



HAL
open science

Molecular Computation for Molecular Classification

Coline Kieffer, Anthony J Genot, Yannick Rondelez, Guillaume Gines

► **To cite this version:**

Coline Kieffer, Anthony J Genot, Yannick Rondelez, Guillaume Gines. Molecular Computation for Molecular Classification. *Advanced Biology*, 2023, 7 (3), 10.1002/adbi.202200203 . hal-04287205

HAL Id: hal-04287205

<https://hal.science/hal-04287205>

Submitted on 15 Nov 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Molecular Computation for Molecular Classification

Coline Kieffer, Anthony J. Genot, Yannick Rondelez,* and Guillaume Gines*

DNA as an informational polymer has, for the past 30 years, progressively become an essential molecule to rationally build chemical reaction networks endowed with powerful signal-processing capabilities. Whether influenced by the silicon world or inspired by natural computation, molecular programming has gained attention for diagnosis applications. Of particular interest for this review, molecular classifiers have shown promising results for disease pattern recognition and sample classification. Because both input integration and computation are performed in a single tube, at the molecular level, this low-cost approach may come as a complementary tool to molecular profiling strategies, where all biomarkers are quantified independently using high-tech instrumentation. After introducing the elementary components of molecular classifiers, some of their experimental implementations are discussed either using digital Boolean logic or analog neural network architectures.

1. Introduction

Living organisms synthesize and process a large range of chemically diverse molecules (proteins, enzymes, peptides, metabolites, ions, nucleic acids, etc.). The expression level of these compounds is continuously monitored in space and time, and adjusted by dedicated regulatory systems and feedback mechanisms. The levels of some of these molecules get dysregulated (up- or down-regulated) when the organism faces a disease, such as cancers, neurodegenerative diseases, or infections.^[1–4] Assays that monitor a single marker, that is, a molecule that, on its own, is representative of a cellular state, have been intensely developed for diagnostics,^[5] but in many cases, the detection of

a panel of disease-related markers is necessary to provide a robust and informative diagnostic (Figure 1A).^[6–8] For example, an *in silico* multi-marker classifier based on gene expression profiles from whole blood samples has been proposed to discriminate bacterial from viral cause in acute respiratory infections.^[9] Transcriptional signatures of isolated leukocytes have been used to identify specific pathogens in acute infections.^[10] By combining 61 genetic markers with eight circulating protein levels, it was possible to detect and stratify eight common cancer types from a blood sample, while simultaneously identifying the tissue of origin.^[6] Various SARS-Cov2 viral strains could be identified by Raman spectroscopy, which scrutinizes multiple slight variations at the molecular level.^[11]

More generally, the use of multidimensional molecular signatures may improve early diagnosis, prognosis, therapy monitoring, and personalized treatments. The current dominant strategy consists in profiling a patient sample, that is, measuring multiple genomic, transcriptomic, proteomic, or metabolomic markers to construct a multidimensional profile, which is then compared to typical disease signatures identified from statistical analysis or *in silico* training (Figure 1B). This approach involves multidimensional assays such as multiplex PCR or ELISA, next generation sequencing, microarrays, mass spectrometry, or microfluidics instruments, and generally requires costly equipment, lengthy protocols, skilled technicians, and complex bioanalysis pipelines to accurately quantify each of these markers independently and interpret the result.

Comparatively, gene regulatory networks empower cells with the ability to integrate multiple stimuli such as concentrations of internal or external molecules, and trigger an appropriate response. For instance, the bone morphogenetic protein pathway involves a large set of receptors able to interact with a wide range of ligands, playing a fundamental role in organizing tissue architecture, beyond osteogenesis.^[12] Specific receptors are also at play to provide immune cells with pattern recognition capabilities for recognizing foreign elements.^[13] At the tissue level, neural networks compute complex classification problems allowing for pattern recognition. A well-known example is the olfactory circuit which can classify thousands of different odors from promiscuous receptor inputs.^[14] In these illustrations, cells and brains do not measure in the analytical sense each of the input stimuli; they “directly recognize” a multidimensional pattern, which may correspond to a threat, a smell, etc.

Inspired by these biological examples, where pattern recognition and classification are performed at the molecular

C. Kieffer, Y. Rondelez, G. Gines

Laboratoire Gulliver

UMR 7083

CNRS

ESPCI Paris

PSL Research University

10 rue Vauquelin, Paris 75005, France

E-mail: yannick.rondelez@espci.fr; guillaume.gines@espci.fr

A. J. Genot

LIMMS

CNRS-Institute of Industrial Science

IRL 2820

University of Tokyo

Tokyo 153-8505, Japan

 The ORCID identification number(s) for the author(s) of this article can be found under <https://doi.org/10.1002/adbi.202200203>.

© 2023 The Authors. Advanced Biology published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

DOI: 10.1002/adbi.202200203

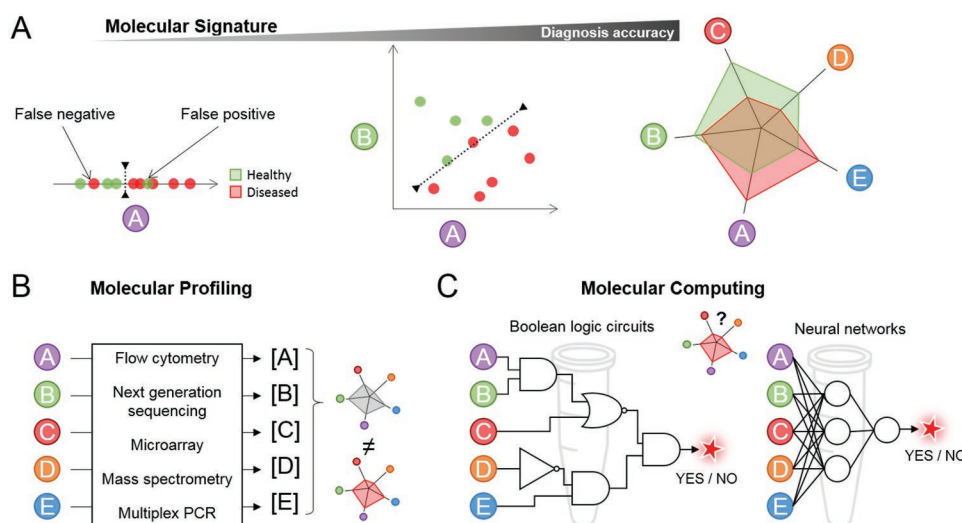


Figure 1. A) A molecular signature consists of a panel of biomarkers (A, B, C, D, and E in this example) that, taken together, allow accurate diagnosis, prognosis, or patient stratification. B) Molecular profiling refers to the separate quantification of these biomarkers to construct the patient molecular profile, which can be compared to the disease signature(s). C) Molecular classifiers represent a class of artificial (bio)chemical circuits emulating Boolean logic or neural networks, executed to classify samples using in-moleculo computation from the biomarker inputs.

level, man-made molecular classifiers have been created for multi-dimensional biosensing applications (Figure 1C).^[15–17] In this review, we will describe these early examples of artificial (bio)chemical networks engineered to classify samples based on multiple input sensing and computation. Because in these systems the integration of multiple chemical inputs into a low dimensional output (e.g., “healthy”/“disease”) is performed *in moleculo*, we will refer to them as molecular classifiers. Most notably, DNA nanotechnology, a field, inaugurated 30 years ago,^[18] is perfectly fitted for building molecular classifiers for several reasons: i) It can be interfaced with most biological molecules, either by direct hybridization with nucleic acids (NA), or via antibody, aptamers or transcription factors for proteins or small metabolites; ii) it can be connected to a variety of readouts including optical (colorimetric, fluorescent) and electronic signals for low-cost or even instrument-free setups; iii) DNA reaction networks are based on simple interaction rules between DNA strands; iv) an ever-expanding catalog of biocatalysts is available for modifying at will the biopolymer, bringing countless design possibilities on top of DNA hybridization rules; v) synthetic DNA is cheap, an argument most certainly central for the development of screening tests and clinical assays accessible to all. We discuss here the latest advances in DNA nanotechnology that are specifically targeted at solving classification problems at the molecular level, with a particular emphasis on their application to diagnosis. Section 2 recapitulates the elementary modules that compose a molecular classifier. We then briefly describe in Section 3 the molecular toolboxes that are compatible with the design of classification algorithms. Section 4 is dedicated to experimental demonstrations of molecular classifiers, using either Boolean or neural-like architectures. Section 5 concludes this review and give some considerations on the future of molecular computing for diagnosis.

2. Elementary Modules of Molecular Classifiers

Molecular classifiers are built from elemental functions that can be wired together in complex circuits, whose topology and dynamics are tailored to the computational task. These elemental functions include target recognition and transduction (interface between inputs and circuit), connections (gates, fan-in, etc.), amplification, thresholding, signal restoration, and readout (Figure 2).

2.1. Target Recognition and Signal Transduction

Target recognition and signal transduction are often encoded in the same molecule or molecular complex, used as an interface between the input biomarker(s) and a nucleic acid signal that can be further processed through DNA computation. In the case of nucleic acid detection, signal transduction may be encoded for instance in toehold-mediated strand displacement (TMSD) gates,^[17,27] riboswitches,^[28] DNazymes or ribozymes.^[29] Proteins may also be the input in strand-displacement cascades.^[30–32] For small molecules and other metabolites and proteins, signal transduction may utilize antibody-DNA conjugates,^[33–35] aptamers,^[36] or aptaswitches, but also transcription factors^[37,38] or enzymes. Enzymes themselves may be used as inputs of DNA circuits using modified nucleic acid substrates.^[39]

2.2. Signal Amplification

Often in molecular diagnosis, the concentration of input signals is orders of magnitude below analytical instrumentation detection capacities (aM to pM versus nM to μ M) and calls for molecular-level signal amplification. Many nucleic acid amplification systems have been developed and are detailed in other reviews.^[40–44] Enzyme-free systems, including

