nanotechnology Goes <i>In Vivo</i>	287
	Acknowledgment	289
	References	289
15	Development and Application of Catalytic DNA in Nanoscale Robotics	293
	<i>David Arredondo, Matthew R. Lakin, Darko Stefanovic, and Milan N. Stojanovic</i>	
15.1	Introduction	293
15.2	Brief History of DNAzymes	293
15.3	Experimental Implementations	296
15.4	DNAzyme Walkers	298
15.5	Statistical Mechanics and Simulation	300
15.6	Conclusions	302
	References	304
16	DNA Origami Transformers	307
	<i>Reem Mokhtar, Tianqi Song, Daniel Fu, Shalin Shah, Xin Song, Ming Yang, and John Reif</i>	
16.1	Introduction	307
16.2	Design	312
16.3	Experimental Demonstrations	316
16.4	Applications	318
16.5	Conclusion	322
	Acknowledgment	322
	References	322
17	Nanopore Decoding for DNA Computing	327
	<i>Hiroki Yasuga, Kan Shoji, and Ryuji Kawano</i>	
17.1	Introduction	327
17.2	Application of Nanopore Technology for Rapid and Label-Free Decoding	330
17.3	Application of Nanopore Decoding in Medical Diagnosis	335
17.4	Conclusions	339
	References	339
18	An Overview of DNA-Based Digital Data Storage	345
	<i>Xin Song, Shalin Shah, and John Reif</i>	
18.1	Introduction	345
18.1.1	Durability and Energy Efficiency	345
18.1.2	Density and Coding Capacity	345
18.1.3	Availability of Supporting Technologies	346
18.2	Components of a DNA Storage System	346

18.2.1	Data Encoding	346
18.2.2	Data Writing	346
18.2.3	Data Storage	348
18.2.4	Data Retrieval	348
18.2.5	Data Decoding	349
18.3	Conclusions and Outlook	350
	Acknowledgments	350
	References	350

19 **Interfacing Enzyme-Based and DNA-Based Computing Systems: From Simple Boolean Logic to Sophisticated Reversible Logic Systems** 353

Evgeny Katz

19.1	Interfacing Enzyme-Based and DNA-Based Computing Systems is a Challenging Goal: Motivations and Approaches	353
19.2	Bioelectronic Interface Transducing Logically Processed Signals from an Enzymatic System to a DNA System	354
19.3	The Bioelectronic Interface Connecting Enzyme-Based Reversible Logic Gates and DNA-Based Reversible Logic Gates: Realization in a Flow Device	362
19.4	Enzyme-Based Fredkin Gate Processing Biomolecular Signals Prior to the Bioelectronic Interface	363
19.5	Reversible DNA-Based Feynman Gate Activated by Signals Produced by the Enzyme-Based Fredkin Gate	368
19.6	Conclusions and Perspectives	371
19.A	Appendix	373
19.A.1	Oligonucleotides Used in the System Mimicking Feynman Gate	373
	References	374

20 **Conclusions and Perspectives: Further Research Directions and Possible Applications** 379

Evgeny Katz

Index	383
--------------	-----

Preface

The use of biomolecular systems for processing information, performing logic operations, computational operations, and even automata performance is a rapidly developing research area. The entire field was named with the general buzzwords, “biomolecular computing” or “biocomputing.” Exciting advances in the area include the use of various biomolecular systems including proteins/enzymes, DNA, RNA, DNAzymes, antigens/antibodies, and even whole biological (usually microbial) cells operating as “hardware” for unconventional computing. The present book concentrates on DNA and RNA molecules utilized for information processing (biocomputing). Extensive ongoing research in the DNA- and RNA-based biocomputing has been motivated by speeding up computation, at least for solving some special problems, due to massive parallel operation of numerous biomolecules. The advantages of the DNA and RNA computing systems are also in their ability to operate in a biological environment for solving biomedical problems in terms of diagnostics and possibly therapeutic action, operating as nanorobots in living organisms. DNA molecules are also applicable as memory material with extremely high data density storage.

The present book summarizes research efforts of many groups in different universities and countries. The book reviews and exemplifies these developments, as well as offering an outlook for possible future research foci. The various topics covered highlight key aspects and the future perspectives of the DNA- and RNA-based computing. The different topics addressed in this book will be of high interest to the interdisciplinary community active in the area of unconventional biocomputing. The readers can find additional complementary material on molecular [1], biomolecular [2], and enzyme-based [3] computing published recently by Wiley-VCH (see book cover pages below). It is hoped that the present book will be important and beneficial for researchers and students working in various areas related to biochemical computing, including biochemistry, materials science, computer science, and so on. Furthermore, the book is aimed to attract young scientists and introduce them to the field while providing newcomers with an enormous collection of literature references. I, indeed, hope that the book will spark the imagination of scientists to further develop the topic.

I would like to conclude this preface by thanking my wife Nina for her support in every respect in the past 49 years. Without her help it would not have been

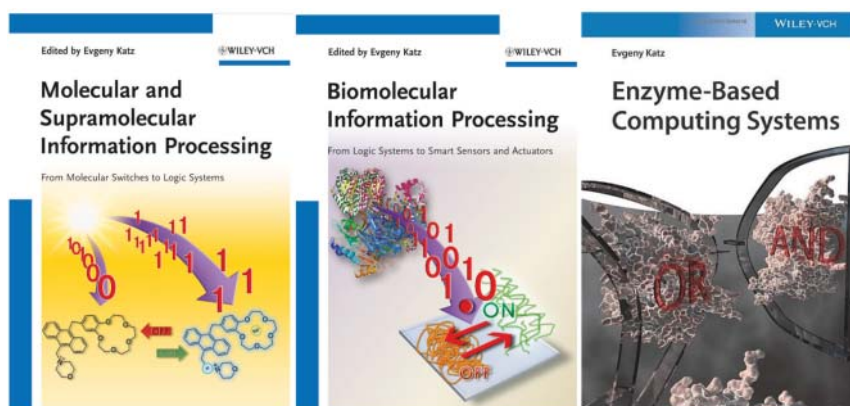
possible to complete this work. Also, cooperation and hard work of all authors working together with me on this edited volume are highly appreciated.

