

# Computational Biosensors: Molecules, Algorithms, and Detection Platforms

Elebeoba E. May, Jason C. Harper and Susan M. Brozik

**Abstract** Advanced nucleic acid-based sensor-applications require computationally intelligent biosensors that are able to concurrently perform complex detection and classification of samples within an in vitro platform. Realization of these cutting-edge computational biosensor systems necessitates innovation and integration of three key technologies: molecular probes with computational capabilities, algorithmic methods to enable in vitro computational post processing and classification, and immobilization and detection approaches that enable the realization of deployable computational biosensor platforms. We provide an overview of current technologies, including our contributions towards the development of computational biosensor systems.

**Keywords** Biomolecular logic systems · Digital biosensors · DNA computing · Computational biosensors · Intelligent biosensor · Application- and Substrate-specific algorithms · DNA detection · Deoxyribozymes · Enzymes

## 1 Introduction

The ability to detect and discriminate nucleic acid sequences is necessary for a wide variety of applications: high throughput screening, mutation tracking for disease emergence, genetically modified organism (GMO) monitoring, molecular computing, biometrics fingerprinting, and various genotype associated studies. Traditional sensor systems are multistep platforms that often times rely heavily on

---

In memory of Dr. Susan M. Brozik, deceased January 2014.

---

E.E. May (✉)  
University of Huston, Huston, TX, USA  
e-mail: eemay@uh.edu

J.C. Harper · S.M. Brozik  
Sandia National Laboratories, Albuquerque, NM, USA

off-platform post-processing to determine the outcome of detection or classify the biomolecule detected. Current high-throughput systems use traditional or silicon-based computing for interpretation of molecular recognition events. Complex bioinformatics algorithms are tasked with de-noising and processing sensor output signals. This approach can be error prone and not easily integrated into emerging microsystem technologies for hand held, lab-on-chip systems. Portable computational biosensor systems would be particularly useful for monitoring and diagnostic applications in resource-limited settings or situations, such as those encountered in developing countries and in military medical support applications.

The objective of this chapter is to provide an overview of capabilities needed for creating advanced, computationally intelligent, nucleic acid-based microsensor platforms that are reliable, deployable, and limits reliance on off-platform computational post-processing. In this chapter we review and discuss technologies and algorithmic advances necessary for the realization of integrated computational biosensor systems including: Nucleic acid based probes that enable detection and computation; Algorithm development for probe selection and target classification that transfers computational intelligence from silicon-based processing to nucleic acid based computational elements; Development of immobilization and detection approaches that permit the implementation and integration of computational biosensors on deployable platforms.

Nucleic acid based sensors use various types of probes, with single-stranded DNA, such as those used in microarrays, remaining the more widely employed. As an alternative to single stranded DNA technologies, we will discuss the use of molecular beacon, single-stranded oligonucleotide probes that incorporate structure to enhance functionalization [33, 118, 127]. Molecular beacons are able to detect mutations in target sequences and can be multiplexed; these properties make molecular beacons effective platforms for detecting genetically modified targets in various biosensor systems. Catalytic molecular beacons, such as deoxyribozymes, integrate enzymatic properties and modularity into standard beacons [98]. We will review characteristics of molecular beacon based probes and discuss their viability as molecular substrates for constructing computational biosensor platforms.

A major challenge in the development of computational biosensors is in the *de novo* development of application and substrate-specific algorithms that enable the transfer of computational intelligence from a silicon-based system to a nucleic acid based system. The algorithmic challenge is multifaceted and linked, with application-specific library design being the first aspect, and the second being the development of computational primitives that can systematically compute a classification outcome using the molecular substrate. We will discuss various approaches to address this challenge, including information-based methods for probe design and classification algorithm generation.

One of the drawbacks of optical-based detection platforms for molecular beacons is the reliance on laboratory-based equipment for analysis. Secondly, fluorescence detection may limit the integration of the computational sensor with traditional silicon-based microprocessors used in hand held, portable sensor systems. Electrochemical detection provides significant advantages over other

common detection and signal transduction methods. Specifically, it is sensitive, rapid, and can provide greater specificity and sensitivity over optical detection methods, as interfering background fluorescence does not adversely affect the electrochemical signal. Electrochemical detection systems are also more amenable to miniaturization and integration given that the signal is already in a form that can be interpreted by silicon microprocessors and integrated circuit fabrication methods can be leveraged to produce the miniaturized components. We discuss the development of immobilization and detection technologies that enable the design and implementation of application specific integrated computational biosensor systems.

## 2 Computational Biosensors

Computational biosensors have roots in the emerging field of biocomputing. Increased research in biocomputing has resulted in significant advances in processing of chemical and biochemical information; this success is due, in part, to the innate high specificity and selectivity of biological molecules (e.g. enzymatic catalysis of substrate to product; nucleic acid binding of Watson-Crick base pairs). Further, the general compatibility of biomolecules allows for the intimate assembly of differing biomolecules within cascading networks capable of performing diverse reactions. Biocomputing systems have been reported that are assembled from proteins/enzymes [96, 113], RNA [124], DNA [101], and even whole cells [94]. Exploiting properties intrinsic to biomolecules for unconventional computing has resulted in greater success over chemical computing strategies alone [42].

Biocomputing shows perhaps the greatest promise for use with analytical systems, particularly for biomedical applications [60, 64, 115]. Use of biocomputing and logic operations could yield a novel class of computationally intelligent biosensors that can accept diverse input signals and systematically compute and output an actionable signal (Wang and Katz 2010). Additionally, signal output from the biosensors can be coupled to signal-responsive materials or processes that allow the system to sense, and then act [81]. Such ‘intelligent’ biosensors would offer significant advantages over traditional biosensors which typically accept only a single input, and then output a signal that must be post-processed and further analyzed by a well-trained operator for meaningful conclusions to be drawn.

Successful implementation of biocomputing principles with biosensing has led to significant and exciting advances in intelligent computational biosensors. Often referred to as ‘digital biosensors,’ or ‘bio-logic’ analysis, the vast majority of reported systems rely on either enzyme cascades or networks of DNA probes. Enzyme-based computational biosensors will be introduced and briefly discussed. A thorough review of enzyme logic systems lies outside the scope of this chapter which focuses on nucleic acid probe-based computational biosensors.

## 2.1 *Enzyme-Based Computational Biosensors*

Enzyme logic systems rely on multiple enzymes (at least 2, typically 3 or more) to develop a signaling cascade that results in the processing of chemical information in a manner that is comparable with Boolean logic operations (e.g. AND, OR, NAND, NOR, XOR, XNOR) [7, 82, 106, 132]. Combining the logic operations in a small logic network results in biosensor output in the form of a Yes/No response. Inputs to the sensor are typically chemicals and/or enzyme(s), and the Yes/No output is typically a change in solution pH, color, optical density, or electrochemical properties. This approach has been used to perform simple arithmetic functions [6]. More interestingly, enzyme logic systems have been coupled to signal-responsive materials forming switchable membranes [110], emulsions [70], nanoparticle assemblies [69], or modified electrode surfaces [83, 134].

Enzyme-based computational biosensors have been developed that focus on biomedically relevant markers/signaling molecules at their physically relevant concentrations. For example, an enzyme-based computational biosensor was developed to process biochemical information related to the pathophysiological conditions associated from traumatic brain injury and hemorrhagic shock [78]. Other biomedically relevant enzyme-based computation biosensors have been reported, including one capable of discriminating biomarkers characteristic of liver injury, soft tissue injury, and abdominal trauma [133]. Of note is the robustness of this system which was designed to operate in serum solutions spiked with injury biomarkers in order to mimic real medical samples.

Transitioning enzyme-based computational biosensors from *in vitro* to *in vivo* analysis is particularly challenging as operations are composed of several biocatalytic steps that must each be optimized to minimize signal noise, ensure adequate signal amplification, and reduce the impacts of many different interferants present in real biological fluids. The long-term instability of enzymes under ambient conditions further complicates development of practical, fieldable sensing systems relying on enzymes for signal processing. For further information of enzyme-based computational biosensors, the reader is referred to these excellent review articles [119, 120].

## 2.2 *Nucleic Acid-Based Computational Biosensors*

DNA-based biocomputing is a well-developed field, due in part to the high stability and rich information capacity of DNA molecules [101]. Typically, DNA biocomputing systems solve analytical problems using a combinatorial approach. DNA sequences are generated that encode all possible solutions to the problem. The correct solution is selected from the DNA library using polymerase chain reaction (PCR), affinity-purification and gel electrophoresis designed to only amplify/select DNA strands that meet the criteria for a correct solution. This strategy was first used

by Adleman in 1994 [2] to solve a NP-complete problem (no efficient algorithms are known to solve NP-problems). Researchers have proposed several applications using a DNA computing framework including algorithms for breaking the Data Encryption Standard, DNA encryption methods, and techniques to investigate nature's cellular computing processes [3, 41, 46, 56]. More recently, Adleman and coworkers solved a 20 variable 3-SAT problem (a NP-complete problem requiring exponential time to solve, with 1,048,575 possible truth assignments) using DNA biocomputing [12].

DNA-based biocomputing has also been successfully employed for biomedically relevant applications. A stochastic computing system was reported in which a varying concentration of biomedical marker inputs provided competition between two alternative biochemical pathways that output a DNA molecule encoding the result [1]. This technique was subsequently used to analyze the concentration of cancer-related molecular indicators (i.e. mRNA). In the presence of the marker, the system output was a single stranded antisense DNA therapeutic against the cancer mRNA [8]. Such 'intelligent' systems combining sensing, computation, and action hold great promise for in situ medical diagnosis and treatment [93].

In nucleic acid-based computation and biosensing, the structure of the DNA probe(s) can have a profound impact on the functionality of the given system. The advantages and limitations of single-stranded DNA, DNA molecular beacons, and catalytic DNA molecular beacons are discussed below.