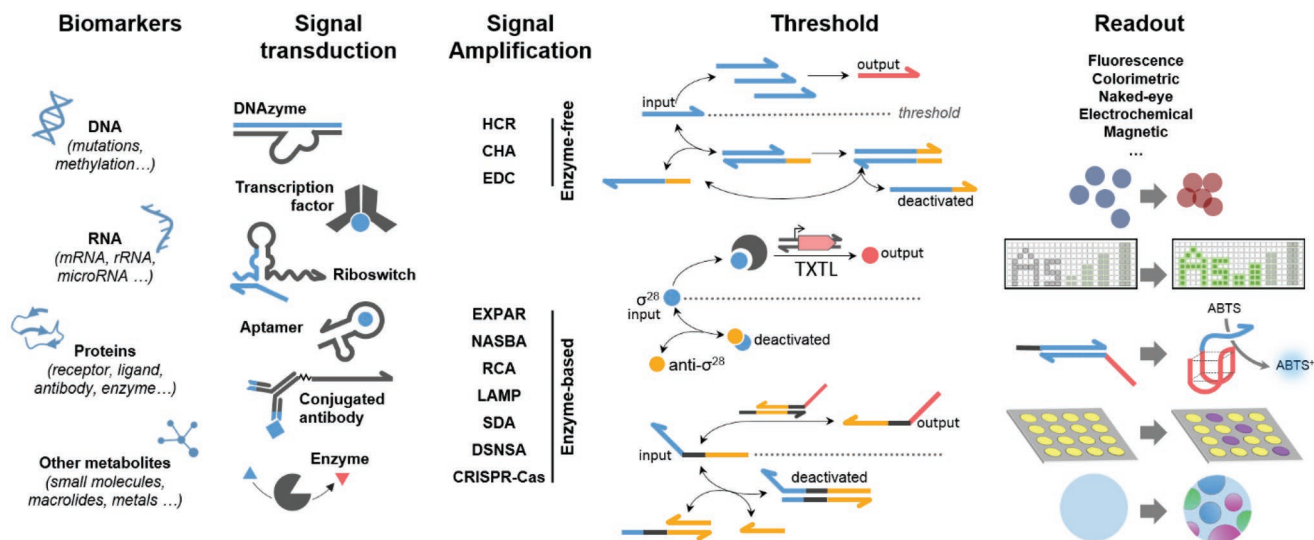


Figure 2. Basic modules of molecular classifiers. The relevant biomarker levels are converted to DNA inputs via signal transduction strategies, and amplified using a variety of enzyme-based or enzyme-free amplification modules. Threshold function may be obtained using template-based enzymatic deactivation of signal strand in PEN-DNA networks,^[19] signal molecule scavenging in genetic circuits,^[20] or irreversible strand exchange in strand-displacement-based circuits.^[21] Examples of readout strategies include gold nanoparticle aggregation (colorimetric),^[22] volume bar-like display (fluorescent),^[23] peroxidase DNAzyme (colorimetric),^[24] paper-based readout (colorimetric),^[25] or liquid–liquid phase separation.^[26]

hybridization-chain reaction (HCR), catalytic hairpin assembly (CHA), and entropy-driven catalysis (EDC) present the advantages of being robust to the media composition and may be used directly in cell extract or plasma.^[15] However, these systems provide a relatively low amplification factor ranging from a dozen to a few thousands fold, insufficient for most biosensing applications. By contrast, enzyme-powered amplification, which leverages the unrivaled catalytic efficiency of protein biocatalysts, enables amplification levels unachievable by enzyme-free circuits. Polymerase chain reaction is by far the most used NA amplification method in clinical settings, owing to its design simplicity, amplification efficiency, and sensitivity. However, the thermocycling process makes this method tricky to adapt to computation (although it was a central tool in the seminal work from Adleman,^[18] and still is in many others^[45,46]). A dozen of isothermal amplification chemistries and probably hundreds of variations and hybrid concepts are now available and have been or could be adapted in DNA computing.^[41,47] These include loop-mediated isothermal amplification (LAMP; using only a DNA polymerase), rolling circle amplification (RCA; DNA polymerase, optionally an endonuclease), exponential amplification reaction (EXPAR; DNA polymerase and nicking enzyme), duplex-specific nuclease signal amplification (DSNSA; duplex specific nuclease), helicase-dependent amplification (helicase, DNA polymerase), strand-displacement amplification (DNA polymerase and nicking enzyme), recombinase polymerase amplification (recombinase, DNA polymerase), nucleic acid sequence based amplification (reverse transcriptase, RNA polymerase, RNase H), among others.

Most, if not all amplification systems are prone to leak, meaning that a positive signal is eventually generated even in absence of the input sequence, albeit at a lower rate. This limitation makes exponential mechanisms particularly unstable and affects their sensitivity because the signal amplified from the unspecific leaky

events is undisguisable from the specific reaction and shields low concentrations of input. More generally, leak limits the sensitivity of a system and also complicates circuit scaling, because errors propagate through the circuits and may affect the output computation.^[48,49] A number of strategies have been implemented to tackle this problem, consisting in reducing the leak, absorbing it or compensating for it. Careful strand design to minimize base breathing and purification of the gates reduces leakage in strand displacement circuits.^[50–52] “Shadow cancellation” compensates for the leak using a complementary circuit with the same leak rate that outputs an anti-leak molecule.^[53]

Spurious reactions also occur in DNA-enzyme circuits, owing to nonspecific interaction of the enzyme with a template or undesired, partial hybridization of primers and/or templates.^[54] A strategy conceptually similar to the shadow cancellation mechanism has been proposed for RCA, where a single-stranded DNA is produced from nonspecific reaction and activates the digestion of the correctly amplified product mediated by Cas12a, therefore compensating for background amplification.^[55] Wan et al. have incorporated in an arsenic cellular biosensor an ingenious nonlinear degradation mechanism that allows to reduce the background signal without affecting the maximum output response.^[23] A similar nonlinear degradation mechanism of the signal oligonucleotide was proposed to absorb the leak in an EXPAR-like system, reducing by 3 orders of magnitude the limit of detection of microRNAs.^[19] Dozens of other strategies or additives are summarized in a comprehensive review by Özyay and McCalla.^[56]

2.3. Thresholding

Threshold mechanisms somehow resemble leak-absorption mechanisms in the sense that they scavenge or deactivate a

certain fraction of the signal molecule before it is transmitted to the next layer of the circuit. Such thresholds are particularly useful in Boolean circuits to make a gate respond to a certain concentration of input, above or below the threshold, rather than just its presence or absence. A threshold function is also essential in analog mode as it controls the bias of neurons in neural computation. In combination with an amplifier, thresholding allows signal restoration, that is, the amplification of an output signal up to a given concentration, a crucial mechanism to avoid background noise as mentioned above, but also to deal with signal loss in circuits of large size.^[57]

2.4. Readout

Readout strategies depend on the type of signal generated (optical such as fluorescence or colorimetry, electrochemical, mass, etc.), the way it is measured (real-time, end-point, digital, single-molecule, etc.) and the user interface or instrument that is used (plate reader, lateral flow assay, mass spectrometry, digital display, naked-eye, etc.). While molecular profiling generally requires high-tech instrumentation such as sequencers, flow cytometers or mass spectrometers, DNA computing has the unique ability to replace this technological overhead by molecular devices, with the promise of lower costs and simpler readouts. Even though a vast majority of nucleic acid circuits are currently tracked using fluorescence, the field may move in the future toward instrument-free or naked-eye readout, well-adapted to point-of-care devices. Particle aggregation has been used for naked-eye detection by a change of color^[58] using notably the plasmonic properties of gold nanoparticles, or their agglutination at the bottom of V-shaped wells (mostly used for rapid immunoassays).^[59] Paper-based assay^[25,60] and paper-based microfluidics^[61,62] are proposed as simple, low-cost, and low-tech interfaces suitable for screening campaigns in disease outbreaks. Shi et al. explore the possibility to grow large, self-assembled and easy-to-detect structures from single-molecule seeds.^[63] The low-cost arsenic sensing platform mentioned above generated a visual bar-like pattern for easy interpretation of the results.^[23] Gong et al. reported an exotic readout based on phase separation of droplets conditionally to a Boolean logic algorithm run on microRNA inputs.^[26] Single-molecule nanopore

sensing^[64,65] may also be of interest for multiplexing the readout of many computations at a time.

Of note, the properties of CRISPR effector proteins regarding both cleavage specificity and collateral activity have been explored for signal transduction and signal amplification, respectively.^[66] This has triggered since the last decade a wave of CRISPR-based diagnosis assays, for example SHERLOCK,^[67] DETECTR,^[68] or LEOPARD,^[69] among others.

3. Molecular Toolboxes

Depending on the underlying (bio)chemistry, the basic functions described above are integrated in various flavors of molecular toolboxes, which allow to rationally build networks with specific topologies and functions.^[70,71] We focus here on those that have been proved to be compatible with molecular classification tasks.

3.1. Genetic Circuits

As touched on above, cells have evolved a modular and sophisticated way of organizing molecular—genetic—circuits (Figure 3A). These can sense and integrate a large panel of internal or external cues and in turn generate a variety of tailored responses (division, motion, death, etc.). A little more than 20 years ago, two seminal works have shown the possibility to wire genetic building blocks to reshape metabolic pathways into artificial networks, in these cases implementing a toggle switch^[72] and an oscillator.^[73] Since then, the field of synthetic biology went on exploring the versatility of rewired genetic networks. These have led to a vast array of demonstrations spanning dynamical behaviors, computation,^[74] pattern formation, and biosensing.^[75,76] However, whole cell biosensors, despite their demonstrated capabilities in solving complex classification problems, suffer from a few pitfalls including biosafety and storage issues, incompatibility with toxic species, or limited diffusion of the inputs through the cell membrane. This have prompted the development of cell-free biosensors to bypass—at least partially—these problems.^[70,77] These are still relatively costly systems that need cautious storage conditions (maintenance of cold-chain or lyophilization). Yet, both

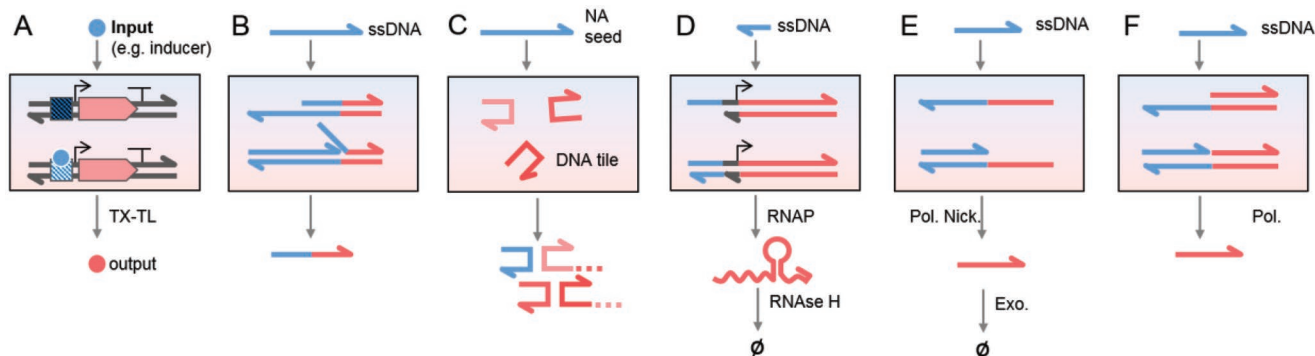


Figure 3. Simplified scheme of some biomolecular toolboxes. A) Genetic circuits. B) Toehold-mediated strand displacement. C) DNA tile self-assembly. D) Genelet. E) PEN-DNA. F) Polymerase-driven strand displacement.

cell-based and cell-free systems can coordinate a generous panel of transducers (membrane receptors, transcription factors, riboswitches, aptaswitches, etc.), amplifiers (polymerases or ribosomes), and output modules (fluorescent or bioluminescent proteins, enzyme-based, or aptamer-based readouts).^[78]