Potsdam, NY, USA
January 2020

Evgeny Katz

References

- 1 Katz, E. (ed.) (2012). *Molecular and Supramolecular Information Processing: From Molecular Switches to Logic Systems*. Weinheim: Wiley-VCH.
- 2 Katz, E. (ed.) (2012). *Biomolecular Information Processing – From Logic Systems to Smart Sensors and Actuators*. Weinheim: Wiley-VCH.
- 3 Katz, E. (2019). *Enzyme-Based Computing Systems*. Weinheim: Wiley-VCH.



DNA Computing: Origination, Motivation, and Goals – Illustrated Introduction

Evgeny Katz

Clarkson University, Department of Chemistry and Biomolecular Science, Potsdam, NY 13699, USA

1.1 Motivation and Applications

Exponential development of computing systems based on silicon materials and binary algorithms formulated as “Moore’s law” [1] (Figure 1.1) is coming to the end being limited by further component miniaturization and by the speed of operation. Conceptually novel ideas are needed to break through these limitations. The quest for novel ideas in the information processing has resulted in several exciting directions in the general area of unconventional computing [2–4], including research in quantum computing [5] and biologically inspired molecular computing [6–9]. Molecular computing systems, generally motivated by mimicking natural biological information processing [10, 11], are not necessarily based on biomolecules and could be represented by synthetic molecules with signal-controlled switchable properties. Synthetic molecular systems and nano-species have been designed to mimic operation of Boolean logic gates and demonstrate basic arithmetic functions and memory units. However, despite progress achieved in assembling synthetic molecular systems performing basic Boolean operations and simple computations [6–9], these systems have limited complexity, and further increase of their complexity is very challenging. A new advance in the development of molecular information systems has been achieved with use of biomolecular species [12] (Figure 1.2) such as deoxyribonucleic acid (DNA) and ribonucleic acid (RNA) [13–16], oligopeptides [17], proteins [18], enzymes [2, 19, 20], antigens/antibodies [21], and even whole biological cells/organisms [22–24] capable of operating in a biological environment [25], borrowing some ideas from systems biology [26]. The advantage of the biomolecular computing systems is their ability to be integrated in artificially designed complex reacting processes mimicking multistep information processing networks. These systems are still far away from the natural information processing in cells but are already much more complex than pure synthetic molecular systems. In fact, biochemical reactions are at the core of the mechanism of life itself, and therefore one could set rather ambitious expectations for how far can (bio)chemical reaction systems be scaled up in complexity, if not speed, for information processing. While in a long perspective

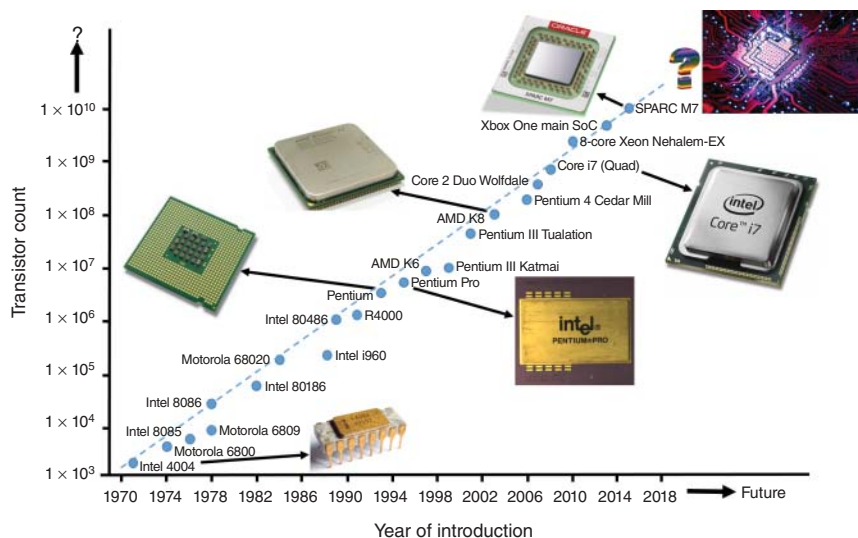


Figure 1.1 Moore's law – exponential increase of transistors on integrated circuit chips. (The plot shown in the figure is based on the data provided by Wikipedia: https://en.wikipedia.org/wiki/Moore%27s_law.) Source: From Katz [2]. Reprinted with the permission of John Wiley and Sons.

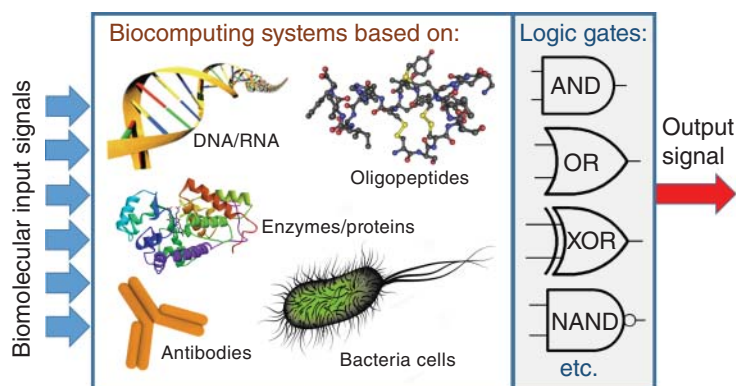


Figure 1.2 Biomolecular computing systems mimicking operation of different Boolean logic gates and circuitries can be based on various species including oligopeptides, enzymes/proteins, DNA/RNA, antibodies, and even whole biological (e.g., microbial) cells. Source: From Katz 2019 [2], Boolean Logic Gates Realized with Enzyme-Catalyzed Reactions – Unusual Look at Usual Chemical Reactions. ChemPhysChem © 2018. Reproduced with the permission of John Wiley & Sons.

a “biocomputer” might become a reality [27], particularly for some special applications, e.g., for solving complex combinatorial problems [28], potentially promising to have an advantage over silicon-based electronic computers due to parallel computing performed by numerous biomolecular units, the present level of technology does not allow any practical use of biomolecular systems for real computational applications. For achieving any practical result soon, some other

applications, different from making a biocomputer, should be considered using the (bio)molecular systems with a limited complexity. One of the immediate possible applications for molecular logic systems is a special kind of biosensing [29–31] where the multiple input signals are logically processed through chemical reactions resulting in YES/NO decisions in the binary (0,1) format. In this subarea of biomolecular logic systems, practical results are already possible at the present level of the system complexity, particularly for biomedical applications [32–35]. Overall, the research in molecular/biomolecular information processing, which has been motivated originally to progress unconventional computing applications, is broadly developing to areas not directly related to computing in its narrow definition. This research is bringing us to novel areas in sensing/biosensing [29–31], switchable “smart” materials controlled by logically processed signals [32–36], bioelectronic devices (e.g., biofuel cells) controlled by external signals [37, 38], signal-controlled release processes [39–43], etc.