### 2.2.1 Single-Stranded Nucleic Acid Probes

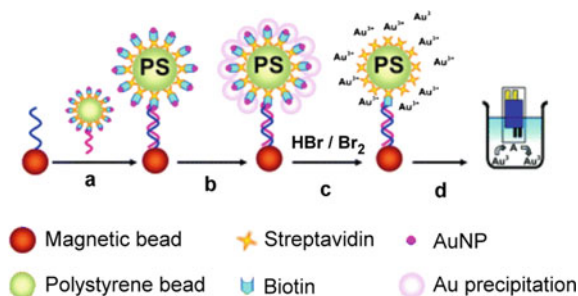
The most widely employed DNA probe remains single stranded DNA. Used in DNA biocomputation systems, simple DNA-based biosensors, and DNA microarrays, single-stranded DNA probes are effective at detecting complementary DNA and RNA sequences due to strong Watson-Crick base pair binding [118]. Base pair binding results in exceptionally high specificity for the target sequence. Various methods are used to indicate that hybridization has occurred between the probe and the complementary target strand. These signal transduction methods are discussed in Sect. 4 of this chapter.

Despite the exceptionally high specificity intrinsic to DNA-based biosensors, the detection limit for most sensors is not sufficient to detect the target sequence at the very low concentrations typically found in real-world samples. This is partly due to the 1:1 target:signal stoichiometry which limits the number of signal events to the amount of target present in the sample. Thus, nearly all DNA-based biosensors rely on either (1) amplification of the target prior to introduction to the sensor, or (2) post-hybridization labeling that significantly amplifies the signal from each hybridization event. For example, a recent article reviewed detection technologies for identifying genetically modified organisms (GMO) in food and feed [5]. The authors stated that DNA-based biosensors are the 'leading edge' technology for simple, rapid and inexpensive testing for GMOs. However, they reported that for successful detection nearly all DNA-based GMO sensors required PCR

amplification of samples extracted from raw ingredients and processed food prior to introduction to the sensor. The requirement of DNA sample pre-amplification via PCR makes these sensors difficult to implement as platforms for autonomous or field deployable biosensing.

Regarding post-hybridization amplification, a multitude of DNA-based sensors have been reported that use nanoparticles, dendrimers, supramolecular assemblies, and the like for labeling DNA hybridization events [79]. Detection of the label provides many orders of magnitude enhancement in the signal from each hybridization event. For example, Kawde and Wang used a triple amplification technique resulting in attomolar DNA detection limits [43]. As shown in Fig. 1, streptavidin coated polystyrene spheres (PS) decorated with biotinylated gold nanoparticles (AuNPs) were used as labels for DNA hybridization. Hybridization between the magnetic bead-bound DNA probe and the labeled target sequence (step a) was followed by further catalytic deposition of gold onto the AuNPs (step b). Detection of the binding event occurred by dissolution of the Au (step c) and stripping voltammetry in which the gold ions are accumulated onto a carbon electrode over many seconds, and then stripped from the electrode by reversing the applied potential. The multi-step strategy yielded a 300 attomolar ( $10^{-18}$  mol/L) DNA detection limit, eliminating the need for sample pre-amplification by PCR. However, the steps required for this, and other similar amplification strategies, significantly increase the complexity of the sensor making it impractical as a reliable and robust fieldable device.

Notwithstanding these limitations, the ability to multiplex DNA-based sensors into microarrays has resulted in the development of powerful analytical tools commonly referred to as gene chips, or biochips. DNA microarrays are widely used for high throughput analysis of DNA and RNA samples for differential gene expression levels, diagnosis of genetic diseases, detection and identification of infectious agents, determination of DNA-protein interactions, drug screening and forensic analysis [32, 118]. These hybridization chips are fabricated by



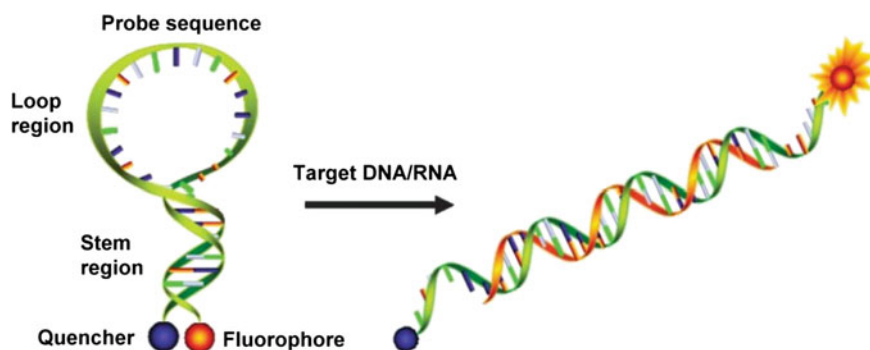
**Fig. 1** Schematic illustration of amplified DNA detection employing probe DNA modified magnetic beads. **a** Treatment of probe with target DNA labeled with streptavidin conjugated polystyrene beads loaded with biotinylated AuNPs, followed by **b** Au precipitation onto AuNP seeds, **c** dissolution of the Au and **d** detection via electrochemical stripping. (Adapted with permission from [43]. Copyright 2004, Wiley.)

immobilization, or in situ synthesis, of thousands of single stranded DNA probes (up to  $10^6$  unique probes/cm<sup>2</sup>) on a planar support (glass, silicon, or plastic) [118]. Target samples are introduced to the array and hybridization reactions are performed under conditions of high stringency such that mishybridizations are prevented. The vast majority of commercial gene chips utilize fluorescence for monitoring hybridization events using a bench-top fluorescent scanner, followed by software post-processing and expert analysis. Thus, these systems are not amenable for portable field use.

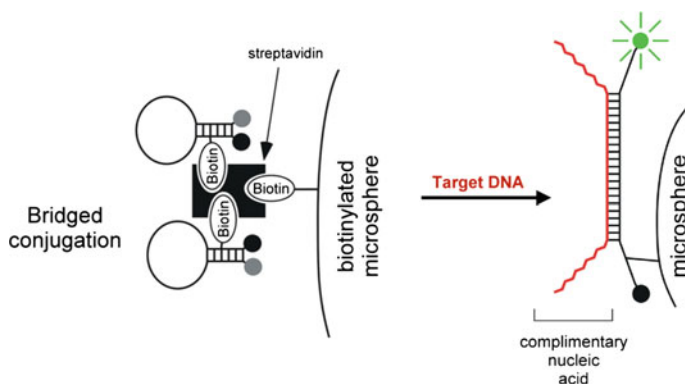
### 2.2.2 Molecular Beacon Nucleic Acid Probes

Molecular beacons are single-stranded oligonucleotide probes in which five to seven base pairs on either end of the strand are complementary to each other. The complementarity results in the strand ends hybridizing and forming a stem-loop structure [112], as shown in Fig. 2. The loop region of the strand contains a probe sequence that is complementary to the target sequence. Traditional molecular beacons contain a fluorophore and quencher on each arm of the stem. In the absence of target molecules, the stem-loop structure forms and the fluorophore is brought into close proximity to the quencher, suppressing fluorescence. When target DNA/RNA is present, hybridization occurs with the loop, breaking the stem region and separating the fluorophore and quencher. This results in a strong increase in probe fluorescence (see Fig. 2).

As opposed to traditional single-stranded DNA based biocomputing and biosensing systems, incorporation of molecular beacons that exploit structure, and changes in structure, adds another dimension of functionality to the system. Molecular beacons do not require additional labeling steps post-hybridization as the hybridization event itself results in an action (conformational change) that increases



**Fig. 2** Structural characteristics of a typical DNA molecular beacon probe showing formation of a stem-loop structure in the absence of target DNA/RNA. Breaking of the stem portion occurs when target is present and binds to the probe sequence loop region. (Adapted with permission from [107]. Copyright 2005, Royal Society of Chemistry.)



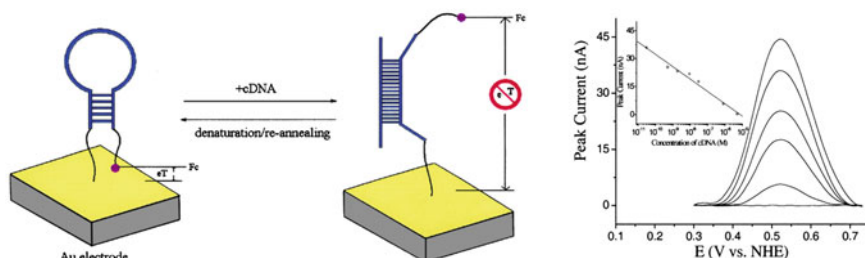
**Fig. 3** Streptavidine bridged conjugation of biotinylated molecular beacons to a biotinylated microsphere. Introduction of target DNA and subsequent hybridization results in a conformational change, separating the fluorophore and quencher. (Adapted with permission from [33]. Copyright 2005, Oxford Journals.)

fluorescent output. This simplifies the system, increasing the likelihood that such a system could be developed into a deployable biosensor. Also, unlike microarray detection, molecular beacon probes can be used to detect the presence of target DNA or RNA sequences in complex mixtures [93].

Molecular beacons have found wide use in bioanalytical systems [107]. In one example, Horejsh et al. immobilized molecular beacons onto solid supports for multiplex detection of unlabeled nucleic acids in solution [33]. As shown in Fig. 3, biotinylated molecular beacons were conjugated to biotinylated microspheres using streptavidin. This streptavidin ‘bridged’ conjugation resulted in a nearly two fold improvement in signal to noise ratio over directly conjugating biotinylated molecular beacons to streptavidin coated microspheres. Using two different bead sizes and molecular beacons with two fluorophore colors allowed for flow cytometry based detection of nucleic acid sequences indicative of three respiratory pathogens: SARS coronavirus, parainfluenza virus type 3 (PIV-3), and respiratory syncytial virus (RSV). Micron and submicron molecular beacon-based DNA sensors have been reported using molecular-beacons immobilized on optical fibers [51]. Molecular beacons have also been immobilized on quantum dots and used for live intracellular monitoring [131], expanding their potential use for in vivo diagnostic and therapeutic applications.