3.2. Enzyme-Free Systems

By harnessing the predictability and simplicity of base pairing between nucleic acid sequences, a number of molecular algorithms were made from hybridization of DNA sequences alone (Figure 3B).^[15] These circuits rely on the control of output release by strand exchange reactions,^[16] following an irreversible (TMSD)^[57,79,80] or reversible process (toehold-mediated strand exchange, based on seesaw gates).^[21,81] Functional nucleic acids such as aptamers^[82,83] or (deoxy) ribozymes^[84] have been used to build Boolean circuits,^[85,86] expanding the range of input molecules compatible with nucleic acid circuits.^[87]

3.3. DNA Tile Assembly

The programmable self-assembly of DNA tiles—that is single or multi-stranded DNA bricks interacting via “sticky ends”—has been largely exploited to build macrostructures from the top-down (Figure 3C).^[88] The interactions between different tiles is predetermined and can be designed so that the growth from a seed structure is controlled by logical rules, eventually forming a given output pattern.^[89] Of note regarding the application of DNA tile assembly for diagnostics, sensitive seed detection has been demonstrated using kinetically controlled single-stranded DNA slat assembly, robust to spontaneous nucleation.^[63] Interestingly, this framework can be combined with strand displacement reactions, opening the way to spatial and temporal control of the growth.^[90]

3.4. Transcriptional Systems

Genelets are transcription-based systems powered by rNTPs polymerization, which rely on the dynamical production and degradation of RNA strands by an RNA polymerase and an RNase, respectively (Figure 3D). Pseudo-genes are encoded in partially double-stranded DNA templates whose expression rate is modulated by the presence of input DNA activators and RNA inhibitors. By connecting different pseudogenes, various transcriptional circuits such as oscillators^[91] or bistable switches^[92] have been implemented. It was even theoretically demonstrated that neural network architectures with adjustable biases and weights can be assembled from genelets, using RNA species as neurons and pseudogenes as synapses.^[93] Such networks are amenable to transcriptomic pattern recognition, or multiomic analysis using functional RNA such as aptamers,^[94] ribozymes, or riboswitches.

Single-stranded hairpin pseudogenes have also been proposed to build transcription-only circuits.^[95] Here, the transcription of a gene can be inhibited by nucleic acids binding in

the loop to induce the dehybridization of the promoter region, similar to the molecular beacon mechanism.^[96]

Another modular design inspired by retrovirus replication, Reverse transcription and TRanscription-based Autonomous Computing System (RTRACS),^[97–99] makes use of 4 enzymes (reverse transcriptase, RNaseH, RNA polymerase, and DNA polymerase) to convert RNA inputs into RNA outputs. It was suggested that RTRACS could be combined with small molecule inputs using aptamers or aptazymes, or protein outputs if run in a cell-free expression system.^[98]

3.5. DNA-Enzyme Systems

In the early 2010s, the PEN-DNA toolbox was proposed to construct dynamical circuits such as oscillatory, multistable, or excitable systems (Figure 3E).^[100–102] PEN-DNA systems rely on DNA strands, termed templates, to encode the connectivity of the network, and on 3 DNA-processing enzymes (Polymerase, Exonuclease, Nickase) to catalyze the production and degradation of short oligonucleotides used as signals. These circuits support linear and nonlinear amplification (e.g., exponential) kinetics,^[101] threshold function,^[19] and can transduce various biomolecular inputs including DNA or RNA strands,^[103] small molecules and enzymes.^[38] Reif et al. proposed a polymerase-only system, where output DNA oligos are released by a DNA polymerase with strand-displacement activity (Figure 3F).^[104,105] This single-enzyme system comes at the cost of increased leak, because output strands are present in the mixture from the very beginning of the reaction, even if initially nonreactive.^[104] Other demonstrations of DNA computations using restriction enzymes have been reported. For instance, Benenson et al. and Zhao et al. leveraged the activity of type IIS restriction enzyme FokI to build a decision tree^[106] and a Boolean algorithm,^[107] respectively.

4. Molecular Computers for Sample Classification

We now describe a selected number of nucleic acid circuits that have shown applicability in sample classification. We can classify these circuits in two categories according to whether they rely on Boolean logic or neural network architectures (Table 1). Boolean circuits refer to a type of digital computation inspired by their electronic counterparts, where—molecular—gates output either 0 or 1 according to combinations of inputs and encoded conditions. Logic circuits operate on digitalized input values, simply presence or absence, or below or above a pre-defined threshold. Binariness makes them robust to noise, but it does not reflect the wide heterogeneity in the concentration values of physiological biomarkers. In addition, the handling of continuous concentrations is cumbersome with digital circuits, and requires a large budget of logic gates ranging from dozens to hundreds of gates.^[108,109] On the contrary, summation is straightforward in analog molecular networks, which directly manipulates concentrations. This feature, together with the possibility to adjust individually each weight makes analog networks more compact (e.g., using less nucleic acid strands) than those using logic gates.

Table 1. Features of some representative molecular logic circuits and neural networks.

Type	Nodes	Strands	Computation	Input conc	Input type	Chemistry	Input number	Reference
Logic	5 gates	33	miRNA pattern classification	>10 nM	miRNA	TMSD	6	[57]
Logic	12 gates	130	4-bit square root circuit	>10 nM	ssDNA	TMSD	4	[21]
Logic	7 gates	ND	2 inputs classification	>μM	Metabolites	TMSD	2	[37]
Logic	10 gates	34	4-bit square root circuit	>nM	ssDNA	Polymerase-driven SD	4	[105]
Logic	1 gate	ND	Glucose and NO _x classifier	>μM	Metabolites	In cellulo genetic circuit	2	[76]
Logic	7 gates	ND	miRNA pattern classification	fM to pM	miRNA	In cellulo genetic circuit	6	[118]
Logic	8 gates	1	RNA transcript classifier	ND	RNA transcripts	In cellulo TMSD + TXTL	12	[119]
Logic	7 gates	24	4-bit square root circuit	>μM	ssDNA	TMSD	4	[113]
Logic	7 gates	100	Even-odd number sorting	100 nM	ssDNA	Tile assembly	6	[116]
Neural	4 neurons	112	4-neuron Hopfield NN	>10 nM	ssDNA	TMSD	4	[81]
Neural	6 neurons	305	Handwritten digit recognition	>10 nM	ssDNA	TMSD	100	[121]
Neural	1 neuron	≈20	Healthy versus lung cancer	>nM	Circular ssDNA	TMSD	4	[46]
Neural	2 neurons	62	Bacterial versus Viral infection	>nM	mRNA	TMSD	7	[124]
Neural	1 neuron	ND	AND, OR, or MAJORITY	ND	DNA genes	TXTL	3	[20]
Neural	1 neuron	ND	4-input linear classifier	>μM	Metabolites	TXTL	4	[127]
Neural	1 neuron	13	Majority voting	>nM	ssDNA	PEN-DNA	10	[125]
Neural/logic	2 neurons, 1 gate	13	Nonlinear space partitioning	pM-nM	miRNA	PEN-DNA	2	[125]
Neural	1 neuron	ND	Spatial pattern recognition	>μM	Metabolites	Multicellular genetic network	2	[126]
Neural	2 neurons	250	Handwritten symbol recognition	>10 nM	ssDNA	TMSD	144	[123]
Neural	8 neurons, 1 gate	512	Handwritten symbol recognition	>10 nM	ssDNA	TMSD	148	[123]

4.1. Boolean Logic and Digital Circuits

Boolean molecular logic circuits are inspired from Boolean logic rules currently used for electronic computing to mimic silicon computer systems.^[110,111] Boolean logic allows digital computations and can be adapted for several purposes such as the classification of binary inputs or the conception of complex algorithmic tasks. A Boolean logic circuit is composed of multiple logic gates, where each takes one or multiple inputs as an entry and accordingly returns a binary output having either 0 or 1 value. The “1” value can be associated as TRUE/YES answer whereas the “0” can be considered as FALSE/NO answer. When inputs are taken in a continuous interval, analog-to-digital conversion is performed by thresholding the input values. Based on this paradigm, many digital computing systems operating at molecular level have been designed in the past few decades.^[17]

Seelig et al. have proposed a Boolean logic approach for miRNA patterns classification.^[57] The largest network includes 5 logic gates, encoded in 12 DNA gates (including threshold, amplification, and translator functions), representing ≈30 strands. The circuit was shown to work properly with dozens of nM of synthetic RNA inputs spiked in total RNA extract. However, its low sensitivity would require input preamplification to be compatible with biologically relevant concentrations. Qian and Winfree used the so-called seesaw gates to scale up this approach, based on a dual-rail logic representation

(Figure 4A).^[21] The circuit, composed of 130 DNA strands for conversion (“seesawing”), thresholding and reporting, encodes 6 Boolean operations for computing the square root of 4-bit input number. Song et al. encoded the same algorithm in a more compact polymerase-driven strand displacement circuit made of 37 DNA strands (Figure 4B).^[105] The DNA polymerase, which considerably speeds up the strand displacement reactions (as compared with TMSD) combined with the lower leak of single-stranded gate (as compared with the seesaw gate motif) allowed to decrease the computation time down to 30 min, an important improvement regarding the applicability of these circuits to the clinical world.

Wang et al. drew inspiration from switching circuits, an architecture originally proposed by Shannon in 1938^[112] and nowadays widely used in fast and high-bandwidth telecommunication. The proposed architecture relies on a single DNA switch motif, activated upon a strand displacement reaction by an input strand, therefore allowing high modularity and easy programmability, with the advantage of having uniformity in the molecular design.^[113] As in electronics, two switches in series encode an AND gate, while they compute a OR gate when wired in parallel. Cascading these two functions, the authors constructed the square root molecular computer for a 4-bit input with 24 strands, that reaches a time to half completion in less than 10 min. Although the computation is fast, the sensitivity (dozens of nM of inputs) is currently limited by the lack of an efficient catalytic mechanism.