1.2 DNA- and RNA-Based Biocomputing Systems in Progress

While the general topics of the biomolecular computing [12] and specifically the enzyme-based computing [44] have been covered with recently published books, the present book is concentrated on the use of DNA and RNA molecules in computing systems, broadly defined as information processing systems. From the time (1953) when James D. Watson and Francis H.C. Crick (Figure 1.3) discovered chemical structure of DNA (Figure 1.4) [45], the progress in the DNA study resulted in many novel fundamental scientific concepts [46–48] and highly important practical applications [49]. Among many other, mostly biomedical applications, DNA molecules have been extensively studied over last two decades for unconventional biomolecular computing [13, 15, 50–55], following the pioneer work (1994) by Leonard M. Adleman [28, 56] (Figure 1.5). In his seminal work Adleman demonstrated for the first time computational use of DNA molecules for solving a “traveling salesman problem,” Hamiltonian path problem. Actually, this work initiated (bio)molecular computing research not necessary using DNA molecules.

The “traveling salesman problem” asks the following question [57–59]: “Given a list of cities and the distances between each pair of cities, what is the shortest possible route that visits each city and returns to the origin city?” It is a problem in combinatorial optimization, important in theoretical computer science. It is frequently used to test computational algorithms and computer hardware. In general, the traveling salesman problem is hard to solve, particularly when the number of the visited cities is increasing. Adleman solved the problem for seven cities only (Figure 1.6), which was rather a trivial task, but importantly it was solved using computational power of DNA reactions [28, 56]. The DNA molecules hybridized in a special way to solve the problem, and the computation was performed by numerous DNA sequences (actually rather short oligonucleotides) operating in parallel. This was important conceptual difference from

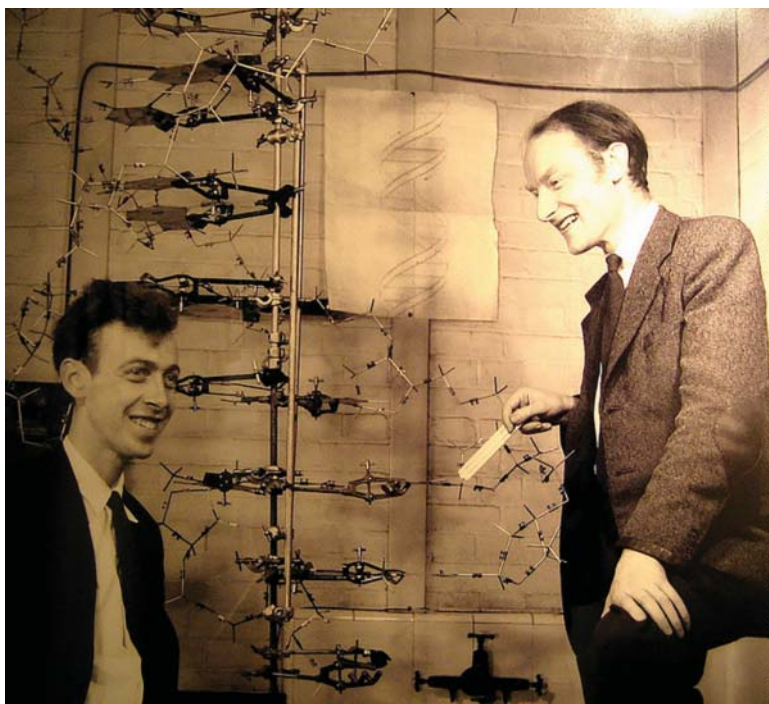


Figure 1.3 The discoverers of the structure of DNA. James Watson (b.1928) at left and Francis Crick (1916–2004), with their model of part of a DNA molecule in 1953. Photographed in the Cavendish Laboratory, University of Cambridge, UK, in May 1953. Source: From Watson and Crick [45]. <https://cnx.org/contents/8M7b3dzJ@2/DNA-Structure>. Licensed Under CC BY 4.0.

Si-based electronic computers that perform all operations in a sequence. The computation in Adleman's experiment was performed at the speed of 1014 operations per second, a rate of 100 teraflops or 100 trillion floating point operations per second (comparable to the fastest presently available quantum computer) – all because of massively parallel processing capabilities of the DNA computing operation [54, 60]. The promise for extremely fast computation ignited the interest to the DNA computing concept, then being extended to a broader area of molecular [8] and biomolecular [12] computing. Despite the fact that the practical results have not been obtained after almost 25 years of the active research, optimistic expectations for building DNA computers are still present [27, 55, 61, 62]. The advantages of the DNA and RNA computing systems are not only in their potentially high speed of operation due to the parallel information processing but also in their ability to operate in a biological environment for solving biomedical problems in terms of diagnostics and possibly therapeutic action (theranostics) [16, 63], for example, for logic control of gene expression [64]. RNA-based computing systems are particularly promising for *in vivo* operation, thus being excellent candidates for nanomedicine with implemented Boolean logic [65]. DNA computers can operate as a Turing machine [51] and can be sophisticated enough to mimic neural network computations similar