The utility of molecular beacon-based oligonucleotide probes is not limited to fluorescence detection systems. Fan and coworkers reported the use a traditional stem-loop molecular beacon that was modified with an electroactive group (ferrocene) on one end, and a thiol molecule on the other end [26]. The thiol moiety facilitated self-assembly and immobilization of the molecular beacon onto a gold electrode surface, as shown in Fig. 4. In the absence of complementary DNA, the hairpin loop brings the ferrocene group in close proximity to the gold electrode surface, facilitating efficient redox reaction electron transfer. Upon hybridization the



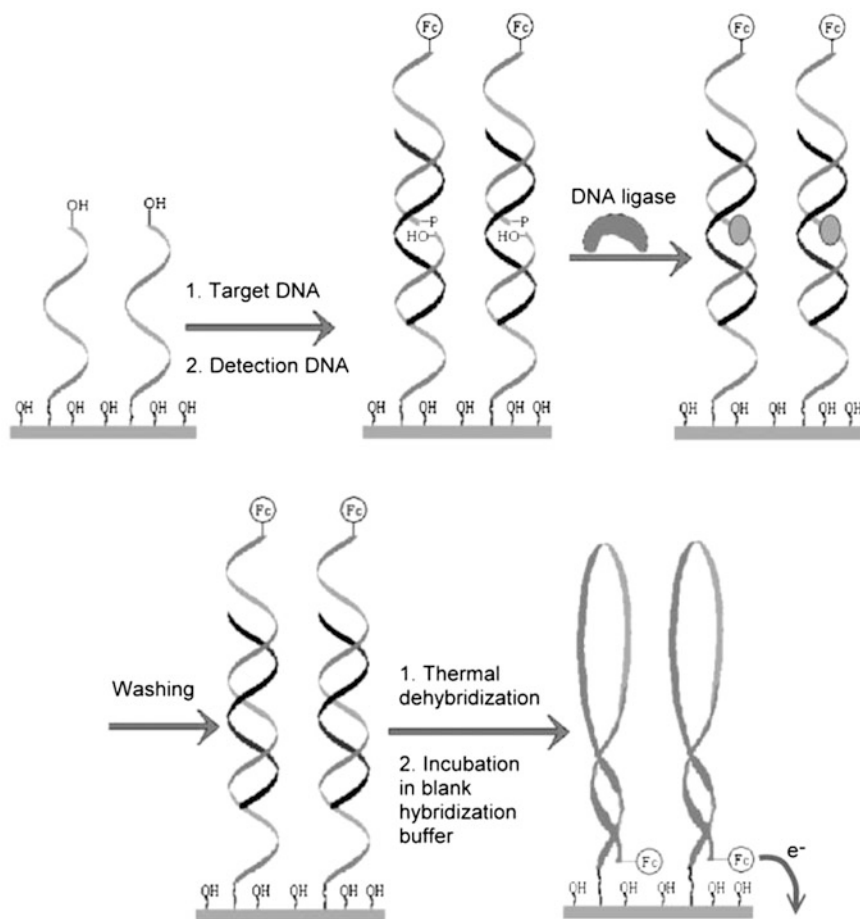


**Fig. 4** Electrochemical detection using a molecular beacon oligonucleotide modified with a redox active ferrocene group on the 5' end, and a thiol group on the 3' end. The probe is immobilized via self-assembly between the thiol moiety and the Au electrode surface. Electron transfer ( $eT$ ) is disrupted in the presence of target DNA. The *right* most panel shows background subtracted voltammograms with peak current decreasing as complementary DNA concentration increases. (Adapted with permission from [26]. Copyright 2003, PNAS USA.)

stem-loop structure is broken and the distance from the ferrocene tag to the electrode surface is increased. This reduced redox reaction associated electron transfer. This sensor is reusable and reagentless, making it well suited for device development and continuous monitoring of a given analyte. However, 'signal-off' sensors are more prone to false positive responses. Also, the differential in current response from no analyte present, to fully saturated with target analyte, was only 45 nA. Such a low signal differential can make analyzing the sensor output challenging.

Molecular beacon based sensors have also been reported with single nucleotide specificity, capable of identifying the presence of point mutations in the target sequence. For example, Wu and co-workers developed a 'signal-on' electrochemical DNA sensor based on DNA ligase and a 'reverse' molecular beacon [127]. In this approach, a single-stranded DNA capture probe was first immobilized onto a gold electrode surface via gold-thiol self-assembly. Introduction of the target DNA resulted in hybridization between the capture probe and a portion of the target DNA. The remaining portion of the target DNA is complementary to a detection DNA strand, and formed a 'sandwich.' As shown in Fig. 5, the detection DNA contains a phosphoryl group at the 5' end, and a ferrocene tag at the free 3' end. Addition of DNA ligase to the system joins the detection DNA to the surface immobilized capture DNA, forming a single DNA strand. Washing and thermal dehybridization removes the protein and target DNA. Incubation in blank hybridization buffer results in formation of a stem-loop structure as the detection DNA strand contains a 6-base sequence near the ferrocene group that is complementary to the 5' end of the capture probe. This 'reverse' molecular beacon results in an increase in electron transfer when target DNA was present.

Wu and co-workers also reported that this system was capable of discriminating mutations in DNA sequences. Eight single-base mutations in the target sequence were evaluated. Point mutations at the nick location provided the lowest signal response, with G:T mismatches inducing slight higher currents. Relatively higher currents were observed when point mutations were on the 5'-side of the nick versus



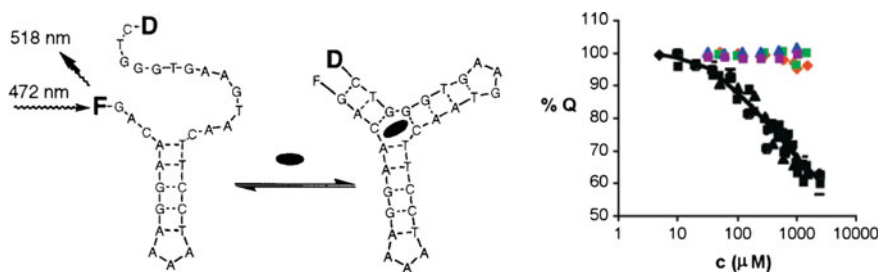
**Fig. 5** Schematic illustration of a 'reverse' molecular beacon DNA detector with an electrochemical 'signal-on' mechanism. Target DNA is 'sandwiched' between the surface immobilized capture probe DNA and ferrocene label detection DNA. Enzymatic ligation joins the capture and detection sequences, resulting in a hairpin structure that positions the ferrocene label in close proximity to the electrode surface, permitting ready electron transfer. (Adapted with permission from [127]. Copyright 2007, Wiley.)

the 3'-side. Finally, the relatively highest currents from point mutation strands were measured when single-base mutations were at the third and fourth position from the ligation point. Still, the currents measured for these single-base pair mismatches were between only 0.6 and 5% of the signal obtained from the fully complementary target sequence which was approximately 70 nA at the highest target concentration. Using these subtle differences in the low nA current range to discriminate and identify point mutations would be challenging. Further, this detection scheme required several incubation/washing steps and addition of reagents, again making it difficult to employ in a setting outside the laboratory.

### 2.2.3 Aptamer Beacon Nucleic Acid Probes

Another important advance in structure-based nucleic acid biosensor systems came with the discovery of aptamers. First reported in 1990 by three independent groups [24, 86, 111], aptamers are single-stranded DNA or RNA molecules that fold into secondary and tertiary structures that bind to a target molecule with very high affinity (nanomolar to picomolar dissociation constants are typical). Often described as homologous to antibodies, aptamers can bind molecules from small inorganic ions, proteins, to even living cells [35]. Unlike antibodies, aptamers are more stable, simple to modify with functional groups, easy to purify and produce, can be developed against virtually any molecule, and do not require animals or cell lines to generate. A potential disadvantage of aptamers is that conditions under which the target molecule and aptamer are introduced (e.g. ionic strength, pH, temperature, protein content) cannot deviate significantly from those used during the SELEX selection process (systematic evolution of ligands by exponential amplification) or the aptamer will not fold and bind with high avidity.

'Aptamer beacons' [37] are similar to traditional beacons, typically contain a fluorophore and a quencher, and undergo a conformational change upon binding to the target that can be monitored by fluorescence signal readout. An early example is the aptamer beacon sensor for cocaine developed by Stojanovic et al. [98]. As shown in Fig. 6, a nucleic acid-based aptamer that recognizes cocaine was engineered with instability in one stem of a three-way junction that binds cocaine. The short stem was labeled with a fluorophore and a quencher that in the absence of cocaine remains open and emits fluorescence. Introduction of cocaine results in a conformational change that brings the fluorophore and quencher in close proximity, reducing fluorescence output (see Fig. 6, right panel). The aptamer beacon was selective over cocaine metabolites and could operate in serum, although it was



**Fig. 6** An aptamer beacon with a fluorophore (fluorecene, F) on the 5' end, and a quencher (dabcyl, D) on the 3' end. In the open conformation fluorescence is observed. As the target, cocaine, is introduced the aptamer binds resulting in a closed conformation that quenches fluorescence (smaller font F). The right panel shows percent fluorescence quenching versus concentration of cocaine (*black*, seven different runs within two months), cocaine metabolites benzoyl-econine (*red*) and ecgonine methyl ester (*green*), and control aptamer (*blue*). (Adapted with permission from [98]. Copyright 2001, American Chemical Society.)

stable under only brief exposure to the serum, and the sensitivity ( $\sim 10 \mu\text{M}$  detection limit) was not high enough for clinical or forensic applications.