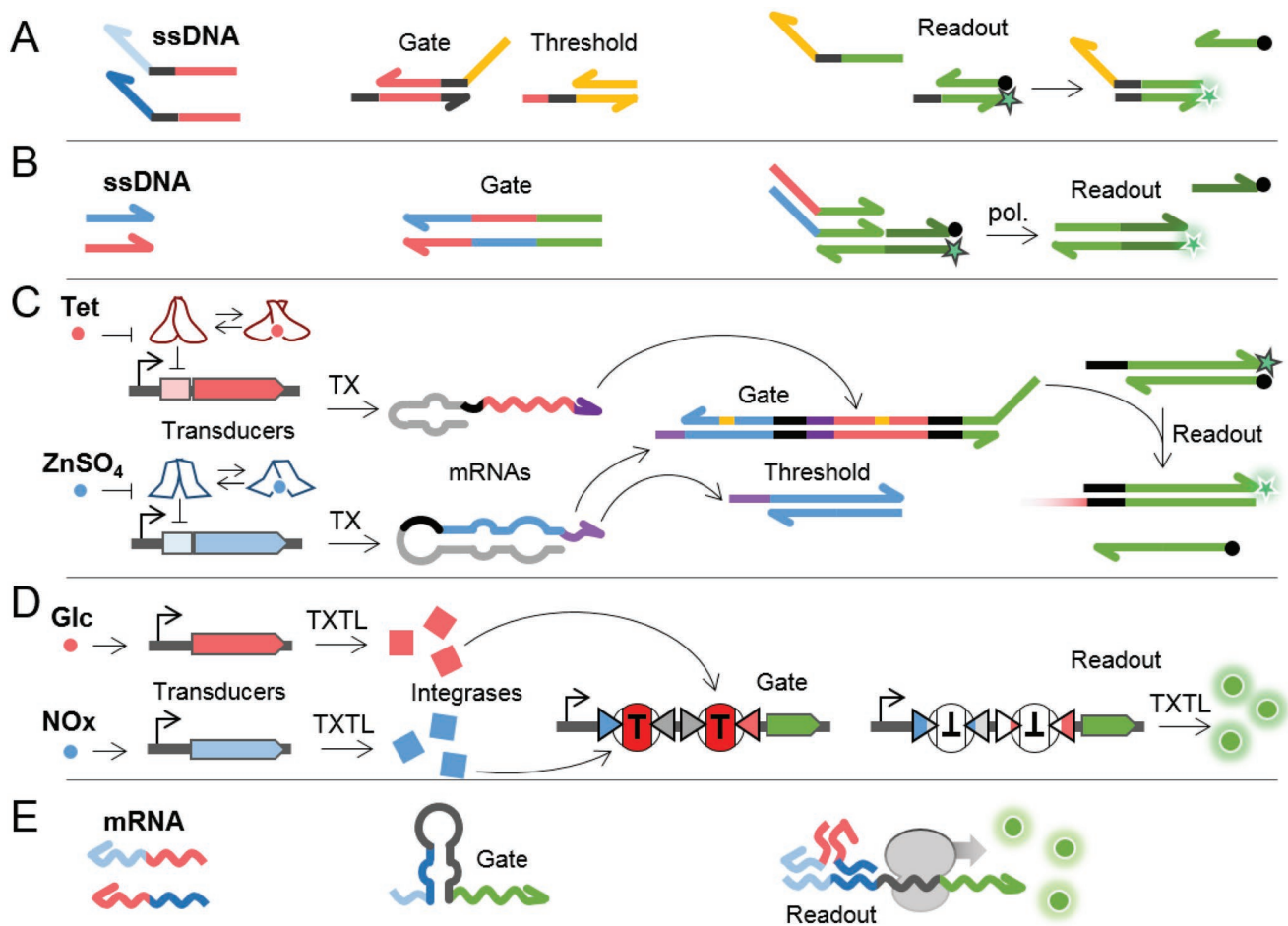


Figure 4. Molecular 'AND' Logic Gates. A) See-saw gates digital computing.^[21] B) Polymerase-driven AND gate.^[105] C) ROSALIND coupled to toehold-mediated strand displacement reaction.^[37] D) Bacteria engineered to perform logic computation using Integrase-controlled transcription.^[76] E) Ribocomputing device.^[119]

Takeuchi et al. developed a miRNA classifier based on a 5-input DNA AND gate and a nanopore readout.^[114] The gate simply consists of the 5 microRNA binding sites and a 5' hairpin structure to guide the direction of the gate inside the nanopore (α -hemolysin) from the 3' end. The binding of all 5 miRNAs to the gate leads to a distinct nanopore signal subsequently decoded in silico. Although the computation is not performed by a Boolean DNA circuit, this work demonstrates the possibility to use nanopore readout to detect subfemtomolar concentration of nucleic acid targets.

Using allosteric transcription factors (aTF), the Lucks' group developed a range of biosensors for small molecules and metals, termed ROSALIND (RNA Output Sensors Activated by Ligand INDuction, Figure 4C).^[115] In this case, DNA templates are engineered to produce a fluorescent RNA aptamer upon in vitro transcription, which is regulated by the binding/unbinding of aTF. More specifically, in absence of input, the aTF is bound to the template, preventing the RNA polymerase from escaping the upstream promoter region. When the input metabolite is present, it binds to its cognate aTF, pulling it out of its binding site and eventually removing the inhibition exerted on the operator for the aptamer transcription. In

a subsequent work, the same group interfaced this sensing layer with strand displacement computation.^[37] To that goal, the aptamer output is replaced by an invader RNA that initiates the strand displacement reaction on AND or OR DNA gates. A NOT gate was also constructed by producing, upon the presence of the input, an output RNA that sequesters a constitutively transcribed invader strand (with no aTF binding site), therefore preventing the gate to be activated. Combining these elementary logic gates, they built more complex systems including NOR, IMPLY, NAND, and NIMPLY gates, taking tetracycline and zinc ions as inputs.

Woods et al. designed a Boolean framework based on DNA tile self-assembly that can run a wide variety of 6-bit algorithms.^[116] Each algorithm is encoded in a specific set of tiles selected from an ingeniously designed collection of 355 single-stranded DNA tiles. Starting from a DNA origami seed that encodes the 6-bit input, the tiles self-assemble iteratively to form an extended nanotube, and the result of the computation is read by atomic force microscopy after labeling biotinylated strands with streptavidin.

Courbet et al. have engineered bacteria to perform logic computation.^[117] The design is based on the inversion or deletion

of gene terminator or promoter by integrase input(s) to control the output transcription rate of a gene gate. By playing with the gene architecture, location and orientation of the recombination site, the authors demonstrated the high programmability of the approach for designing various logical operations. By incorporating sensitive transcriptional promoters upstream to the integrase gene, the expression of the integrase becomes conditional to the promoter inputs, allowing to interface the genetic circuit with other molecules (nitrogen oxides, glucose, Figure 4D).^[76] They demonstrated this “bactosensor” to be resistant to the presence of a high percentage of serum or urine concentration in the culture medium. This aspect is non negligible since the toxicity of samples for living cells remains one of the main drawbacks of whole-cell devices. Because the output depends directly on modifications at the genomic level, these bacterial sensors can store the result of the computation for long-term, the latter being recovered by re-growing the cells or by analyzing their DNA by PCR or sequencing. This unique feature may provide a useful approach for results preservation and/or to delay the readout in remote clinical settings.

In 2011, Benenson’s team built a cell-type classifier based on transcriptional/post-transcriptional synthetic regulatory circuits that senses the expression level of endogenous miRNAs.^[118] Interestingly, they substituted the fluorescence output by the production of a protein (hBax) that induces apoptosis in HeLa cells which expresses the correct miRNA pattern, without affecting surrounding cells. This proof-of-concept represents a “tour-de-force” in terms of optimization of complex genetic circuits running inside mammalian cells from endogenous inputs. Green et al. reported another cellular computer based on ribocomputing, where engineered riboswitches respond to RNA transcripts expressed from transformed plasmids (Figure 4E).^[119] They ultimately constructed a 12-input Boolean classifier made of AND, OR and NOT gates. However, this strategy requires the transfection/transformation of synthetic circuits for internal sensing and actuation (green fluorescent protein (GFP) production^[119] or apoptosis triggering^[118]) on targeted cells, which seems not directly applicable for therapeutic purposes.

4.2. Neural Networks

Molecular neural networks (MNN) have an architecture that abstracts their biological counterparts (BNN) with a mathematical formulation borrowed from the field of artificial intelligence (ANN, Figure 5). They are built by connecting individual molecular “neurons,” each performing a weighted sum on an upstream input collection and returning an output signal according to the embedded—nonlinear—activation function;

very much like a biological neuron receives input current from dendrites, processes these signals in the cell body and outputs to the next neuron(s) via its axon. Contrary to logic circuits, which work on digital inputs, neural networks are governed by ordinary differential equations. As such they are related to chemical reaction networks, which, in the deterministic limit, are also naturally described by ordinary differential equations. In this analogy, continuous values of the variables in the neural network are represented by continuous ranges of concentrations in the chemical mixture. Although neural networks were proposed early in the field of DNA computing,^[120] experimental demonstrations of a DNA-encoded chemical “neuron” appeared only a decade ago (cf. below). The most elementary neural network is a perceptron composed of a unique neuron that switches ON if the weighted sum of its inputs exceeds a given threshold and stays OFF otherwise. A perceptron can classify linearly separable data, that is, sets of data that can be separated by a line or a plane in two or three dimensions, respectively—or a hyperplane for patterns of higher dimensions. Multiple perceptrons can be cascaded to assemble multi-layered architectures capable of sophisticated classification algorithms.

In 2011, Qian et al. experimentally demonstrated a neural network architecture based on reversible strand displacement gates (see-saw gates).^[81] In this initial demonstration, the authors achieved a single layer NN of 4 fully connected neurons mimicking a Hopfield neural network, capable of recognizing 4 different patterns. Each pattern corresponds to the answers to 4 questions for playing the game “guess who” with 4 scientists. More recently, Cherry and Qian successfully implemented a neural network capable of recognizing handwritten digits (100-bit patterns, Figure 6A).^[121] This more compact architecture utilizes annihilation gates allowing simple and powerful winner-take-all computation.^[122]

Pei et al. also proposed a compact approach based on the sparse topology of convolutional neural networks.^[123] In this work, the classification is performed based on the transformation of the inputs (binary 2D matrices) by a convolutional kernel, followed by summation and subtraction to finally activate the output neuron(s) (Figure 6B). All inputs assigned to the same region of the kernel share the same weight, each represented by a unique weight tuning molecule, therefore reducing the size of the DNA network. They demonstrated the recognition of 8 handwritten symbols encoded in a 12×12 binary input matrix. Adding an upstream coarse logical circuit, DNA-based convolution MNN could recognize 32 symbols from 4 categories (8 symbols borrowed from the Arabic numbers, Chinese oracles, English alphabets, and Greek alphabets): in this implementation, 4 “pixels” are used to tag the symbols, each tag corresponding to the symbol category. A set of DNA gates encoding a Boolean algorithm first reacts with the tag strands, outputting

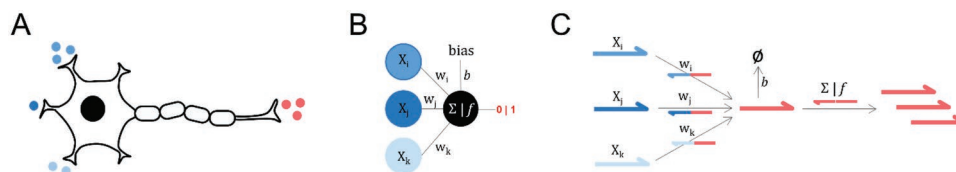


Figure 5. A) Biological, B) Artificial, and C) Molecular neurons used in BNN, ANN, and MNN, respectively.