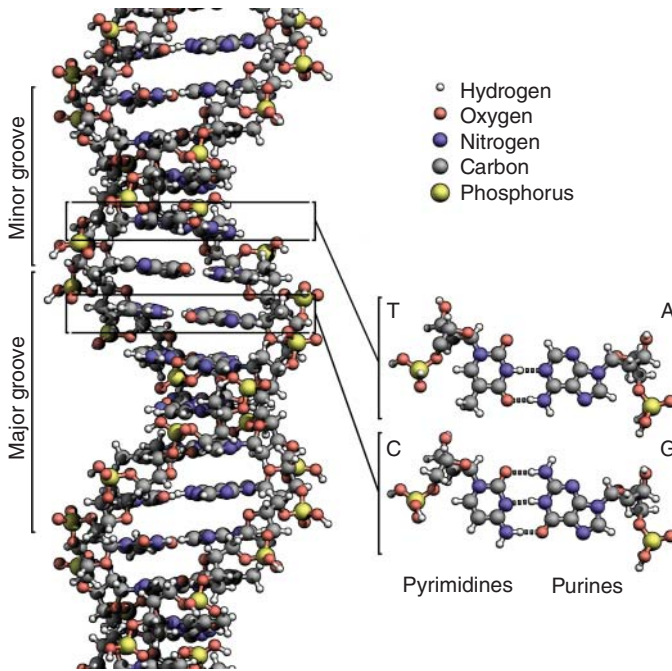


Figure 1.4 The structure of the DNA double helix. The atoms in the structure are color-coded by element, and the detailed structures of two base pairs are shown in the bottom right. Source: From Watson and Crick [45]. Also adapted from Zephyris, DNA Structure, Wikimedia commons, 2011. Public Domain.

Figure 1.5 Leonard Adleman – a pioneer of the biomolecular computing; the photo of 1993 when the first experiments on DNA computing were running. Source: Courtesy of Prof. Leonard Adleman.



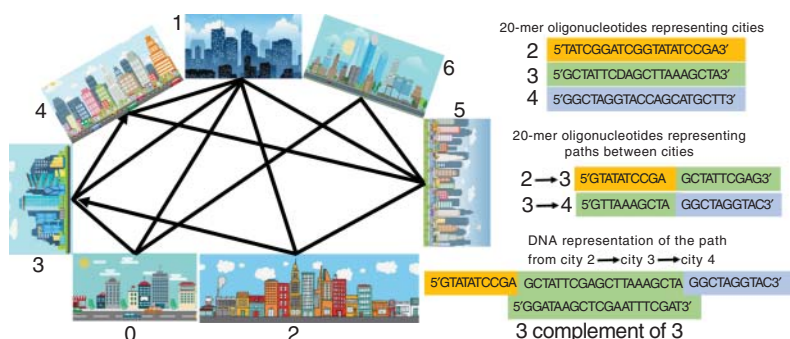


Figure 1.6 The principle of Leonard Adleman's DNA computer to solve the "traveling salesman problem" (see detailed explanation in Ref. [54]). Source: Based on Parker [54].

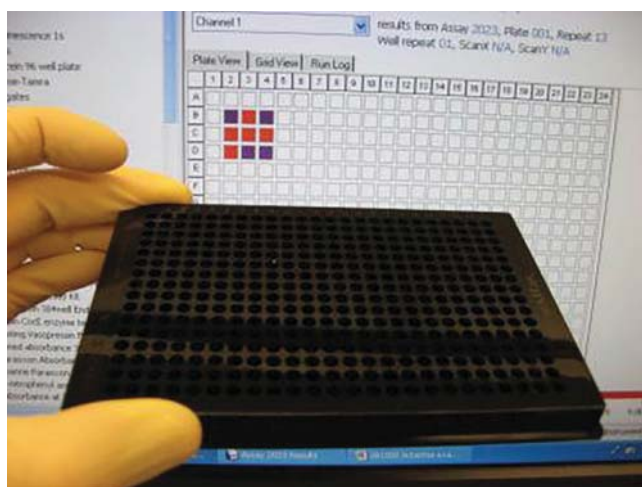


Figure 1.7 The DNA computer playing the tic-tac-toe game. Shown in the foreground is a cell culture plate containing pieces of DNA that code for possible "moves." A display screen (background) shows that the computer (red squares) has won the game against a human opponent (blue). Source: Courtesy of Prof. Milan Stojanovic, Columbia University.

to human brain, obviously in a very simplified way [66]. The DNA computing systems playing a tic-tac-toe game against human have been "smart" enough to win [13, 67–69] (Figure 1.7).

The progress in the DNA computing has been based on three major developments: (i) the use of sophisticated DNA structures (e.g., origami), (ii) the use of more powerful instrumentation for automatic operation of DNA computing steps (DNA chips), and (iii) specialized programming languages specifically developed for the DNA computing. The invention of the DNA origami structures [70, 71] – nanoscale folding of DNA resulting in nonarbitrary two- and three-dimensional shapes [72, 73] (Figure 1.8) – resulted in further sophistication of the DNA computing systems [74], capable of operating as nanorobots in living organisms [75–77]. The use of DNA microarrays (DNA chips [78]) allowed

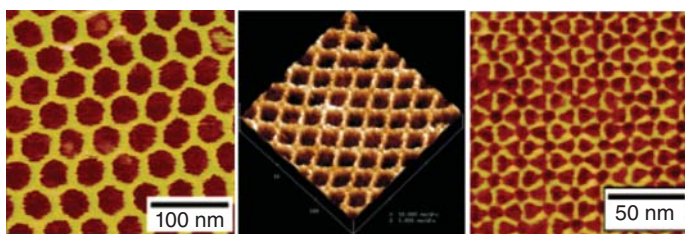


Figure 1.8 Atomic force microscopy (AFM) images of DNA origami with different shapes. Source: From Hong et al. [73]. Reprinted with the permission of American Chemical Society.

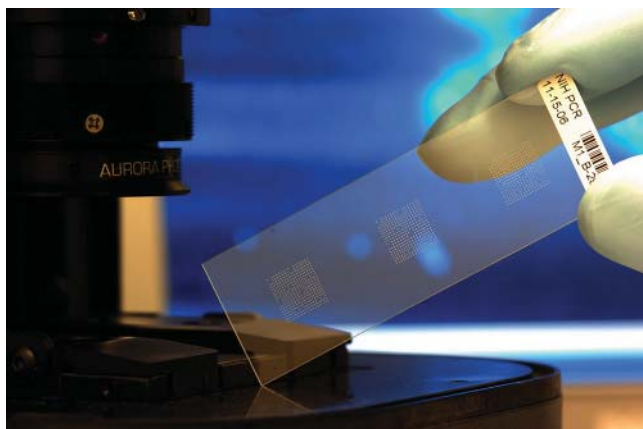


Figure 1.9 An example of a DNA chip used in the DNA sensing and computing. The chip represents a DNA microarray as a collection of microscopic DNA spots attached to a solid surface. Each DNA spot contains picomoles of a specific DNA sequence. The chip allows simultaneous analysis of many DNA probes. The analysis of the probes can be performed optically (as it is in the present example) or electrochemically (then the chip should be based on a microelectrode array). Source: Courtesy of Argonne National Laboratory and Mr. Calvin Chimes.