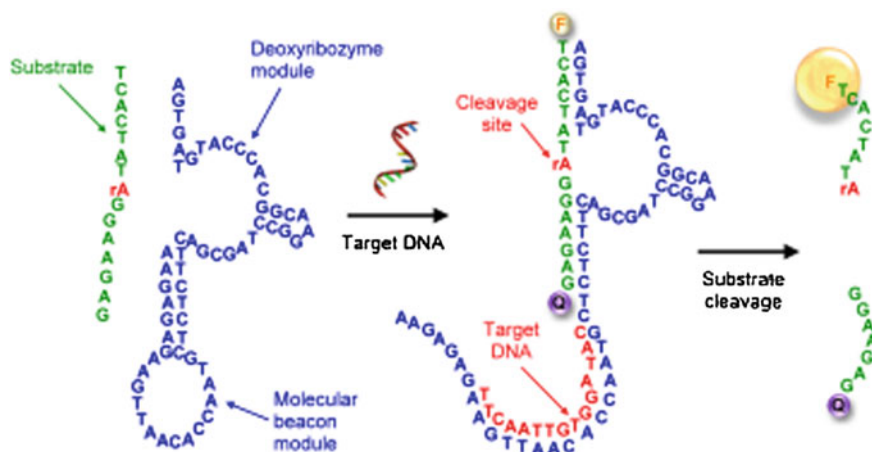
Aptamer beacons have also been immobilized onto solid supports and modified with redox active groups for reagentless and label free electrochemical detection of small molecules [87] and proteins [52, 128]. Typically, the conformational change of the aptamer brings the electroactive group closer or farther away from the electrode surface in a manner similar to that shown in Fig. 4 of Sect. 2.2.2. The change in conformation results in an increase or decrease in redox current. Although these sensors provide a simple one-step protocol for detection with no labeling steps or addition of reagents, they again show small changes in current response between no analyte present, to fully saturated with target analyte (typically 10–200 nA differential) making analysis of the sensor output difficult.

#### 2.2.4 Catalytic Molecular Beacon Nucleic Acid Probes

In the early 1980s it was discovered that naturally occurring single stranded ribonucleic acid molecules are capable of catalyzing certain biochemical reactions in a manner similar to enzymes. Termed ‘ribozymes’, or ‘catalytic RNA,’ these biological catalysts participate in RNA ligation, cleavage, synthesis, alkylation and acyl-transfer reactions, and n-glycosidic and peptide bond formation [123]. A well-studied prototype catalytic RNA molecule is the hammerhead ribozyme which recognizes specific RNA sequences and catalyzes cleavage and ligation reactions at a specific site in the recognized region [34, 80]. Hammerhead ribozymes are also capable of self-cleavage.

Although DNA molecules are much more stable and serve as the primary carrier of genetic information in nature, naturally occurring catalytic DNA molecules have not yet been identified. However, single-stranded DNA molecules capable of performing catalytic reactions similar to RNA and enzymes can be engineered. Catalytic DNA, or deoxyribozymes, have been synthesized in the laboratory via an *in vitro* iterative selection process [14, 89]. Integrating this catalytic capability of single-stranded DNA with the structural dependent functionality of molecular beacons opens intriguing opportunities for development of complex, intelligent computational biosensors.

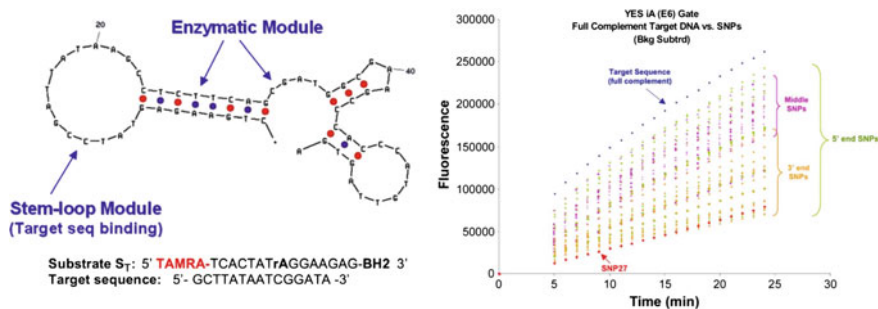
Like traditional molecular beacons, so called catalytic molecular beacons typically contain a catalytic module and a beacon module [99]. As shown in Fig. 7, the stem-loop structure of the beacon module is complementary to a portion of the catalytic module, completely inhibiting catalytic activity in the absence of target. In the presence of target oligonucleotide, the stem-loop structure is broken. The change in structure allosterically activates the deoxyribozyme complex. Leveraging the structural modularity of catalytic molecular beacons Stojanovic and colleagues have produced AND, NOT, XOR, YES and other DNA-based logic gates and circuits [54, 100, 102–104]. In the YES gate example presented in Fig. 7, the deoxyribozyme module is a hammerhead-type that cleaves an oligonucleotide substrate only in the presence of DNA target. This substrate is labeled with a



**Fig. 7** A deoxyribozyme-based catalytic molecular beacon (blue) [99]. In the absence of target DNA (red), the stem-loop region blocks access of the substrate (green) to the catalytic region of the deoxyribozyme. In the presence of target DNA, hybridization occurs with the loop region of the molecular beacon. The conformational change exposes the catalytic region to the substrate, activating catalytic cleavage of the substrate. Labeling the substrate on either end with a fluorophore and a quencher allow catalytic activity and substrate cleavage to be monitored as an increase in fluorescence. (With permission from [64]. Copyright 2008, IEEE.)

fluorophore on the 5' end, and a quencher on the 3' end. Once the catalytic molecular beacon is activated by target DNA and substrate is cleaved, fluorescence output significantly increases.

Thus, in combination with DNA information processing techniques, DNA logic gates derived from catalytic molecular beacons are viable candidates for constructing intelligent computational biosensors. They can potentially integrate several molecular inputs, perform logical operations, and output an actionable signal and/or autonomously output the appropriate therapeutic. We have reported the development of a deoxyribozyme-based intelligent computational biosensor for the detection of genetic modifications in avian influenza [64, 65]. First, the deoxyribozyme  $YES_{iA}$  (E6) gate was constructed, based on a previously reported modular design [99], that appended a 15 nucleotide molecular beacon recognition loop into a hammerhead-type deoxyribozyme (Fig. 8, left panel). Using a substrate labeled with a fluorophore and quencher, the  $YES_{iA}$  (E6) deoxyribozyme exhibited a 14-fold increase in fluorescent signal intensity, resulting in a detection limit of 22.7 pM target DNA that was resolvable in only 3 min. This impressive detection limit and response time was a consequence of the catalytic behavior of the deoxyribozyme molecular beacon. Unlike traditional molecular beacons that are limited by a 1:1 target:signal stoichiometry, a single target DNA activates a given catalytic beacon, resulting in the cleaving of several substrate molecules. Thus, PCR amplification of target DNA prior to introduction to the sensor may not be



**Fig. 8** YES<sub>IA</sub> (E6) deoxyribozyme, along with the substrate and target sequence (*left panel*) and the fluorescence output over time in the presence of the full complementary target sequence vs. 45 SNP sequences representing every possible single point mutation in the 15-mer substrate (*right panel*). (Black-Hole-2 quencher, **BH2**). (Adapted with permission from [64]. Copyright 2008, IEEE.)

required for such systems, greatly simplifying development of a portable and robust diagnostic instrument.

We also showed that this system was very sensitive at discriminating the perfect complementary target input from sequences containing single base mutations. We compiled 45 single nucleotide polymorphism (SNP) containing target sequences, representing all possible single point mutations for each of the 15 positions in the target sequence. As shown in the right panel of Fig. 8, discrimination between the targets and SNP-containing sequences was possible in only 5 min. Further, the kinetic behavior of the YES<sub>IA</sub> (E6) deoxyribozyme was impacted by the location and type of mutation in a quasi-predictable manner. This opened the exciting possibility of algorithm based analysis of the sensor output for predicting the exact location and composition of a given mutant [64]. This is potentially a powerful capability for DNA based sensors which typically cannot identify mismatches with high fidelity, and are not effective at detecting near neighbors for which the sensor was not designed. Development of a substrate-specific algorithm will be discussed more fully in Sect. 3.2.

In addition to the single input gate, a two input AND gate was developed for detecting H5N1 avian influenza virus DNA that was capable of differentiating point mutations that would increase human susceptibility. We designed the H5N1<sub>HA</sub>AND<sub>PB2</sub> gate to detect the mutated codon sequence in hemagglutinin (HA) protein that allows the virus to more readily bind to human cell receptors [129], and the mutated codon sequence in polymerase basic 2 protein (PB2) that allows for efficient replication in mammalian cells [91]. Early detection of viruses containing either or both mutations is critical to averting a pandemic. The H5N1<sub>HA</sub>AND<sub>PB2</sub> gate showed the ability to discriminate each normal input target (HA only, PB2 only, HA + PB2) and each mutant input sequence (normal HA/mutated PB2, normal PB2/mutated HA, mutated HA/mutated PB2). The normalized rate of fluorescence output was inversely proportional to the number of total mismatches [64].

Other interesting intelligent biosensors exploiting the computational ability of deoxyribozyme probes include the work of Yashin et al. in which communication between deoxyribozyme gates immobilized on beads was demonstrated [130]. The detection of an orally administered drug inducing the release of a therapeutic peptide was reported by Taylor et al. [108]. For more information on deoxyribozymes for sensing and logic gate applications the reader is referred to the review by Willner et al. [122].

### 2.2.5 DNA Molecular Nanodevices

Aptamers, deoxyribozymes, and other physical properties of single stranded nucleic acids have been used to develop integrated intelligent biosystems. Capable of complex molecular detection and amplification schemes under autonomous and isothermal conditions, such systems are often described as ‘molecular machines’, ‘nanodevices’, or ‘autonomous machines.’ For example, by combining deoxyribozyme and DNA origami technologies, Lund et al. developed a DNA walker capable of autonomous behavior [53].

Detection of the small molecule, AMP, or the protein, lysozyme, was reported by Willner and co-workers using an aptamer-deoxyribozyme nanodevice [49]. The device consists of an aptamer sequence (region I) designed to target AMP or lysozyme. The aptamer recognition sequence is integrated with a deoxyribozyme sequence (region II) that in the presence of heme, will catalyze hydrogen peroxide reduction. The catalytic activity by the deoxyribozyme has been described as a horseradish peroxidase mimic. A second DNA strand that is complementary to portions of both the aptamer and deoxyribozyme regions is used to prevent catalytic activity except in the presence of the aptamer target and heme. Catalytic turnover of hydrogen peroxide was monitored colorimetrically using ABTS, or via chemiluminescence using luminol.