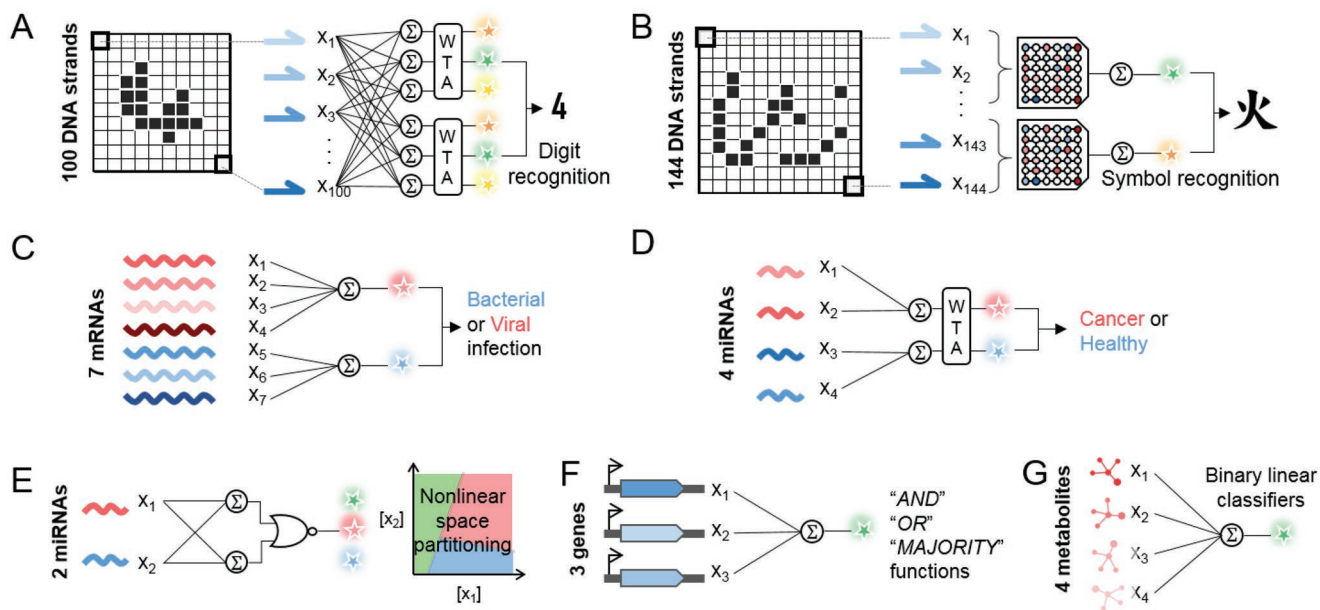


Figure 6. Experimentally demonstrated neuromorphic molecular computing. A) Winner-take-all (WTA) computation designed for handwritten digit recognition.^[121] B) Molecular convolutional network for the recognition of symbols.^[123] C) Classifier of bacterial versus viral infection from host RNA transcripts.^[124] D) Neural sensor of lung cancer patient from 4 input circulating miRNAs.^[46] E) Neural-logic hybrid network capable of nonlinear space partitioning.^[125] F) Riboregulator-based perceptron for the classification of gene inputs.^[20] G) Metabolic perceptron from enzymatic conversion and benzoate biosensor.^[127]

both a fluorescence signal corresponding to each category and a set of weight tuning molecules using a fan-out mechanism. The weight tuning molecules are captured on magnetic beads and released to activate the downstream MNN for completing the pattern classification, that is, choosing among the 8 symbols of the corresponding category. As a result, each of the 32 symbols is recognized from a combination of the 4 digital circuit outputs and 8 MNN outputs.

These works have shown the potential of DNA strand displacement cascade for classification tasks. However, because they use high input concentrations (typically dozens of nanomolars), it remains unclear whether these circuits could become compatible with typical nucleic biomarkers concentration level.

This sensitivity issue can be partially addressed using pre-amplification of the inputs. In 2015, Lopez et al. reported a DNA classifier that distinguishes bacterial versus viral respiratory infection (Figure 6C).^[124] The classifier is also based on strand displacement reactions and therefore requires in vitro transcription of cDNA mixtures to raise the input concentration from pM to nM. The classifier consists of two disjointed single layer perceptrons that are connected to bacterial or virus-related sequences and report a ROX or FAM signal if the weighted sum on each neuron exceeds the defined threshold. The absolute weights values are encoded in the numbers of probe binding sites selected for each transcript sequence and their sign depends on the output strand of the following strand displacement reaction for summation (one sequence for positive weights, one for negative weights). At the end of the computation, the measured ROX signal is subtracted from the FAM signal, *ex-moleculo*, allowing output normalization and accurate discrimination of bacterial and viral infections.

Zhang et al. elaborated a complex sample pre-processing procedure to convert low microRNA concentrations (fM to pM range) into higher concentration (>nM) of DNA strands using an asymmetric PCR (termed LATE PCR for Linear After Exponential PCR).^[46] The complete workflow involves the reverse transcription of the miRNAs, LATE PCR, followed by the circularization of the product into “Loop DNA” which are finally purified by gel electrophoresis. These converted Loop DNA are then fed as inputs to a 4-miRNA linear classifier, which is able to properly distinguish between clinical serum samples from healthy people and lung cancer patients (Figure 6D). For each wire of the molecular classifier (i.e., connection between a Loop DNA sequence and the integrator), the absolute weight value corresponds to the number of probe binding sites introduced into the Loop DNA during the PCR step, and its sign depends on the output sequence. As in Cherry and Qian, a final “winner-take-all” annihilation reaction between two outputs, one from negative weights, the other from positive weights, returns the final digital output, ‘healthy’ or ‘diseased’. Although the multistep sample treatment remains tedious, this appears to be, to our knowledge, the first time a MNN is used on real patient samples.

The PEN-DNA toolbox also possesses the necessary building blocks to assemble a perceptron, namely input transduction, weighting, summation, and thresholded nonlinear activation function.^[125] Weights are tuned by introducing decoy unproductive templates that compete for the inputs with cognate activator-producing templates: As a result, the ratio of the two templates defines the weight of a given input. Negative weights can be implemented by using templates that produce anti-activator strands, and thus does not require additional annihilation gates as used in enzyme-free networks. The bias term can be easily

tuned by varying the concentration of the template that thresholds the amplification. Okumura et al. reported a majority voting linear classifier working on 10-bit inputs, demonstrating robust weight equalization for all 10 inputs and sharp decision margin. The authors also implemented the first 2-layer neural network and a neural-logic hybrid network capable of nonlinear classification on picomolar concentrations of microRNA inputs without preamplification (Figure 6E). All these experiments have been performed on synthetic samples, and their application on actual patient samples requires further validation.

Van der Linden et al. employed a cell-free expression system (in vitro transcription translation) to run a perceptron on genetic inputs (Figure 6F).^[20] The RNA transcripts resulting from gene expression are interfaced with protein production summation layer via riboswitches, whose respective concentration encodes the weight on the different inputs. Upon activation by their cognate transcript, all riboswitches produce the sigma factor σ^{28} , which in turn activates the expression of GFP under transcriptional control of the P28a promoter. The activation threshold can be ingeniously controlled by adding anti- σ^{28} to scavenge a fraction of the output σ^{28} (Figure 2). On a three-input pattern, the perceptron with adjustable bias can compute a OR (no threshold), MAJORITY (moderate threshold), and AND (high threshold) functions. The authors notified that such classifiers may be tricky to tune manually for any input set due to the inherent variability at the transcriptional and translational level, as well as on the regulatory elements. This problem may however drive the development of strategies to (semi-) autonomously train a MNN by continuously adjusting the weights of the network until the desired outcome is obtained. In addition, and despite the additional complexity as compared to enzyme-free systems, TXTL-based computation possesses the unique advantage of being naturally compatible with a large array of inputs (e.g., small molecules) and outputs (e.g., luminescent proteins).

Li et al. implemented a perceptron-like genetic network that allows pattern recognition in bacterial consortia composed of sender and receiver cells.^[126] In a simplified representation, sender cells are induced by an input molecule (acyl-homoserine lactone, OC6), which triggers the production of a quorum sensing molecule (acyl-homoserine lactone OCH14) via the activation of the P_{Lux} promoter through the binding of LuxR transcription factor. OCH14 can diffuse in the receiver population equipped with a quorum sensing genetic network that perform the weighted sum of OCH14 from the different sender populations and activate the production of a fluorescent protein if it exceeds a given threshold. Interestingly, they demonstrated the tuning of the weight using—mutated— P_{Lux} promoters of different strengths for each sender population. Negative weights were also implemented by placing the P_{Lux} promoter downstream a constitutively active promoter, consequently blocking the RNA polymerase when bound to the OC6-LuxR transcription factor. Using this population computing method in combination with an in silico gradient descent training algorithm, they showed accurate classification of 3×3 -bit patterns.

While all above-mentioned neural networks are based on DNA actuation, either being in vitro genetic circuits, DNA-enzyme or strand-displacement-based systems, it may be interesting as well to leverage metabolic circuits that naturally

integrate and return small molecules as inputs and outputs. Pandi et al. approximate a single-layer perceptron using such strategy, where inputs (hippurate, cocaine, benzamide, and biphenyl-2,3-diol) are all converted into benzoate in a cell-free transcription translation media (Figure 6G).^[127] The summation and actuation layer is encoded in a benzoate sensor, that is, a GFP expressing plasmid under the control of benzoate transcriptional activator. They built a computational model that allows to tune each weight independently by controlling the concentration of the corresponding transducing enzymes (more precisely, their gene). Although limited to positive weight only and lacking control of the activation function (unique threshold), this work sets up the foundations for more complex, possibly multilayer metabolic perceptrons.

5. Conclusion and Perspectives

The integration of multiple molecular clues to provide more precise and personalized diagnostics is a strong trend, as emphasized in this review. However, to collect the necessary multidimensional data, current approaches rely on cumbersome, often costly procedures involving high-tech equipment, skilled staff and the final classification task is typically performed in silico or according to expert knowledge. Integrated molecular computers may bring simple though robust and informative assays to the patient bedside. However, this field still needs to address a couple of challenges before it can deliver generalizable and scalable tools for molecular medicine. One of them is the relatively low sensitivity of most molecular classifiers, which hardly meets the requirements of clinical settings. Target pre-amplification may solve this issue,^[46,124] but it adds extra steps that conflict with the benefits of molecular computers in terms of cost and speed. Enzyme-driven systems have shown amplification efficacy rivaling with PCR-based methods and could either be used as sensitive computational frameworks or combined with enzyme-free chemical reaction networks.^[128]

Another asset of molecular computation is that it can be directly executed inside the body or cells.^[118,129,130] In principle, this theragnostic approach (combining diagnosis and therapeutics) could control the spatio-temporal release of drugs, for instance to selectively kill cancer cells or pathogens, or to regulate imbalanced molecular pathways in neurodegenerative diseases. Yet this view is still in its infancy,^[75] and we still need to learn how to make robust molecular programs which operate in an unknown, fluctuating, and possibly hostile environment.