simultaneous analysis of large numbers of DNA probes [53], thus introducing a powerful hardware for the DNA computing (Figure 1.9). A special computational language, DNA strand displacement (DSD) tool, similar to programming languages used in electronic computers, has been developed by scientists at Microsoft Research for programming DNA computing [79, 80] (Figure 1.10). The language uses DSD as the main computational mechanism, which allows devices to be designed solely in terms of nucleic acids. DSD is a first step toward the development of design and analysis tools for DSD and complements the emergence of novel implementation strategies for DNA computing. The DNA computation can be performed in living cells by DNA-encoded circuits that process sensory information and control biological functions. A special computing language, “Cello,” has been developed for programming DNA logic operations *in vivo* [81]. Overall, the use of computing languages simplified the design of DNA computing systems of high complexity.

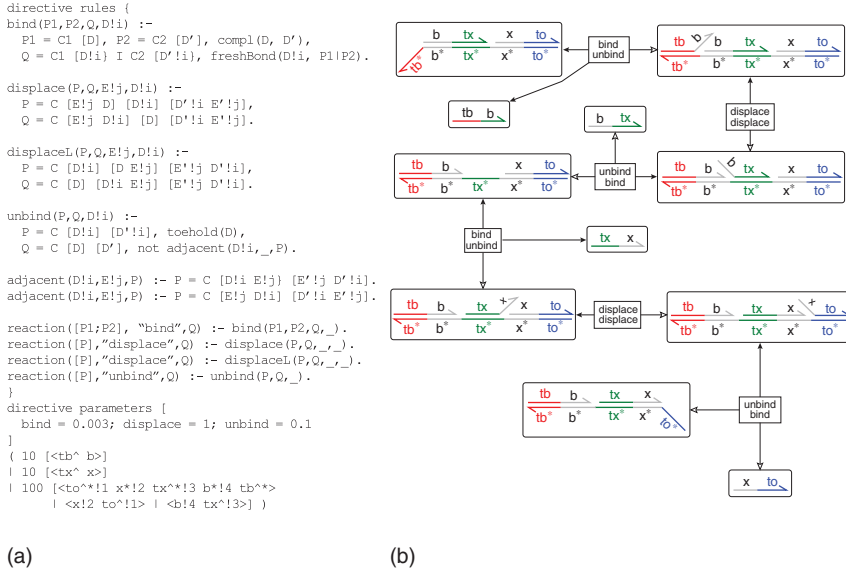


Figure 1.10 Logic program (a) and automatically generated chemical reaction network (b) for a DNA strand displacement example. Source: Adapted from Spaccasassi et al. 2019 [80] with permission; open access paper.

1.3 DNA-Based Information Storage Systems

Human civilization generates huge amount of information increasing exponentially and required to be stored. The total digital information today amounts to 3.52×10^{22} bits globally and at its consistent exponential rate of growth is expected to reach 3×10^{24} bits by 2040 [82]. Data storage density of silicon chips is limited, and magnetic tapes used to maintain large-scale permanent archives begin to deteriorate within 20 years. Alternative methods/materials for storing high density/large amount of information with reliable preservation over long period of time are urgently needed. DNA has been recognized as a promising natural medium for information storage [83]. Indeed, the DNA molecules were created by nature to keep the genetic code, which can be easily “written” and “read” by biomolecular systems. With information retention times that range from thousands to millions of years, volumetric density 10^3 times greater than flash memory, and energy of operation 10^8 times less, DNA is a memory storage material viable and compelling alternative to electronic memory. Recent research in the area of information storage with DNA molecules resulted in the proof-of-the-concept systems [82, 84–87], while the practical use of the DNA memory systems is only limited by technological problems. Both processes in the information storage with DNA, “writing” and “reading” information, are available, but they are not as simple as needed to be implemented with the present computer technology. In other words, the DNA memory is technically possible, but it is not convenient enough to be integrated with standard Si-based computers operated by end users.

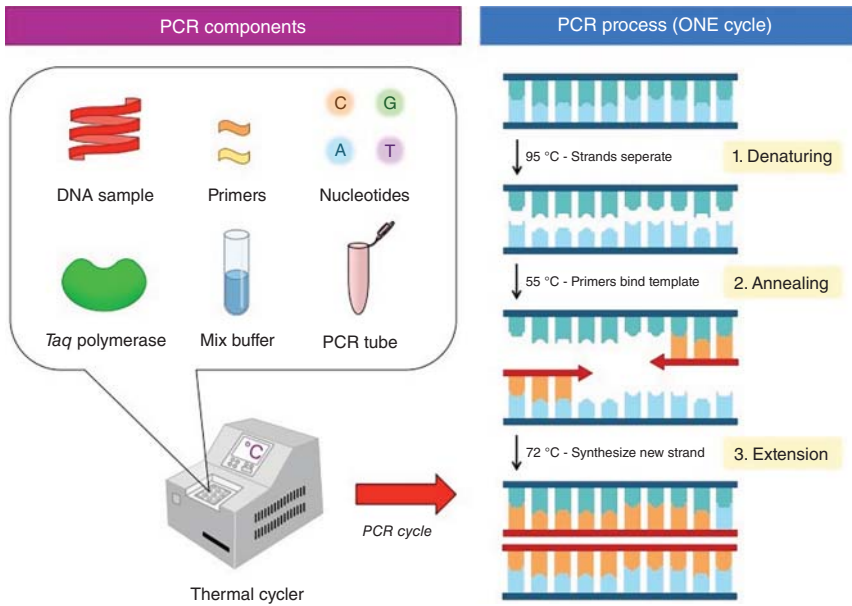


Figure 1.11 PCR method for copying DNA molecules: a thermal cycler, components of the reaction mixture, and reaction steps. For detailed description of the method and instrument see Refs. [89–92]. Source: From Keagile Bati, Polymerase Chain Reaction: Innovation that Revolutionized Molecular Biology, Nov 2018. Public Domain.