Other similar DNA molecular nanodevices employing amplification schemes have been reported by Willner’s group for detection of single stranded DNA [126], cocaine [92], and even larger objects including viruses [9, 125]. For more information, the reader is referred to the recent and thorough review of nucleic acid based molecular devices by Krishnan and Simmel [44].

## 3 Application and Substrate-Specific Algorithm Development

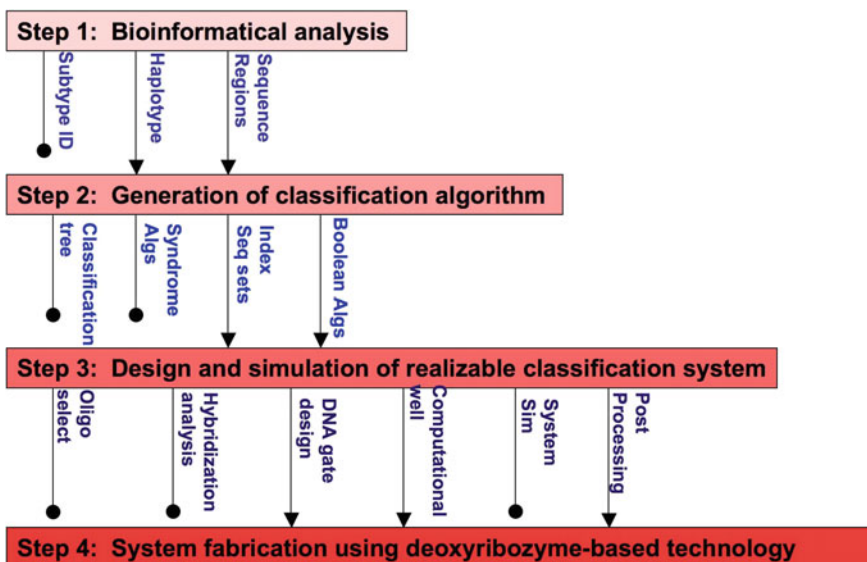
The integration of DNA computing principles with recent advances in DNA-based catalytic molecular probes and nanodevices holds great promise for revolutionizing the field of bioanalytics and computational biosensors. However, application and

substrate-specific algorithm development is key for integration within a platform viable for in-field detection.

Rapid development and deployment of application-specific microsensor detection systems requires algorithmic tools to map the biological detection challenge to the biosensor technology. Additionally computational biosensor-dependent technologies must incorporate algorithmic intelligence into the microsensor platform, enabling the execution of concurrent detection and classification tasks, such as systematic identification of nucleic acid sequences or polymorphisms that indicate the existence of microbial pathogens or disease-related genes in the presence of non-lethal agents.

We have outlined four general steps needed in the design of a computational biosensor (depicted in Fig. 9), where the first three steps focus on the development of computational algorithms for customizing the biological application to the sensor platform. Steps include: (1) Bioinformatics analysis to generate candidate probes; (2) Generation of probe-based classification algorithms; (3) Design and simulation of realizable biosensor and classification system; (4) System fabrication using computational biosensor-based substrate.

In this section we provide an overview of computational methods and algorithms used in the first three steps of the development process. We also provide an overview of methods used specifically for mapping biosensor applications to nucleic acid based computational sensor platforms.



**Fig. 9** Procedure for de novo generation of an application specific “intelligent” DNA sensor system



### ***3.1 Bioinformatics Analysis and Subsequence Selection Algorithms***

#### **3.1.1 Methods for Subsequence Selection for Candidate Probes**

Nucleic acid-based biosensor applications generally attempt to detect the presence of a subsequence target or set of subsequences that correspond to key components of a genome or transcription product. Identification of the most informative regions of a target that enable high detection of targets in a potentially noisy background and accurate classification of targets or variations in the target can be computationally challenging. Nucleic acid sensors and molecular beacons are viable options for genotyping, haplotyping and various gene and mRNA identification applications due to their relatively short target recognition site and their ability to discriminate polymorphisms in target sequences [17, 50, 67, 68, 77]. Various algorithmic methods such as those used in microarray design, haplotyping, and polymorphism detection have been explored for target identification and generation of candidate subsequences for use in nucleic acid based biosensors.

Subsequence selection is a multifaceted challenge requiring analysis and classification methods suited for the given biosensing application. Single nucleotide polymorphism (SNP) based applications have used information entropy to characterize and create initial sets of candidate SNPs for haplotyping and phenotype classification [47, 48]. Statistical and entropy based heuristics can yield subsequence groups with high correlation and minimal redundancy. Linear regression, Bayesian approaches, machine learning methods, genetic algorithms, and support vector machines are examples of methods used to search for the optimal SNP or subsequence subsets with maximal phenotype detection and classification performance [10, 30, 48].

#### **3.1.2 Information Theory Based Subsequence Selection for Molecular Beacon Probes**

Molecular beacons and catalytic molecular beacons have short target recognition regions. Our method for finding highly informative subsequences was based on the integration of sequence alignment and information theory algorithms. As an example application we consider the design of a deoxyribozyme-based computational sensor for identification and monitoring of H5N1, the causative agent of influenza (May et al. [65]). Surveillance plus antigenic and genetic analysis have been used to monitor known H5N1 variants and identify emerging sublineages [97]. Large-scale sequencing efforts, such as the St. Jude Influenza Genome Project [71], provide genetic information for the design of computational biosensors that can aid in tracking and characterizing regional strains of H5N1. Since the genomic sequence of influenza viruses is highly mutable [27, 72] given a set of subtypes or strains, it was challenging to find several highly conserved regions that sufficiently

distinguish the groups. In designing the FluChip diagnostic microarray, Mehlmann et al. manually inspected a phylogenetic tree of influenza sequences to identify clades that had sufficient sequence similarity [66]. These clades were then used in an automated probe extraction strategy. To encourage process uniformity and minimize design time, we developed extraction methods that minimize the manual intervention in the probe identification process.

Using 164 H5N1 hemagglutinin (HA) gene sequences (NCBI Influenza Virus Resource) from samples spanning five different geographic locales (Egypt, Hong Kong, Indonesia, Thailand, and Vietnam), we generated a multiple alignment and consensus sequence for each locale using MUSCLE [21]. We then created a multiple alignment of the five consensus sequences, in order to index the nucleotides across all sequences uniformly. To find regions of low variation within a locale, we used an information theory approach. We set a window size of  $k = 5$  nucleotides, and for each index, we calculated the  $k$ -mer (words of length  $k$ ) sequence entropy using the standard entropy equation [18]

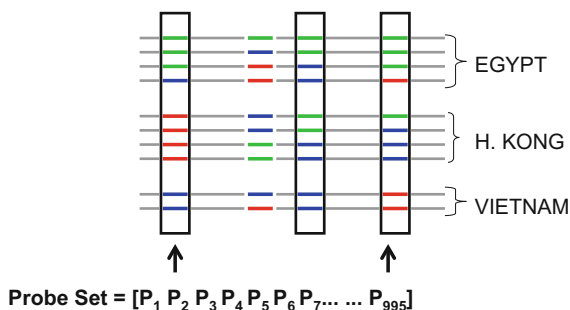
$$E_i = - \sum_{w \in \text{words}_i} p_i(w) \log_2 p_i(w)$$

where the sum is over all  $k$ -length words occurring at position  $i$  in the alignment, and  $p_i(w)$  is the proportion of words found at position  $i$  that are equal to  $w$ . Thus, positions with low  $k$ -mer sequence diversity—e.g., only one or two unique words occurring throughout the alignment at that position—will have low entropy, while positions that have high sequence diversity will have high entropy. We included only the ungapped words in the entropy calculation and introduce a uniform gap penalty to disfavor gapped positions in our calculation.

Once each alignment was annotated with the  $k$ -mer entropy per position, we extracted indices that had low sum-entropy spanning 15 nucleotides, the length of the target-binding region of the deoxyribozyme. Using an empirically determined entropic cutoff of  $E^{15} < 8$  bits, we collated all of the unique 15-nucleotide words at each low sum-entropy position; a percentile cutoff could also be employed. We reduced the candidate probe set by removing the most infrequent words, such that at least 90% of the individuals in the alignment contained one of the words in the reduced set. While using a larger  $k$ -mer would seem the most direct method, we found that large  $k$ -mers decreased the resolution of the algorithm and the resulting candidate probe set (Fig. 10).

The initial probe set was subjected to additional heuristic filters. Candidate target subsequences were: (1) Checked to ensure that they were not cross reactive with non-HA H5N1 gene sequence; (2) Annotated according to their maximum classification sensitivity over all regions/locales (defined as  $\frac{\text{\#individuals in a locale containing the word}}{\text{total \#individuals in a locale}}$ ) and a specificity measure equal to the entropy of the distribution of classification sensitivities. Candidate probes with sensitivity less than 0.75 and specificity scores less than 1 (maximum is  $-\log_2 \frac{1}{2}$ ) were removed. We subsequently performed a final filtering step to identify probe

**Fig. 10** Use of alignment and information algorithms for candidate probe selection for strain classification of regional H1N1 strains



sequences that had a high probability of structural incompatibility with the deoxyribozyme gate architecture (see Sect. 3.3). We applied our algorithmic pipeline to identify candidate probes for distinguishing influenza subtypes and were able to extract 392 candidate marker sequences on the HA gene spanning the five regional subtypes. Our results suggest that our method is robust to sequence groups with high variation.