Boolean logic circuits have been widely explored in the field of DNA nanotechnology.^[17] Despite the simple abstraction scheme borrowed from silicon-based computational paradigm, Boolean functions can be layered to solve complex classification problems. As they work on binary input combinations however, they are not naturally adapted to many diagnostic tasks, where input values are continuous biomarker concentrations and may vary substantially between individuals. Threshold gates and fuzzy logic architectures^[131] somehow mitigate the problem of digitalization, but it remains that building a desired classification algorithm is challenging. By contrast, MNN accepting variable weights and thresholds can be (re)programmed to

fit a given classification problem without changing a set of strands.^[121] These same features authorize compact—and therefore experimentally realistic—networks that compute complex linear and nonlinear data classification on large sets of input biomarkers.

At present, molecularly encoded classifiers are trained in silico before being hard-coded in a chemical reaction network.^[46,118,121,124] Making neural networks that learn in vitro is still a formidable and open-ended task, and it is uncertain if what we learned from in silico training will be directly applicable.^[132] The standard algorithm for training ANN, stochastic gradient descent, uses a myriad of complex mathematical operations such as a differentiation for computing the gradient of the backpropagation step, and it is undefined if those operations can be easily transposed in a chemical setting (even the biological relevance of backpropagation has been debated^[133]). Lakin, Stefanovic et al. have proposed a theoretical biochemical adaptive circuit based on self-inhibiting DNAzyme that can learn linear functions, that is, the conversion of an input species to output species at a given rate.^[134] The learning function relies on a feedback circuit that updates the weights by controlling the activation level of the DNAzyme actuators, based on the difference between the output value (concentration of the output strand) and the expected value (concentration of a DNA species that react in an annihilation reaction with the output strand). The same group also suggested to use buffered—inactive—DNA gates activated by a feedback loop to similarly learn the weights in DNA strand displacement circuits.^[135]

A whole new set of ideas and tools will likely complement this approach to train MNN—directed evolution and the massive parallelism of microfluidics are probably among the most promising candidates. In this scheme, millions of networks with varying topologies and weights are encapsulated in microcompartments (e.g., droplets or chambers), and tested against a range of inputs. After each round, the best instantiations are kept and iterated—possibly with variations, and the round of testing performed again until the system reaches the desired objective.

The question of molecular learning will also force us to design networks that learn without strong supervision, probably by taking inspiration from the brain rather than artificial neural networks. The brain is not thought to use backpropagation to learn.^[133,136] Rather, it encodes computations and information in the temporal domain, by sending spikes of electrical activity between neurons. This sparse temporal encoding is the basis of Hebbian learning, which learns connections between neurons by amplifying weights between the ones that fire together. In the future, MNN may exploit this natural learning to train networks in a massively parallel way, and without relying on assistance from an external operator. Temporal information processing has been investigated for instance using DNA strand displacement^[137] or genetic circuits logic gates^[138] and may also make use of spiking biochemical circuits.^[139]

Another option is to quantitatively model the experimental system, measuring the activation function and the chemical weight modulation, to then perform the training in silico using standard optimization algorithms.^[126] However, this approach may not be possible for complex biochemical systems, for which obtaining a comprehensive modeling is challenging. It

was also proposed to use translational regulation as a mechanism to adjust in real-time the weights during learning.^[20]

Reservoir computing is an approach to reduce the burden of training.^[140–143] In this framework, a recurrent network with many interacting and nonlinear units is perturbed with inputs, provoking a complex response of the network. This perturbation essentially maps the set of inputs to a high-dimensional state and can be seen as a computation. But contrary to common neural networks, no effort is made to adjust the weights between the units. The only training is to learn a linear function that maps back the high-dimensional response to the desired output. While reservoir computing has not really won against the brute force approach of deep learning in the in silico world, it may be a natural contender in the molecular world. Chemical systems genuinely have complex responses, with many layers of intended and unintended feedbacks between the chemical species. In addition, finely tuning chemical interactions may be cumbersome, depending on the underlying chemistry. Since reservoir computing does not need to tune internal weights, it may be well adapted to molecular learning.

In addition, reservoir computing may open up a new dimension to pattern recognition: the temporal domain. Indeed, most molecular diagnosis is static, in the sense that it only looks at a profile of concentrations at a fixed point in time. Yet in the body, concentrations evolve dynamically, and many species are expressed periodically or in bursts. Reservoir computers can naturally handle time-varying inputs, and for instance detect the recurrence of a pattern in the temporal signal of a chemical species. Alternatively, implementing a resetting function may be of interest for the quasi-continuous monitoring of disease progression, relevant for implantable devices that monitor at specific time points the composition of biofluids. This may be performed for instance by limiting a chemical fuel so that the computation consumes it all and the system relaxes to the initial state after readout. A fuel shot is reintroduced each time a new sampling is performed, making it possible to reuse a molecular computer many times. Computing on surfaces is also compatible with resetting as the support can be restored to its initial step by simply washing and flowing back the necessary components for rebooting the molecular program. Computing on surface has other advantages as it may limit errors, avoid cross-reactivity and accelerate reaction rates. As reactive elements are spatially separated on distinct locations, it allows to (re)use the same components (e.g., molecular gates or neurons) on different nodes of the network, – mitigating the tediousness of sequence design faced when scaling up –, and enables to run multiple algorithms simultaneously.^[144,145]

Acknowledgements

This work was supported by the European Research Council under the framework program H2020 for research and innovation (grant project MoP-MiP, No. 949493).

Conflict of Interest

The authors declare no conflict of interest.

Keywords

Boolean circuits, DNA circuits, molecular classification, molecular programming, neural networks

Received: July 19, 2022

Revised: September 28, 2022

Published online:

- [1] Z. Zhao, K. Wang, F. Wu, W. Wang, K. Zhang, H. Hu, Y. Liu, T. Jiang, *Cell Death Dis.* **2018**, *9*, 475.
- [2] C. Liu, J. Zhao, F. Tian, L. Cai, W. Zhang, Q. Feng, J. Chang, F. Wan, Y. Yang, B. Dai, Y. Cong, B. Ding, J. Sun, W. Tan, *Nat. Biomed. Eng.* **2019**, *3*, 183.
- [3] S. Quinlan, A. Kenny, M. Medina, T. Engel, E. M. Jimenez-Mateos, in *International Review of Cell and Molecular Biology* (Eds: L. Galluzzi, I. Vitale), Academic Press, Cambridge **2017**, p. 309.
- [4] V. Tiwari, D. M. Wilson, *Am. J. Hum. Genet.* **2019**, *105*, 237.
- [5] D. A. Giljohann, C. A. Mirkin, *Nature* **2009**, *462*, 461.
- [6] J. D. Cohen, L. Li, Y. Wang, C. Thoburn, B. Afsari, L. Danilova, C. Douville, A. A. Javed, F. Wong, A. Mattox, R. H. Hruban, C. L. Wolfgang, M. G. Goggins, M. D. Molin, T.-L. Wang, R. Roden, A. P. Klein, J. Ptak, L. Dobbyn, J. Schaefer, N. Silliman, M. Popoli, J. T. Vogelstein, J. D. Browne, R. E. Schoen, R. E. Brand, J. Tie, P. Gibbs, H.-L. Wong, A. S. Mansfield, et al., *Science* **2018**, *359*, 926.
- [7] G. A. Calin, C. M. Croce, *Nat. Rev. Cancer* **2006**, *6*, 857.
- [8] T. Muinao, H. P. D. Boruah, M. Pal, *Heliyon* **2019**, *5*, e02826.
- [9] E. L. Tsalik, R. Henao, M. Nichols, T. Burke, E. R. Ko, M. T. McClain, L. L. Hudson, A. Mazur, D. H. Freeman, T. Veldman, R. J. Langley, E. B. Quackenbush, S. W. Glickman, C. B. Cairns, A. K. Jaehne, E. P. Rivers, R. M. Otero, A. K. Zaas, S. F. Kingsmore, J. Lucas, V. G. Fowler, L. Carin, G. S. Ginsburg, C. W. Woods, *Sci. Transl. Med.* **2016**, *8*, 322ra11.
- [10] O. Ramilo, W. Allman, W. Chung, A. Mejias, M. Ardura, C. Glaser, K. M. Wittkowski, B. Piqueras, J. Banchereau, A. K. Palucka, D. Chaussabel, *Blood* **2006**, *109*, 2066.
- [11] G. Pezzotti, F. Boschetto, E. Ohgitani, Y. Fujita, M. Shin-Ya, T. Adachi, T. Yamamoto, N. Kanamura, E. Marin, W. Zhu, I. Nishimura, O. Mazda, *Adv. Sci.* **2022**, *9*, 2103287.
- [12] Y. E. Antebi, J. M. Linton, H. Klumpe, B. Bintu, M. Gong, C. Su, R. McCardell, M. B. Elowitz, *Cell* **2017**, *170*, 1184.
- [13] S. Gordon, *Cell* **2002**, *111*, 927.
- [14] C. Bushdid, M. O. Magnasco, L. B. Vosshall, A. Keller, *Science* **2014**, *343*, 1370.
- [15] C. Jung, A. D. Ellington, *Acc. Chem. Res.* **2014**, *47*, 1825.
- [16] D. Y. Zhang, G. Seelig, *Nat. Chem.* **2011**, *3*, 103.
- [17] S. Zhao, L. Yu, S. Yang, X. Tang, K. Chang, M. Chen, *Nanoscale Horiz.* **2021**, *6*, 298.
- [18] L. M. Adleman, *Science* **1994**, *266*, 1021.
- [19] K. Montagne, G. Gines, T. Fujii, Y. Rondelez, *Nat. Commun.* **2016**, *7*, 13474.
- [20] A. J. van der Linden, P. A. Pieters, M. W. Bartelds, B. L. Nathalia, P. Yin, W. T. S. Huck, J. Kim, T. F. A. de Greef, *ACS Synth. Biol.* **2022**, *11*, 1510.
- [21] L. Qian, E. Winfree, *Science* **2011**, *332*, 1196.
- [22] I.-H. Lee, K.-A. Yang, J.-H. Lee, J.-Y. Park, Y. G. Chai, J.-H. Lee, B.-T. Zhang, *Nanotechnology* **2008**, *19*, 395103.
- [23] X. Wan, F. Volpetti, E. Petrova, C. French, S. J. Maerkl, B. Wang, *Nat. Chem. Biol.* **2019**, *15*, 540.
- [24] G. Eckhoff, V. Codrea, A. D. Ellington, X. Chen, *J. Syst. Chem.* **2010**, *1*, 13.
- [25] K. Pardee, A. A. Green, T. Ferrante, D. E. Cameron, A. DaleyKeyser, P. Yin, J. J. Collins, *Cell* **2014**, *159*, 940.
- [26] J. Gong, N. Tsumura, Y. Sato, M. Takinoue, *Adv. Funct. Mater.* **2022**, *32*, 2202322.
- [27] F. C. Simmel, B. Yurke, H. R. Singh, *Chem. Rev.* **2019**, *119*, 6326.
- [28] K. Pardee, A. A. Green, M. K. Takahashi, D. Braff, G. Lambert, J. W. Lee, T. Ferrante, D. Ma, N. Donghia, M. Fan, N. M. Daringer, I. Bosch, D. M. Dudley, D. H. O'Connor, L. Gehrke, J. J. Collins, *Cell* **2016**, *165*, 1255.
- [29] I. Willner, B. Shlyahovskiy, M. Zayats, B. Willner, *Chem. Soc. Rev.* **2008**, *37*, 1153.
- [30] A. Bertucci, A. Porchetta, E. Del Grosso, T. Patiño, A. Idili, F. Ricci, *Angew. Chem., Int. Ed.* **2020**, *59*, 20577.
- [31] A. P. Diaz, S. Bracaglia, S. Ranallo, T. Patino, A. Porchetta, F. Ricci, *J. Am. Chem. Soc.* **2022**, *144*, 5820.
- [32] W. Engelen, L. H. H. Meijer, B. Somers, T. F. A. de Greef, M. Merckx, *Nat. Commun.* **2017**, *8*, 14473.
- [33] T. Sano, C. L. Smith, C. R. Cantor, *Science* **1992**, *258*, 120.
- [34] B. Schweitzer, S. Wiltshire, J. Lambert, S. O'Malley, K. Kukanskis, Z. Zhu, S. F. Kingsmore, P. M. Lizardi, D. C. Ward, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 10113.
- [35] D. Zhao, Y. Kong, S. Zhao, H. Xing, *Top. Curr. Chem.* **2020**, *378*, 41.
- [36] R. M. Dirks, N. A. Pierce, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 15275.
- [37] J. K. Jung, C. M. Archuleta, K. K. Alam, J. B. Lucks, *Nat. Chem. Biol.* **2022**, *18*, 385.
- [38] V. A. Shenshin, C. Lescanne, G. Gines, Y. Rondelez, *Nucleic Acids Res.* **2021**, *49*, 7765.
- [39] N. Farag, R. Mattossovich, R. Merlo, Ł. Nierzwicki, G. Palermo, A. Porchetta, G. Perugino, F. Ricci, *Angew. Chem.* **2021**, *133*, 7359.
- [40] L. M. Zanolli, G. Spoto, *Biosensors* **2012**, *3*, 18.
- [41] Y. Zhao, F. Chen, Q. Li, L. Wang, C. Fan, *Chem. Rev.* **2015**, *115*, 12491.
- [42] G. Gines, R. Menezes, W. Xiao, Y. Rondelez, V. Taly, *Mol. Aspects Med.* **2019**, *72*, 100832.
- [43] R. Duan, X. Lou, F. Xia, *Chem. Soc. Rev.* **2016**, *45*, 1738.
- [44] M. De Felice, M. De Falco, D. Zappi, A. Antonacci, V. Scognamiglio, *Biosens. Bioelectron.* **2022**, *205*, 114101.
- [45] C. M. Ackerman, C. Myhrvold, S. G. Thakku, C. A. Freije, H. C. Metsky, D. K. Yang, S. H. Ye, C. K. Boehm, T.-S. F. Kosoko-Thoroddsen, J. Kehe, T. G. Nguyen, A. Carter, A. Kulesa, J. R. Barnes, V. G. Dugan, D. T. Hung, P. C. Blainey, P. C. Sabeti, *Nature* **2020**, *582*, 277.
- [46] C. Zhang, Y. Zhao, X. Xu, R. Xu, H. Li, X. Teng, Y. Du, Y. Miao, H. Lin, D. Han, *Nat. Nanotechnol.* **2020**, *15*, 709.
- [47] P. Craw, W. Balachandran, *Lab Chip* **2012**, *12*, 2469.
- [48] I. Zarubiieva, C. Spaccasassi, V. Kulkarni, A. Phillips, *ACS Synth. Biol.* **2022**, *11*, 1931.
- [49] X. Chen, N. Briggs, J. R. McLain, A. D. Ellington, *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 5386.
- [50] B. Wang, C. Thachuk, A. D. Ellington, E. Winfree, D. Soloveichik, *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 12182.
- [51] C. Thachuk, E. Winfree, D. Soloveichik, in *DNA Computing and Molecular Programming* (Eds: A. Phillips, P. Yin), Springer International Publishing, New York City **2015**, p. 133.
- [52] E. Kim, J. Xu, J. Kim, H. Chun, *Analyst* **2022**, *147*, 1899.
- [53] T. Song, N. Gopalkrishnan, A. Eshra, S. Garg, R. Mokhtar, H. Bui, H. Chandran, J. Reif, *ACS Nano* **2018**, *11*, 11689.
- [54] E. Tan, B. Erwin, S. Dames, T. Ferguson, M. Buechel, B. Irvine, K. Voelkerding, A. Niemi, *Biochemistry* **2008**, *47*, 9987.
- [55] B. Tian, G. A. S. Minero, J. Fock, M. Dufva, M. F. Hansen, *Nucleic Acids Res.* **2020**, *48*, e30.
- [56] B. Özyay, S. E. McCalla, *Sens. Actuators Rep.* **2021**, *3*, 100033.
- [57] G. Seelig, D. Soloveichik, D. Y. Zhang, E. Winfree, *Science* **2006**, *314*, 1585.
- [58] M. H. Jazayeri, T. Aghaie, A. Avan, A. Vatankhah, M. R. S. Ghaffari, *Sens. Biosensing Res.* **2018**, *20*, 1.