Synthetic procedure for production of DNA molecules with specific nucleotide sequences is well known in organic chemistry [88] and can be used to “write” information in the DNA. Once the DNA molecules with the encoded information are prepared, they can be multiplied using a polymerase chain reaction (PCR) [89–92] (Figure 1.11), which is a technique to make many copies of a specific DNA *in vitro* (in a test tube rather than an organism). This technique is rather advanced in the instrumental realization but still requires special apparatus that cannot be connected easily to an electronic computer, at least at the end-user convenience. “Reading” the DNA-encoded information (DNA sequencing [93]) was advanced during the Human Genome Project [94] and presently is very technologically effective. Further improvements in sequencing throughput ($>10^4$) and parallelization ($>10^7$) are expected in the next five years [84]. Emerging technologies such as nanopore sequencing [95] will further reduce errors, cost, time, and energetics during reading the DNA-encoded information. While future advances can result in novel technological approaches, already available techniques based on the DNA memo-chips have been tested [96]. A simple chemical, rather than electronic, apparatus operating as the end-to-end automatic DNA data storage was designed and demonstrated the automatic “writing”–“reading” DNA processes [97] (Figure 1.12). The recent research efforts opened the way toward practical, high-capacity, low-maintenance information storage in synthesized DNA [98–100]. As an example, a 5.27-megabit book was stored using DNA microchips and then read the book by using the

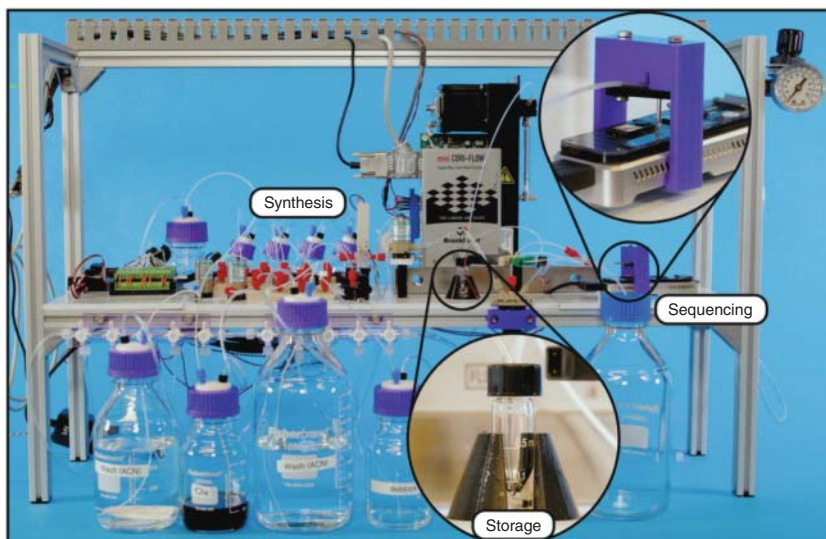


Figure 1.12 Apparatus operating as the end-to-end automatic DNA data storage allowing automatic “writing”–“reading” DNA processes. Source: From Takahashi et al. [97]. <https://www.nature.com/articles/s41598-019-41228-8>. Licensed Under CC BY 4.0.

DNA sequencing [101]. Other, even more impressive, examples demonstrated encoding the pixel values of black-and-white images and a short movie into the genomes of a population of living bacteria and then retrieving them back by the DNA sequencing [102].

1.4 Short Conclusions and Comments on the Book

Overall, the DNA computing is a multidisciplinary research area with major contributions from synthetic biology, nanotechnology, computer science, chemical engineering, biosensing and biotechnology, biology and medicine, etc. Some of the research areas are already reaching the mature states, while others are still in the infancy. It is still not easy to predict in what direction the research will go and what applications will be more benefiting from the DNA computing. In the most probability, practical applications will be in two major subareas: medicine with the DNA information processing nanorobotic systems operating *in vivo* [103, 104] and large data storage systems providing extremely high density of the information storage [84, 105]. Many other applications of the DNA computing are in the research and discussion [106, 107]. However, it is quite unexpected that the DNA computing will come to the end users instead of standard electronic computers, at least in the short perspective.

The present book, composed of the chapters written by the best experts in the field, covers all subtopics of the DNA computing, including the design of Boolean logic gates and circuitries, programming the DNA information processing systems, their biomedical applications and operation *in vivo*, DNA data storage and

nanopore DNA decoding, and interfacing of the DNA computing with enzyme logic systems, and many more detailed explanations on the DNA and RNA computing with many references and illustrations.

References

- 1 Moore, G.E. (1998). *Proc. IEEE* 86: 82–85.
- 2 Katz, E. (2019). *ChemPhysChem* 20: 9–22.
- 3 Calude, C.S., Costa, J.F., Dershowitz, N. et al. (eds.) (2009). *Unconventional Computation*, Lecture Notes in Computer Science, vol. 5715. Berlin: Springer.
- 4 Adamatzky, A. (ed.) (2017). *Advances in Unconventional Computing*, Emergence, Complexity and Computation, 2 volumes. Switzerland: Springer.
- 5 Mermin, N.D. (2007). *Quantum Computer Science: An Introduction*. Cambridge: Cambridge University Press.
- 6 Szacilowski, K. (2012). *Infochemistry – Information Processing at the Nanoscale*. Chichester: Wiley.
- 7 de Silva, A.P. (2013). *Molecular Logic-Based Computation*. Cambridge: Royal Society of Chemistry.
- 8 Katz, E. (ed.) (2012). *Molecular and Supramolecular Information Processing – From Molecular Switches to Logic Systems*. Weinheim: Wiley-VCH.
- 9 Sienko, T. (ed.) (2003). *Molecular Computing* (Series Eds.: Adamatzky, A., Conrad, M., and Rambidi, N.G.). Cambridge, MA: MIT Press.
- 10 Spitzer, N.C. and Sejnowski, T.J. (1997). *Science* 277: 1060–1061.
- 11 Kampfnér, R.R. (1989). *BioSystems* 22: 223–230.
- 12 Katz, E. (ed.) (2012). *Biomolecular Computing – From Logic Systems to Smart Sensors and Actuators*. Weinheim: Wiley-VCH.
- 13 Stojanovic, M.N., Stefanovic, D., and Rudchenko, S. (2014). *Acc. Chem. Res.* 47: 1845–1852.
- 14 Stojanovic, M.N. and Stefanovic, D. (2011). *J. Comput. Theor. Nanosci.* 8: 434–440.
- 15 Ezziane, Z. (2006). *Nanotechnology* 17: R27–R39.
- 16 Xie, Z., Wroblewska, L., Prochazka, L. et al. (2011). *Science* 333: 1307–1311.
- 17 Ashkenasy, G., Dadon, Z., Alesebi, S. et al. (2011). *Isr. J. Chem.* 51: 106–117.
- 18 Unger, R. and Moulton, J. (2006). *Proteins* 63: 53–64.
- 19 Katz, E. and Privman, V. (2010). *Chem. Soc. Rev.* 39: 1835–1857.
- 20 Katz, E. (2015). *Curr. Opin. Biotechnol.* 34: 202–208.
- 21 Halámek, J., Tam, T.K., Chinnapareddy, S. et al. (2010). *J. Phys. Chem. Lett.* 1: 973–977.
- 22 Rinaudo, K., Bleris, L., Maddamsetti, R. et al. (2007). *Nat. Biotechnol.* 25: 795–801.
- 23 Arugula, M.A., Shroff, N., Katz, E., and He, Z. (2012). *Chem. Commun.* 48: 10174–10176.
- 24 Adamatzky, A. (2010). *Physarum Machines – Computers from Slime Mould*. London: World Scientific.