## 3.2 Application and Substrate Specific Algorithm Generation

### 3.2.1 Classification Algorithms for Nucleic Acid Biosensor Applications

Nucleic acid based biodetection applications determine the presence or absence of a target using a subset of the target's genomic space. Depending on the sensor platform, the number of target probes incorporated in the system can vary in number and length. Consequently once an initial set of candidate probes are generated, further refinement of the target subsequence set may be required. Algorithmic or heuristic-based methods must generate a set of probes that when detected can be used to computationally classify the sample as one of multiple targets or a non-target, thus the probe selection and the sensor output classification algorithm are strongly correlated.

Use of genome wide association and polymorphism data to predict disease phenotypes or to identify drug resistant pathogens require the selection of optimal probe or SNP sets for maximal classification outcome. Statistical methods have been used to rank SNPs in order to optimize disease prediction outcome using single or ensemble classification algorithms [59]. Classification methods used include decision trees, logistic regression, support vector machines and Bayesian methods. In addition to statistical methods, regression methods, support vector machines, Mahalanobis distance, and machine learning methods have been used in

developing classification algorithms for various application areas including genotyping, haplotyping, and mortality rate prediction [30, 47, 84].

Many nucleic acid based detection applications rely on classification algorithms that can be implemented “off chip” using traditional silicon-based computing platforms to post-process, analyze and interpret the detection event. However, the ability of molecular beacons, deoxyribozymes in particular, to functionally operate as computational logic gates enables us to concurrently detect and classify probe sets. Ultimately this reduces post-processing cost and potentially reduces classification error due to signal loss. To take advantage of the integrated computing capability of computational nucleic acid molecules [39, 105], we explore new algorithmic methods that couple detection and classification while considering potential computational and design limitations of the molecular beacon.

### 3.2.2 Algorithmic Design of Intelligent Deoxyribozyme Sensors

In order to demonstrate the capabilities of the deoxyribozyme platform for computational biosensing, we used the 392 candidate marker sequences generated from the HA gene spanning three of the five regional H5N1 subtypes (Egypt, Hong Kong, Indonesia, Thailand, and Vietnam) to develop a prototype computational biosensor system [65]. Our system was designed de novo and was able to concurrently perform complex detection/classification tasks in the biological substrate, classifying a nucleic acid sample as belonging to one of the three regional subtypes based on the HA gene. The ability to compute differentiates this system from other molecular sensor systems. By increasing the complexity of the executed algorithm one can potentially increase the sensitivity and specificity of the biosensor, hence improving accuracy through computation.

We reduced the region classification problem to an experimentally tractable form: develop a deoxyribozyme-compliant algorithm that detects and classifies a candidate strain as originating from Egypt, Hong Kong, or Indonesia, using a reduced probe set of 75 candidate probes. The reduced probe set was selected from the candidate probe set based on their maximum classification sensitivity values. Using coding theoretic methods described in prior work [62–65] we formulated the classification and design problem as an inverse error control code (ECC), where we determine the optimal parity check matrix,  $H_{EgHkIn}^T$ , such that

$$H_{EgHkIn}^T * C_{Region_i} = S_{Region_i}$$

where  $C_{Region_i}$  is the set of all sample sequences from Region<sub>*i*</sub>, for  $i = \{Egypt, Hong Kong, Indonesia\}$ ; a sample sequence (a row in  $C_{Region_i}$ ) is a 75-bit binary sequence, where a 1 at location  $p$  indicates that probe sequence  $p$  is present in the sample sequence.  $S_{Region_i}$  is the syndrome matrix for Region<sub>*i*</sub> and contains the syndrome vector for each sequence in  $C_{Region_i}$ . The solution for the optimal  $H_{EgHkIn}^T$  concurrently maximized the inter-region hamming distance  $d_{hamming}(S_{Region_i},$

**Table 1** Deoxyribozyme compliant computational wells for detection and classification of H1N1 regional strains

<i>CompWell</i> <sub>1</sub>	$probe_{21} \oplus probe_{30} \oplus probe_{37} \oplus probe_{42} \oplus probe_{50} \oplus probe_{62} \oplus probe_{73}$
<i>CompWell</i> <sub>2</sub>	$probe_{27} \oplus probe_{32} \oplus probe_{33} \oplus probe_{36} \oplus probe_{42} \oplus probe_{43} \oplus probe_{47} \oplus probe_{74}$
<i>CompWell</i> <sub>3</sub>	$probe_{34} \oplus probe_{38} \oplus probe_{45} \oplus probe_{75}$

$S_{Region_j}$ ) of the regional syndrome matrix, minimized the intra-region hamming distance  $d_{hamming}(S_{Region_i}, S_{Region_i})$ , and maximized the number of zeros in  $H_{EgHkIn}^T$ . These constraints were designed to produce efficient classification algorithms using a minimal set of probes.

We used our coding theory based method to generate an ( $n = 75, k = 72$ ), deoxyribozyme compliant detection and classification algorithm. The algorithm resulted in a three column regional syndrome matrix, which corresponds to three computational wells, and used a genetic algorithm to search the space of feasible linear block codes that satisfied our design constraints. The resulting  $H_{EgHkIn}^T$  used 18 of the 75 probe sequences from the reduced set. Our detection and classification algorithm for each of our three computational wells (CompWell) was reported as follows ( $\oplus$  indicates exclusive-OR):

Table 1 lists the probe sequences used in each computational well. The algorithm produced three regionally distinct average syndrome patterns:

$$S_{Egypt}^{Avg} = (0, 0, 0) \quad S_{HongKong}^{Avg} = (0, 1, 1) \quad S_{Indonesia}^{Avg} = (1, 1, 1)$$

Based on the resulting syndrome patterns the computational biosensor system should be able to correctly distinguish the Egypt and Indonesian strains, but may not be as successful when distinguishing between the Hong Kong and Indonesian strains. A longer syndrome vector, hence additional computational wells, should result in a more accurate computational biosensor system. The algorithm in the computational wells can be used to design multiple deoxyribozyme gates to implement the prototype system.

### 3.3 Design of Nucleic-Acid Based Biosensor Substrates

The desired outcome of the algorithmic design process is the development of a realizable biosensor and classification system. A key to successful system fabrication is the design of structurally and functionally accurate molecular probes. Several computational tools, methods, and integrated pipelines for designing molecular beacons and various nucleic acid-based biosensor systems have been reported. Specific to molecular beacon probes, computational methods for ribozyme design using a random search method and partition function for calculating base-pair bindings was developed by Penchovsky and Breaker [74]. Building on

public RNA secondary structure calculation tools, Hall et al. developed methods for computationally designing aptazymes based on free energy profiles [28]. Additional methods for molecular design and simulation of nucleic acid-based circuits composed of multiple interacting molecular probes have also been explored [31, 75, 85]. Drawing on existing tools and thermodynamic-based methods, we developed computational methods and metrics for evaluating candidate deoxyribozyme gates used to implement in vitro computational biosensors [64, 65].

### 3.3.1 Designing Realizable Deoxyribozyme Gates

As in all probe design steps, the influence of the target sequence on the probe structure can result in undesirable secondary structure. For deoxyribozymes this is particularly critical as the change in conformation is key to the enzymatic and Boolean functionality of the computational gate. Using the E6 catalytic core sequence [54] and the probe sequences generated by our algorithmic pipeline, we screened candidate probes to ensure that the target loop sequences had minimal complementarity to the catalytic portion of the deoxyribozyme, as well as minimal complementarity to other target sequences when multiple loops exist on the same gate.

Based on the form of the equations generated for the three computational wells, we required a series of probe sequences to be combined and incorporated into XOR functions. At the time of our prototype design there did not exist a deoxyribozyme gate that implemented the XOR function, therefore in lieu of an XOR gate we designed two complementary ANDNOT gates with the loops reversed to perform the same logic ( $p \oplus q \Rightarrow (p \wedge \neg q) \vee (\neg p \wedge q)$ ). To determine which loops were pairwise compatible, we optimized selection based on string edit distance, calculated using Smith-Waterman local alignment [121], and hybridization energy calculated using the UNAFold software package. Using a graph theoretic approach we found the optimal probe-pairing using an implementation of Edmonds' Algorithm [22] written by Eppstein [25]. Unpaired probe sequences were incorporated into YES gates. We designed two complementary ANDNOT gates for each paired probe sequence using the E6 deoxyribozyme catalytic sequence. For each gate, we generated the minimum free energy structure using UNAFold and verified its structural similarity to the reported ANDNOT gate structure, however minor structural variations were permitted in the target recognition regions.

### 3.3.2 Computational Modeling and Simulation of Deoxyribozyme Gates

Due to the high cost and long turnaround times associated with experimental biosensor measurements, computational, or "in silico", prediction of biosensor performance is highly desirable prior to system fabrication. Computational predictions can be used to quickly and inexpensively screen proposed new biosensor

designs and provide molecular-level rationale for experimentally observed phenomena. We have demonstrated the ability to computationally predict experimentally observed deoxyribozyme system performance using DNA hybridization thermodynamics information [64].

Ideally the use of molecular simulation could serve as a plausible method for investigation of the atomistic details of deoxyribozyme gate function using CHARMM [55] and AMBER [20] force fields. Both have been developed to include parameters for nucleic acid molecules and have been shown to be reasonably reliable models in representing the structural and dynamic properties of nucleic acids. These force field functional forms are included in massively parallel molecular dynamics (MD) software packages such as LAMMPS (Large-scale atomic/molecular massively parallel simulator; *lammps.sandia.gov*) and have been used in many biomolecular simulation studies. Unfortunately, complete observation of real gate dynamics would require orders of magnitude more simulation time than was feasible even given high performance computing resources. However valuable information can be extracted from short molecular scale simulations including insight into important 3D structural effects in the hybridization thermodynamics predictions.

Using DNA hybridization thermodynamics we can predict properties of candidate gates by accumulating contributions from base pair matches and nearest neighbor interactions. Such calculations are possible using assigned parameters that are fit from extensive experimental measurements and methods pioneered by the SantaLucia lab that has enabled nucleic acid hybridization thermodynamics prediction [76, 88]. Thermodynamics based methods developed by SantaLucia and colleagues have been implemented in two comparable software packages: the freely available HyTher web server and the commercial Oligonucleotide Modeling Platform (OMP) sold by DNA Software, Inc. Our initial prediction pipeline used HyTher, but ultimately we used OMP to develop a pipeline for deoxyribozyme gate and computational well design [65].