- [59] A. Poschenrieder, M. Thaler, R. Junker, P. B. Lippa, *Anal. Bioanal. Chem.* **2019**, *411*, 7607.
- [60] P. B. Allen, S. A. Arshad, B. Li, X. Chen, A. D. Ellington, *Lab Chip* **2012**, *12*, 2951.
- [61] L. Magro, C. Escadafal, P. Garneret, B. Jacquelin, A. Kwasiborski, J.-C. Manuguerra, F. Monti, A. Sakuntabhai, J. Vanhomwegen, P. Lafaye, P. Tabeling, *Lab Chip* **2017**, *17*, 2347.
- [62] P. Garneret, E. Coz, E. Martin, J.-C. Manuguerra, E. Brient-Litzler, V. Enouf, D. F. G. Obando, J.-C. Olivo-Marin, F. Monti, S. van der Werf, J. Vanhomwegen, P. Tabeling, *PLoS One* **2021**, *16*, e0243712.
- [63] D. Minev, C. M. Wintersinger, A. Ershova, W. M. Shih, *Nat. Commun.* **2021**, *12*, 1741.
- [64] J. Kong, J. Zhu, K. Chen, U. F. Keyser, *Adv. Funct. Mater.* **2019**, *29*, 1807555.
- [65] S. Cai, T. Pataillot-Meakin, A. Shibakawa, R. Ren, C. L. Bevan, S. Ladame, A. P. Ivanov, J. B. Edel, *Nat. Commun.* **2021**, *12*, 3515.
- [66] M. M. Kaminski, O. O. Abudayyeh, J. S. Gootenberg, F. Zhang, J. J. Collins, *Nat. Biomed. Eng.* **2021**, *5*, 643.
- [67] J. S. Gootenberg, O. O. Abudayyeh, J. W. Lee, P. Essletzbichler, A. J. Dy, J. Joung, V. Verdine, N. Donghia, N. M. Daringer, C. A. Freije, C. Myhrvold, R. P. Bhattacharyya, J. Livny, A. Regev, E. V. Koonin, D. T. Hung, P. C. Sabeti, J. J. Collins, F. Zhang, *Science* **2017**, *356*, 438.
- [68] J. S. Chen, E. Ma, L. B. Harrington, M. Da Costa, X. Tian, J. M. Palefsky, J. A. Doudna, *Science* **2018**, *360*, 436.
- [69] C. Jiao, S. Sharma, G. Dugar, N. L. Peeck, T. Bischler, F. Wimmer, Y. Yu, L. Barquist, C. Schoen, O. Kurzai, C. M. Sharma, C. L. Beisel, *Science* **2021**, *372*, 941.
- [70] D. Jeong, M. Klocke, S. Agarwal, J. Kim, S. Choi, E. Franco, J. Kim, *Methods Protoc.* **2019**, *2*, 39.
- [71] A. Padirac, T. Fujii, Y. Rondelez, *Curr. Opin. Biotechnol.* **2013**, *24*, 575.
- [72] T. S. Gardner, C. R. Cantor, J. J. Collins, *Nature* **2000**, *403*, 339.
- [73] M. B. Elowitz, S. Leibler, *Nature* **2000**, *403*, 335.
- [74] A. Tamsir, J. J. Tabor, C. A. Voigt, *Nature* **2011**, *469*, 212.
- [75] Y. Benenson, *Nat. Rev. Genet.* **2012**, *13*, 455.
- [76] A. Courbet, D. Endy, E. Renard, F. Molina, J. Bonnet, *Sci. Transl. Med.* **2015**, *7*, 289ra83.
- [77] V. Noireaux, R. Bar-Ziv, A. Libchaber, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 12672.
- [78] T. Wang, Y. Lu, *Biosensors* **2022**, *12*, 318.
- [79] N. Srinivas, J. Parkin, G. Seelig, E. Winfree, D. Soloveichik, *Science* **2017**, *358*, eaal2052.
- [80] B. Yurke, A. J. Turberfield, A. P. Mills Jr., F. C. Simmel, J. L. Neumann, *Nature* **2000**, *406*, 605.
- [81] L. Qian, E. Winfree, J. Bruck, *Nature* **2011**, *475*, 368.
- [82] D. Han, Z. Zhu, C. Wu, L. Peng, L. Zhou, B. Gulbakan, G. Zhu, K. R. Williams, W. Tan, *J. Am. Chem. Soc.* **2012**, *134*, 20797.
- [83] J. Zhu, L. Zhang, Z. Zhou, S. Dong, E. Wang, *Chem. Commun.* **2014**, *50*, 3321.
- [84] J. Elbaz, O. Lioubashevski, F. Wang, F. Remacle, R. D. Levine, I. Willner, *Nat. Nanotechnol.* **2010**, *5*, 417.
- [85] M. N. Stojanovic, T. E. Mitchell, D. Stefanovic, *J. Am. Chem. Soc.* **2002**, *124*, 3555.
- [86] R. Penchovsky, *ACS Synth. Biol.* **2012**, *1*, 471.
- [87] R. R. Breaker, *Curr. Opin. Biotechnol.* **2002**, *13*, 31.
- [88] E. Winfree, F. Liu, L. A. Wenzler, N. C. Seeman, *Nature* **1998**, *394*, 539.
- [89] C. G. Evans, E. Winfree, *Chem. Soc. Rev.* **2017**, *46*, 3808.
- [90] D. Y. Zhang, R. F. Hariadi, H. M. T. Choi, E. Winfree, *Nat. Commun.* **2013**, *4*, 1965.
- [91] J. Kim, E. Winfree, *Mol. Syst. Biol.* **2011**, *7*, 465.
- [92] J. Kim, K. S. White, E. Winfree, *Mol. Syst. Biol.* **2006**, *2*, 68.
- [93] J. Kim, J. Hopfield, E. Winfree, in *Advances in Neural Information Processing Systems 17* (Eds: L.K. Saul, Y. Weiss, L. Bottou), MIT Press, Cambridge **2005**, p. 681.
- [94] J. Kim, J. F. Quijano, J. Kim, E. Yeung, R. M. Murray, *Biotechnol. J.* **2022**, *17*, 2000449.
- [95] S. Kar, A. D. Ellington, *ACS Synth. Biol.* **2018**, *7*, 1937.
- [96] S. Tyagi, F. R. Kramer, *Nat. Biotechnol.* **1996**, *14*, 303.
- [97] M. Takinoue, D. Kiga, K. Shohda, A. Suyama, *Phys. Rev. E* **2008**, *78*, 041921.
- [98] S. Ayukawa, M. Takinoue, D. Kiga, *Acc. Chem. Res.* **2011**, *44*, 1369.
- [99] N. Nitta, A. Suyama, in *DNA Computing* (Eds: J. Chen, J. Reif), Springer Berlin Heidelberg, Berlin, Heidelberg **2004**, p. 203.
- [100] K. Montagne, R. Plasson, Y. Sakai, T. Fujii, Y. Rondelez, *Mol. Syst. Biol.* **2011**, *7*, 466.
- [101] A. Baccouche, K. Montagne, A. Padirac, T. Fujii, Y. Rondelez, *Methods* **2014**, *67*, 234.
- [102] A. Padirac, T. Fujii, Y. Rondelez, *Proc. Natl. Acad. Sci. USA* **2012**, *109*, E3212.
- [103] G. Gines, R. Menezes, K. Nara, A.-S. Kirstetter, V. Taly, Y. Rondelez, *Sci. Adv.* **2020**, *6*, 5952.
- [104] S. Shah, J. Wee, T. Song, L. Ceze, K. Strauss, Y.-J. Chen, J. Reif, *J. Am. Chem. Soc.* **2020**, *142*, 9587.
- [105] T. Song, A. Eshra, S. Shah, H. Bui, D. Fu, M. Yang, R. Mokhtar, J. Reif, *Nat. Nanotechnol.* **2019**, *14*, 1075.
- [106] Y. Benenson, B. Gil, U. Ben-Dor, R. Adar, E. Shapiro, *Nature* **2004**, *429*, 423.
- [107] S. Zhao, Y. Liu, B. Wang, C. Zhou, Q. Zhang, *New J. Chem.* **2020**, *44*, 1931.
- [108] T. K. Lu, A. S. Khalil, J. J. Collins, *Nat. Biotechnol.* **2009**, *27*, 1139.
- [109] N. Roquet, T. K. Lu, *Biotechnol. J.* **2014**, *9*, 597.
- [110] T. Miyamoto, S. Razavi, R. DeRose, T. Inoue, *ACS Synth. Biol.* **2013**, *2*, 72.
- [111] Q. Ma, C. Zhang, M. Zhang, D. Han, W. Tan, *Small Struct.* **2021**, *2*, 2100051.
- [112] C. E. Shannon, *Trans. Am. Inst. Electr. Eng.* **1938**, *57*, 713.
- [113] F. Wang, H. Lv, Q. Li, J. Li, X. Zhang, J. Shi, L. Wang, C. Fan, *Nat. Commun.* **2020**, *11*, 121.
- [114] N. Takeuchi, M. Hiratani, R. Kawano, *JACS Au* **2022**, *2*, 1829.
- [115] J. K. Jung, K. K. Alam, M. S. Verosloff, D. A. Capdevila, M. Desmau, P. R. Clauer, J. W. Lee, P. Q. Nguyen, P. A. Pastén, S. J. Matiassek, J.-F. Gaillard, D. P. Giedroc, J. J. Collins, J. B. Lucks, *Nat. Biotechnol.* **2020**, *38*, 1451.
- [116] D. Woods, D. Doty, C. Myhrvold, J. Hui, F. Zhou, P. Yin, E. Winfree, *Nature* **2019**, *567*, 366.
- [117] J. Bonnet, P. Yin, M. E. Ortiz, P. Subsoontorn, D. Endy, *Science* **2013**, *340*, 599.
- [118] Z. Xie, L. Wroblewska, L. Prochazka, R. Weiss, Y. Benenson, *Science* **2011**, *333*, 1307.
- [119] A. A. Green, J. Kim, D. Ma, P. A. Silver, J. J. Collins, P. Yin, *Nature* **2017**, *548*, 117.
- [120] N. H. Farhat, E. D. M. Hernandez, in *From Natural to Artificial Neural Computation* (Eds: J. Mira, F. Sandoval), Springer, Berlin, Heidelberg **1995**, p. 215.
- [121] K. M. Cherry, L. Qian, *Nature* **2018**, *559*, 370.
- [122] A. J. Genot, T. Fujii, Y. Rondelez, *J. R. Soc., Interface* **2013**, *10*, 20130212.
- [123] X. Xiong, T. Zhu, Y. Zhu, M. Cao, J. Xiao, L. Li, F. Wang, C. Fan, H. Pei, *Nat. Mach. Intell.* **2022**, *4*, 625.
- [124] R. Lopez, R. Wang, G. Seelig, *Nat. Chem.* **2018**, *10*, 746.
- [125] S. Okumura, G. Gines, N. Lobato-Dauzier, A. Baccouche, R. Deteix, T. Fujii, Y. Rondelez, A. J. Genot, *Nature* **2022**, *610*, 496.
- [126] X. Li, L. Rizik, V. Kravchik, M. Khoury, N. Korin, R. Daniel, *Nat. Commun.* **2021**, *12*, 3139.
- [127] A. Pandi, M. Koch, P. L. Voyvodic, P. Soudier, J. Bonnet, M. Kushwaha, J.-L. Faulon, *Nat. Commun.* **2019**, *10*, 3880.