- 25 Kahan, M., Gil, B., Adar, R., and Shapiro, E. (2008). *Physica D* 237: 1165–1172.
- 26 Alon, U. (2006). *An Introduction to Systems Biology: Design Principles of Biological Circuits*. Boca Raton, FL: Chapman & Hall/CRC.
- 27 Benenson, Y. (2009). *Mol. Biosyst.* 5: 675–685.
- 28 Adleman, L.M. (1994). *Science* 266: 1021–1024.
- 29 Katz, E., Wang, J., Privman, M., and Halámková, J. (2012). *Anal. Chem.* 84: 5463–5469.
- 30 Wang, J. and Katz, E. (2011). *Isr. J. Chem.* 51: 141–150.
- 31 Wang, J. and Katz, E. (2010). *Anal. Bioanal. Chem.* 398: 1591–1603.
- 32 Halámková, L., Halámková, J., Bocharova, V. et al. (2012). *Analyst* 137: 1768–1770.
- 33 Halámková, J., Windmiller, J.R., Zhou, J. et al. (2010). *Analyst* 135: 2249–2259.
- 34 Minko, S., Katz, E., Motornov, M. et al. (2011). *J. Comput. Theor. Nanosci.* 8: 356–364.
- 35 Tokarev, I., Gopishetty, V., Zhou, J. et al. (2009). *ACS Appl. Mater. Interfaces* 1: 532–536.
- 36 Pita, M., Minko, S., and Katz, E. (2009). *J. Mater. Sci. - Mater. Med.* 20: 457–462.
- 37 Katz, E. and Minko, S. (2015). *Chem. Commun.* 51: 3493–3500.
- 38 Katz, E. (2010). *Electroanalysis* 22: 744–756.
- 39 Okhokhonin, A.V., Domanskyi, S., Filipov, Y. et al. (2018). *Electroanalysis* 30: 426–435.
- 40 Filipov, Y., Gamella, M., and Katz, E. (2018). *Electroanalysis* 30: 1281–1286.
- 41 Gamella, M., Privman, M., Bakshi, S. et al. (2017). *ChemPhysChem* 18: 1811–1821.
- 42 Gamella, M., Zakharchenko, A., Guz, N. et al. (2017). *Electroanalysis* 29: 398–408.
- 43 Katz, E., Pingarrón, J.M., Mailloux, S. et al. (2015). *J. Phys. Chem. Lett.* 6: 1340–1347.
- 44 Katz, E. (2019). *Enzyme-Based Computing Systems*. Wiley-VCH.
- 45 Watson, J.D. and Crick, F.H.C. (1953). *Nature* 171: 737–738.
- 46 Micklos, D. and Freyer, G. (2003). *DNA Science: A First Course*, 2e. New York: Cold Spring Harbor Laboratory Press.
- 47 Calladine, C.R., Drew, H., Luisi, B., and Travers, A. (2004). *Understanding DNA: The Molecule and How It Works*, 3e. San Diego, CA: Elsevier Academic Press.
- 48 Douglas, K. (2017). *DNA Nanoscience*. Boca Raton, FL: CRC Press.
- 49 Fitzgerald-Hayes, M. and Reichsman, F. (2009). *DNA and Biotechnology*, 3e. Amsterdam, Imprint: Academic Press: Elsevier.
- 50 Boneh, D., Dunworth, C., Lipton, R.J., and Sgall, J. (1996). *Discrete Appl. Math.* 71: 79–94.
- 51 Rozen, D.E., McGrew, S., and Ellington, A.D. (1996). *Curr. Biol.* 6: 254–257.
- 52 Eghdami, H. and Darehmira, M. (2012). *Artif. Intell. Rev.* 38: 223–235.
- 53 Ogiwara, M. and Ray, A. (2000). *Nature* 403: 143–144.
- 54 Parker, J. (2003). *EMBO Rep.* 4: 7–10.