Using OMP to compute DNA hybridization thermodynamics, we developed a Python pipeline for computing ribozyme performance. Experimental conditions can serve as inputs to OMP, including primary sequences of the oligonucleotides, solution temperature, ionic concentrations, and buffer conditions. The output of OMP includes: melting temperature ( $T_m$ ), hybridization free energy change ( $\Delta G$ ), and activated deoxyribozyme concentration ((active)). Although the three thermodynamic quantities correlated well with the experiments, we found that (active), the concentration of activated deoxyribozymes, produces the best agreement. We used the data to relate experimentally-measured changes in fluorescence with respect to time to the concentration of activated ribozymes as follows:

$$\frac{\delta \text{fluorescence}}{\delta t} = k[\text{substrate}][\text{active}]$$

where  $k$  is a rate constant and (substrate) is the concentration of substrate molecules cleaved by the deoxyribozyme gate. We set the concentration threshold for

activated ribozyme to a 1 nM as a cutoff to distinguish between pairs producing a strong fluorescence signal (on or logic 1), and those pairs that produce no distinguishable signal (off or logic 0). Using this approach we were able to simulate individual gate performance and simulate system performance for our prototype system.

We used the algorithmic methods described in Sects. 3.1 through 3.3 to develop and test a simple prototype system and demonstrate the feasibility of using in silico design to build a deoxyribozyme-based computational biosensor. We devised a reduced ( $N = 52$ ,  $K = 47$ ) coding-based algorithm for subtype classification resulting in a three-bit classification algorithm:

$CompWell_1 =$	$probe_{23} \oplus probe_{48}$
$CompWell_2 =$	$probe_4 \oplus probe_{23} \oplus probe_{24} \oplus probe_{34}$
	$probe_{43} \oplus probe_{50}$
$CompWell_3 =$	$probe_{20} \oplus probe_{51}$

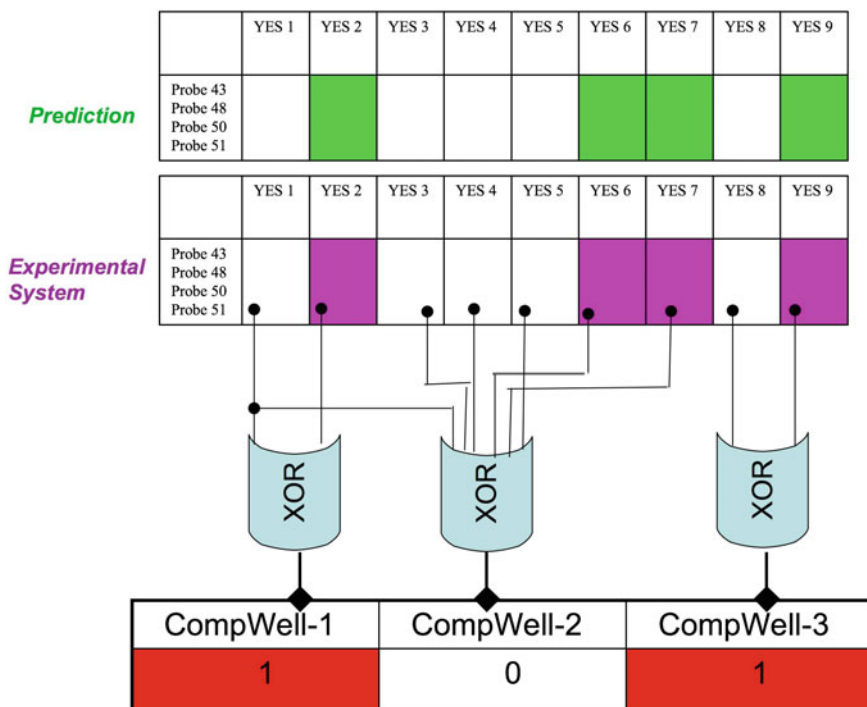
where CompWell represents the computational well, which is realized by combining the discretized binary output of the microplate wells containing each of the YES-gates in the classification algorithm. The YES-gate target probe sequences are listed in Table 2. We used our OMP-based pipeline to design, simulate, and generate a realizable.

YES-gate only system and a YES-gate/AND-NOT gate implementation. The AND-NOT gate implementation uses two AND-NOT gates to implement the exclusive-OR ( $\oplus$ ) function, which we performed in our post-processing step. Our computational biosensor system was implemented, in vitro and we tested the system's performance in H5N1 regional subtype classification using our fluorescence-based detection platform and synthesized input DNA sequences [64, 65]. The results of the experimental tests for the YES-gate configuration and the predicted system output is shown in Fig. 11. The binary microwell values were combined to compute the value of each computational well for a sample sequence

**Table 2** Probe sequences for region identification for the (47, 52) code

Probe Id	Sequence
4	GATCCTCCTTTTTTA
20	GCTATATCAAAACCC
23	CAGACAAGGCTATAT
24	AAGGCTATATCAAAA
34	ACAACATACACCCTC
43	CCAATCATGATGCCT
48	TACTAGACCCAAAGT
50	GCTACTAGACCCAAA
51	CTAGACCCAAAGTAA





**Fig. 11** Schematic for calculation of computational well values from experimental subwells. The probe sequences, representing the genomic components of the viral strain, are added to each well containing a specified YES-deoxyribozyme gate. The predicted (*green*) and experimentally determined (*purple*) outputs concur and are used to calculate the binary values of the computational wells. The three-bit binary syndrome pattern indicates viral strain association for the probe sequences

**Table 3** Syndrome classification patterns for subtype classification

Subtype Class	CompWell <sub>1</sub>	CompWell <sub>2</sub>	CompWell <sub>3</sub>
Egypt	0	1	0
Hong Kong	1	0	1
Indonesia	1	1	1

set as shown. The binary syndrome pattern of the computational wells determined which H5N1 strain the input sequence set belonged (Table 3). Using the binary patterns, we correctly identified the subtype of 97.3% of the simulated sample sequences.

## 4 DNA Detection Platform Development

Thus far we have discussed two key capabilities required for generating advanced computationally intelligent nucleic acid-based microsensor platforms that are reliable, deployable, and limit reliance on off-platform computational post-processing. These two capabilities are use of catalytic nucleic acid-based molecular probes and machines (Sect. 2.2), and application and substrate-specific algorithm development (Sect. 3). We finish this chapter discussing a third key capability, the interface between the biological recognition event and the microsensor, and the transduction of information between the two.

Several methodologies and technologies have been successfully employed for detection of a wide variety of molecules via nucleic acid-based recognition elements. These methodologies innately possess advantages and disadvantages. The choice of which detection modality one applies can therefore have a profound impact on the utility of the sensor.

### 4.1 Detection Modalities

Biosensing detection modalities are broadly characterized into three signal transduction systems: optical, gravimetric, and electrochemical. These three techniques and their characteristics enabling the design and implementation of computational biosensors will be briefly discussed.

#### 4.1.1 Optical Detection

Optical nucleic acid detection methods encompass fluorescence, surface plasmon resonance (SPR), surface enhanced Raman spectroscopy (SERS), luminescence, colorimetric, and similar techniques for transducing the biorecognition event. The most widely used optical biodetection modality is fluorescence. This method relies on the use of fluorescent labels, or the generation of a fluorescent signal, for detection of the biorecognition event. This approach often yields simple to interpret, semi-quantitative data with very low detection limit. Use of fluorophores of differing colors also allows for multianalyte biodetection. However, fluorescence-based detection requires an excitation light source, optics to filter the excitation wavelength from the emitted fluorescence wavelength, and a photo detector sensitive to the emitted fluorescence. These components are typically integrated together into bench-top laboratory equipment (e.g. fluorescence microscope, fluorometer) that is expensive, not portable, and requires training and post-signal processing.

Developments in light emitting diode (LED) technology have led to advances in portable fluorescence-based DNA detection technologies. For example, Wand and

Trau recently reported a generic DNA bioassay using an LED/photodiode system that fits into an aluminum briefcase and can detect PCR products at a concentration as low as 97 fmol [116]. We anticipate continued progress in portable fluorescence-based DNA detection as these technologies are further refined.

Surface plasmon resonance (SPR) is a technique that monitors minute changes in refractive index occurring at the interface between materials with differing refractive indices. Commonly, a prism is coated with a thin gold layer that is functionalized with single-stranded DNA probes. A solution containing the analyte flows over this surface. Infrared or visible light at a single wavelength near the surface plasmon resonance condition is directed incident to the gold surface at the resonance angle, resulting in a minimum of the reflected light intensity. Upon binding of the target DNA to the probe surface, the surface resonance condition is altered, resulting in a change in the refractive index [16, 73]. Thus, a significant advantage of SPR-based DNA detection is that no label is required. However, false positives due to non-specific binding of non-target DNA and other components in the analyte solution are significant challenges. Non-specific binding is a common concern with nearly all label-free based biodetection systems.

Surface enhanced Raman spectroscopy (SERS) is a surface technique that has been used for DNA detection in which rough metallic nanostructures are used to substantially increase very weak Raman scattering. When molecules absorb onto the roughened nanostructures, incident light excites surface plasmons, or forms charge-transfer complexes, resulting in a significant enhancement in excitation intensity which can be as high as  $10^{9-11}$  [45]. This allows for detection of molecules on the SERS substrate that would normally be undetectable using Raman alone. This phenomenon has resulted in the burgeoning field of SERS-based detection of molecules of biomedical interest and the development of many SERS probes/substrates [90]. For example, the Vo-Dinh group has used SERS for detection of DNA hybridization to targets specific for breast cancer and the HIV-1 viral gene [114]. Nanomaterials modified with aptamers have also been widely used for SERS-based biosensing [117]. Although sensitive and specific, this technique can still require target amplification (typically PCR) as the detection limits are not sufficiently low for detection of DNA levels in real-world samples. Also, although portable Raman spectrometers are currently commercially available, these devices can be challenging to operate in the field, must be calibrated to the specific target, can be confounded by complex sample solutions, and are expensive [61].