- [128] B. Li, X. Chen, A. D. Ellington, *Anal. Chem.* **2012**, *84*, 8371.
 [129] S. M. Douglas, I. Bachelet, G. M. Church, *Science* **2012**, *335*, 831.
 [130] L. H. H. Meijer, A. Joesaar, E. Steur, W. Engelen, R. A. van Santen, M. Merx, T. F. A. de Greef, *Nat. Commun.* **2017**, *8*, 1117.
 [131] R. M. Zadegan, M. D. E. Jepsen, L. L. Hildebrandt, V. Birkedal, J. Kjems, *Small* **2015**, *11*, 1811.
 [132] P. Banda, C. Teuscher, D. Stefanovic, *J. R. Soc., Interface* **2014**, *11*, 20131100.
 [133] T. P. Lillicrap, A. Santoro, L. Marris, C. J. Akerman, G. Hinton, *Nat. Rev. Neurosci.* **2020**, *21*, 335.
 [134] M. R. Lakin, A. Minnich, T. Lane, D. Stefanovic, *J. R. Soc., Interface* **2014**, *11*, 20140902.
 [135] M. R. Lakin, D. Stefanovic, *ACS Synth. Biol.* **2016**, *5*, 885.
 [136] J. C. R. Whittington, R. Bogacz, *Trends Cognit. Sci.* **2019**, *23*, 235.
 [137] A. P. Lapteva, N. Sarraf, L. Qian, *J. Am. Chem. Soc.* **2022**, *144*, 12443.
 [138] V. Hsiao, Y. Hori, P. W. Rothmund, R. M. Murray, *Mol. Syst. Biol.* **2016**, *12*, 869.
 [139] N. Lobato-Dauzier, A. Baccouche, G. Gines, T. Lévi, Y. Rondelez, T. Fujii, S. H. Kim, N. Aubert-Kato, A. J. Genot, *BioRxiv* **2022**, <https://doi.org/10.1101/2022.07.06.498929>.
 [140] W. Yahiro, N. Aubert-Kato, M. Hagiya, in *ALIFE 2018: The 2018 Conference on Artificial Life*, **2018**, https://doi.org/10.1162/isal_a_00293.
 [141] N. Lobato-Dauzier, L. Cazenille, T. Fujii, A. Genot, N. Aubert-Kato, in *ALIFE 2020: The 2020 Conference on Artificial Life*, **2020**, https://doi.org/10.1162/isal_a_00293.
 [142] A. Goudarzi, M. R. Lakin, D. Stefanovic, in *DNA Computing and Molecular Programming* (Eds: D. Soloveichik, B. Yurke), Springer International Publishing, Cham **2013**, p. 76.
 [143] X. Liu, K. K. Parhi, *ACS Synth. Biol.* **2022**, *11*, 780.
 [144] A. J. Thubagere, W. Li, R. F. Johnson, Z. Chen, S. Doroudi, Y. L. Lee, G. Izatt, S. Wittman, N. Srinivas, D. Woods, E. Winfree, L. Qian, *Science* **2017**, *357*, eaan6558.
 [145] G. Chatterjee, N. Dalchau, R. A. Muscat, A. Phillips, G. Seelig, *Nat. Nanotechnol.* **2017**, *12*, 920.



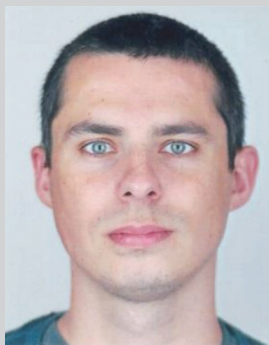
Coline Kieffer is a Ph.D. student at the Gulliver Laboratory in ESPCI-PSL under the supervision of Dr. Y. Rondelez and Dr. G. Gines. She obtained her MSc in Molecular Biology and Biochemistry at Sorbonne University in Paris, where she specialized in RNA biology. She is currently developing of a novel method for the digital and multiplex detection of microRNA based on molecular programming (PEN-DNA toolbox), microfluidics, and flow cytometry.



Anthony Genot is a CNRS Researcher in the International Research Laboratory LIMMS between CNRS and the University of Tokyo. He was trained as a Physicist, and his research interests include the storage and processing of information in DNA, DNA-based materials (DNA hydrogels and DNA-based colloids), as well as microfluidics.



Yannick Rondelez is a Researcher working for the French CNRS at PSL-ESPCI, in Paris, in the field of biomolecular complex systems, enzyme engineering, microfluidics, and diagnostics. His original training concerns physico-chemistry and biochemistry. The inspiration for his work generally comes from biological systems, which use various molecular means, such as informational polymers and biochemical circuits to manipulate information, perform optimization tasks, make decisions, and store data.



Guillaume Gines obtained his Ph.D. in Chemistry and Biology from the University of Grenoble in 2013. He then moved to the University of Tokyo as a Postdoctoral Fellow studying the collective behaviors in populations of DNA-programmed particles. In 2016, he joined the Gulliver laboratory at PSL-ESPCI Paris. Since 2020, he has been a CNRS Researcher whose research activities revolve around DNA nanotechnologies, biosensing, and microfluidics.