- 55 Mehrotra, A. (2015). *Int. J. Latest Technol. Eng. Manage. Appl. Sci. (IJL-TEMAS)* 4 (12): 62–67.
- 56 Adleman, L.M. (1998). *Sci. Am.* 279 (2): 54–61.
- 57 Lawler, E.L., Lenstra, J.K., Rinnooy Kan, A.H.G., and Shmoys, D.B. (eds.) (1995). *The Traveling Salesman Problem: A Guided Tour of Combinatorial Optimization*. Chichester: Wiley-Interscience.
- 58 Laporte, G. (1992). *Eur. J. Oper. Res.* 59: 231–247.
- 59 Applegate, D.L., Bixby, R.E., Chvátal, V., and Cook, W.J. (2006). *The Traveling Salesman Problem – A Computational Study*. Princeton, NJ: Princeton University Press.
- 60 Reif, J.H. (1999). *Algorithmica* 25: 142–175.
- 61 Kumar, S.N. (2015). *Am. J. Nanomater.* 3 (1): 1–14.
- 62 Moe-Behrens, G.H.-G. (2013). *Comput. Struct. Biotechnol. J.* 7 (Art. No.: e201304003).
- 63 Bonnet, J., Yin, P., Ortiz, M.E. et al. (2013). *Science* 340: 599–603.
- 64 Benenson, Y., Gil, B., Ben-Dor, U. et al. (2004). *Nature* 429: 423–429.
- 65 Green, A.A., Kim, J., Ma, D. et al. (2017). *Nature* 548: 117–121.
- 66 Qian, L., Winfree, E., and Bruck, J. (2011). *Nature* 475: 368–372.
- 67 Macdonald, J., Stefanovic, D., and Stojanovic, M.N. (2008). *Sci. Am.* 299 (5): 84–91.
- 68 Stojanovic, M.N. and Stefanovic, D. (2003). *Nat. Biotechnol.* 21: 1069–1074.
- 69 Elstner, M. and Schiller, A. (2015). *J. Chem. Inf. Model.* 55: 1547–1551.
- 70 Seeman, N.C. (1982). *J. Theor. Biol.* 99: 237–247.
- 71 Rothmund, P.W.K. (2006). *Nature* 440: 297–302.
- 72 Jun, H., Zhang, F., Shepherd, T. et al. (2019). *Sci. Adv.* 5 (Art. No.: eaav0655).
- 73 Hong, F., Zhang, F., Liu, Y., and Yan, H. (2017). *Chem. Rev.* 117: 12584–12640.
- 74 Wang, D.F., Fu, Y.M., Yan, J. et al. (2014). *Anal. Chem.* 86: 1932–1936.
- 75 Amir, Y., Ben-Ishay, E., Levner, D. et al. (2014). *Nat. Nanotechnol.* 9: 353–357.
- 76 Kogikoski, S. Jr., Paschoalino, W.J., and Kubota, L.T. (2018). *TrAC, Trends Anal. Chem.* 108 (Art. No.: 88e97).
- 77 Endo, M. and Sugiyama, H. (2018). *Molecules* 23 (Art. No.: 1766).
- 78 Ramsay, G. (1998). *Nat. Biotechnol.* 16: 40–44.
- 79 Phillips, A. and Cardelli, L. (2009). *J. R. Soc. Interface* 6: S419–S436.
- 80 Spaccasassi, C., Lakin, M.R., and Phillips, A. (2019). *ACS Synth. Biol.* 8: 1530–1547.
- 81 Nielsen, A.A.K., Der, B.S., Shin, J. et al. (2016). *Science* 352 (6281): 53.
- 82 Panda, D., Molla, K.A., Baig, M.J. et al. (2018). *3 Biotech* 8 (Art. No.: 239).
- 83 Baum, E.B. (1995). *Science* 268: 584–585.
- 84 Zhirnov, V., Zadegan, R.M., Sandhu, G.S. et al. (2016). *Nat. Mater.* 15: 366–370.
- 85 Ceze, L., Nivala, J., and Strauss, K. (2019). *Nat. Rev. Genet.* 20: 456–466.
- 86 Zhang, S.F., Huang, B.B., Song, X.M. et al. (2019). *3 Biotech* 9 (Art. No.: 342).
- 87 Tomek, K.J., Volkel, K., Simpson, A. et al. (2019). *ACS Synth. Biol.* 8: 1241–1248.

- 88 Ma, S., Tang, N., and Tian, J. (2012). *Curr. Opin. Chem. Biol.* 16: 260–267.
- 89 Behlke, M.A., Berghof-Jäger, K., Brown, T. et al. (2019). *Polymerase Chain Reaction: Theory and Technology*. Poole: Caister Academic Press.
- 90 Garibyan, L. and Avashia, N. (2013). *J. Invest. Dermatol.* 133 (3), art No. e6.
- 91 Alves Valones, M.A., Lima Guimarães, R., Cavalcanti Brandão, L.A. et al. (2009). *Braz. J. Microbiol.* 40 (1): 1–11.
- 92 Powledge, T.M. (2004). *Adv. Physiol. Educ.* 28: 44–50.
- 93 Shendure, J., Balasubramanian, S., Church, G.M. et al. (2017). *Nature* 550: 345–353.
- 94 Dolgin, E. (2009). *Nature* 462: 843–845.
- 95 Agah, S., Zheng, M., Pasquali, M., and Kolomeisky, A.B. (2016). *J. Phys. D: Appl. Phys.* 49 (Art. No.: 413001).
- 96 Stefano, G.B., Wang, F.Z., and Kream, R.M. (2018). *Med. Sci. Monit.* 24: 1185–1187.
- 97 Takahashi, C.N., Nguyen, B.H., Strauss, K., and Ceze, L. (2019). *Sci. Rep.* 9 (Art. No.: 4998).
- 98 Goldman, N., Bertone, P., Chen, S. et al. (2013). *Nature* 494: 77–80.
- 99 Anavy, L., Vaknin, I., Atar, O. et al. (2019). *Nat. Biotechnol.* 37: 1229–1236.
- 100 Organick, L., Ang, S.D., Chen, Y.-J. et al. (2018). *Nat. Biotechnol.* 36: 242–248.
- 101 Church, G.M., Gao, Y., and Kosuri, S. (2012). *Science* 337: 1628.
- 102 Shipman, S.L., Nivala, J., Macklis, J.D., and Church, G.M. (2017). *Nature* 547: 345–349.
- 103 Li, S., Jiang, Q., Liu, S. et al. (2018). *Nat. Biotechnol.* 36: 258–264.
- 104 Ma, W., Zhan, Y., Zhang, Y. et al. (2019). *Nano Lett.* 19: 4505–4517.
- 105 De Silva, P.Y. and Ganegoda, G.U. (2016). *BioMed Res. Int.* 2016 (Art. No.: 8072463).
- 106 Tagore, S., Bhattacharya, S., Islam, M.A., and Islam, M.L. (2010). *J. Proteomics Bioinform.* 3: 234–243.
- 107 Currin, A., Korovin, K., Ababi, M. et al. (2017). *J. R. Soc. Interface* 14 (Art. No.: 20160990).