We already discussed an aptamer-DNAzyme molecular nanodevice that produces a colorimetric signal, or chemiluminescent signal, in the presence of the target analyte and hemin [49], Sect. 2.2.5). Optical and chemiluminescence approaches are advantages in that the results can typically be read by eye, thus they do not require laboratory equipment or significant training to read. However, the detection limit of these systems is often poor as sufficient signal must be generated to be visually discernable. Better detection limits are possible using spectrometers for absorbance/light intensity measurements. However, using such a device negates the advantage of using a detection system that can be read by eye. In the case of the aptamer-DNAzyme system mentioned above, in which the optical signal is

catalytically amplified beyond the 1:1 target to signal ratio, the detection limit of AMP was still only 4  $\mu\text{M}$ . Thus, optical based systems for DNA detection have not yet found significant use outside the laboratory.

#### 4.1.2 Gravimetric Detection

Gravimetric detection modalities rely on measuring small changes in mass upon target binding at the sensor surface. Quartz crystal microbalances (QCMs) are the most wide reported gravimetric biodetection devices. These devices are composed of a quartz wafer sandwiched between two electrodes (typically gold). One of the electrodes is then functionalized with a thin film containing the capture probe. An oscillating electric field is applied between the two electrodes, resulting in a mechanical resonance in the crystal that is very sensitive to mass [15]. The relation between the electrical and mechanical oscillations to the mass on the crystal is described by the Sauerbrey equation:

$$\Delta f = \frac{-2\Delta m n f_0^2}{A\sqrt{\mu\rho}}$$

where  $\Delta f$  is the oscillation frequency,  $\Delta m$  is the mass change,  $A$  is the piezo-electrically active area of the crystal, and  $\mu$  and  $\rho$  are the shear modulus and density of the crystal, respectively.

Such sensors are capable of detecting nanogram changes in mass, in real-time. DNA-based QCM detectors measure the increase in mass due to hybridization of target to the capture probe, resulting in a decrease in resonance frequency. Another significant advantage to this system is the target does not need to be labeled. However, such devices suffer from the same non-specific binding limitation, resulting in false positive measurements, common to most label free sensors [36].

Surface acoustic wave (SAW) devices operate in a similar fashion. A thin piezoelectric crystal is modified with interdigitated electrodes. An oscillating electric field generates a surface acoustic wave, or a bulk acoustic wave, that is launched across the crystal. The crystal surface is functionalized with the capture probe. As the acoustic energy is confined to the thin surface region of the crystal substrate, binding of target molecules to the surface induces perturbations to the wave, altering its amplitude and velocity. Thus, the surface wave device can operate as a mass or viscosity sensor [13, 19].

Most reports of detection of DNA using these devices require PCR amplification. However, some sensors have been reported that detect target DNA in non-amplified samples. For example, Karamollaoglu et al. reported detection of a genetic insert in tobacco plants from PCR amplified DNA samples, and from samples fragmented by digestion and ultrasonication, using a QCM-based DNA biosensor [40].

The surface chemistry used to modify the crystal surface with the DNA capture probe can have a profound impact on the specificity, detection limit and dynamic range of the system. Several reports comparing gold-thiol to thiol-dextran [57], gold-thiol to biotin-streptavidin [58], and gold-thiol to amine-glutaraldehyde [40], all show statistically significant difference in sensitivity and dynamic ranges. Typically, the best results are obtained from the gold-thiol surface as these molecules assemble more densely (increased probe density, reduced sites for non-specific binding) and are comparatively lower in mass (surface chemistry does not contribute as significantly to frequency changes), allowing for more sensitive discrimination target binding induced changes in surface mass.

These devices are amenable for use outside the lab, with portable versions reported [4, 11]. However, they can be high cost, are non-trivial to construct, and require sensitive measurements and data post-processing. This makes production of simple to operate and low cost gravimetric-based DNA detection devices challenging.

### 4.1.3 Electrochemical Detection

Electrochemical detection methods utilize changes in reduction/oxidation (redox) reactions and electron transfer (including conductivity) properties associated with binding of the target molecule to a conducting substrate. Substrates are most often gold, platinum, or carbon. Conducting polymers and other metallic materials have also been successfully employed for electrochemical biosensing. The substrates can be either planar solids, or three dimensional with high surface area and porosity, or nanoscale in dimension. DNA capture probes are immobilized onto the conducting surface, and hybridization of target DNA to the capture probe is monitored directly or indirectly.

Direct methods typically measure changes in current directly from redox reactions with nucleic acid bases in the target DNA, most notably, the direct oxidation of guanine [38] which can be further amplified by using a catalytic mediator such as  $(\text{Ru}(\text{bpy})_3)^{2+}$  [95]. Background signal from the capture probe can be significantly reduced by replacing guanines with uracil, a nucleobase that also binds to cytosine, but oxidizes at a potential much higher than guanine. Thus, only when the target strand hybridizes to the surface is a current from guanine oxidation obtained. Direct detection can also occur by measuring changes in the electrode/electrolyte interface properties upon target hybridization typically by electrochemical impedance spectroscopy (EIS) or conductivity, although these methods are generally not very sensitive without addition of a redox mediator probe (e.g. ferricyanide, methylene blue).

Indirect methods make use of redox or electroactive mediators that can intercalate with hybridized double-stranded DNA. Examples of electroactive double stranded DNA intercalators include ethidium bromide and daunomycin. Alternatively, and most commonly, indirect methods employ labels such as enzymes, metallic or semiconductor nanoparticles, dendrimers, liposomes, carbon nanotubes,

nanowires, etc., which enable a wide variety of electrochemical detection schemes and also significantly amplify the measured signal output [79]. In these schemes it is the label that is detected, or provides the electrochemical signal, indicating hybridization has occurred.

Electrochemical detection provides significant advantages over other common detection and signal transduction methods. Specifically, it can be very sensitive, with detection limits in the pico- and fempto- molar range, detecting DNA in non PCR amplified samples [109]. Also, electrochemical detection can be rapid and provide greater specificity and sensitivity over optical detection methods as interfering background fluorescence does not adversely affect the electrochemical signal. Unlike fluorescent or biological molecules which bleach or degrade with time, electroactive molecules are typically more stable and insensitive to the environment.

Further, multianalyte biodetection is often simpler with electrochemical schemes by using either multiple electroactive labels with differing redox potentials, or an array of individually addressable electrodes that are selectively modified with differing capture probes. For example, we reported the first simultaneous electrochemical detection of DNA and protein on a single platform by selectively immobilizing single stranded DNA probes on five individually addressable electrodes, and immobilizing antibody probes on another four individually addressable electrodes, in a nine element array [29].

Importantly, electrochemical detection systems are also much simpler to miniaturize as integrated circuit fabrication methods can be leveraged to produce the miniaturized components. These components have relatively low power requirements and low mass, making them idea for potable, handheld, or leave behind sensor modules. In another example, we recently reported the development of an inexpensive and low power electrochemical DNA detector that contained an array of nine individually addressable gold electrodes [23]. Using electroaddressable aryl diazonium chemistry, the electrodes were selectively patterned to detect DNA sequences specific to breast cancer, colorectal cancer, and provide information regarding non-specific binding on the array (negative control). The device also contained a resistive platinum heating element that permitted melting of double stranded DNA sample solution, allowing for multianalyte detection from double stranded DNA target on the array. Following detection, the array surface could be renewed via high temperature stripping utilizing the on-chip resistive heating element.

Such inexpensive and simple multianalyte DNA sensors show great potential for integration with nucleic acid probes that incorporate structure and catalytic function, in combination with DNA information processing techniques, resulting in intelligent computational biosensor platforms that can truly revolutionize DNA biodetection.

**Acknowledgements** We would like to thank our colleagues who collaborated on our original computational biosensor work: P. Dolan, P. Crozier, M. Lee, M. Manginell, and R. Polsky.

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

## References

1. Adar R, Benenson Y, Linshiz G, Rosner A, Tishby N, Shapiro E (2004) Stochastic computing with biomolecular automata. *Proc Natl Acad Sci USA* 101(27):9960–9965
2. Adleman LM (1994) Molecular computation of solutions to combinatorial problems. *Science* 266(5187):1021–1024
3. Adleman L, Rothemund P, Roweis S, Winfree E (1999) On applying molecular computation to the data encryption standard. *J Comput Biol* 6(1):53–63
4. Andle JC, Vetelino JF (1994) Acoustic wave biosensors. *Sens Actuators A* 44(3):167–176
5. Arugula MA, Zhang Y, Simonian AL (2014) Biosensors as 21st century technology for detecting genetically modified organisms in food and feed. *Anal Chem* 86(1):119–129
6. Baron R, Lioubashevski O, Katz E, Niazov T, Willner I (2006) Elementary arithmetic operations by enzymes: a model for metabolic pathway based computing. *Angewandte Chemie-International Edition* 45(10):1572–1576
7. Baron R, Lioubashevski O, Katz E, Niazov T, Willner I (2006) Logic gates and elementary computing by enzymes. *J Phys Chem A* 110(27):8548–8553
8. Benenson Y, Gil B, Ben-Dor U, Adar R, Shapiro E (2004) An autonomous molecular computer for logical control of gene expression. *Nature* 429(6990):423–429
9. Beissenhirtz MK, Elnathan R, Weizmann Y, Willner I (2007) The aggregation of Au nanoparticles by an autonomous DNA machine detects viruses. *Small* 3(3):375–397
10. Bevilaqua A, Rodrigues F (2011) L. do Amaral. SNPs classification: building biological high-level knowledge using genetic algorithms. *Integrated computing technology communications in computer and information*. *Science* 165:50–58
11. Bisoffi M, Hjelle B, Brown DC, Branch DW, Edwards TL, Brozik SM, Bondu-Hawkins VS, Larson RS (2008) Detection of viral bioagents using a shear horizontal surface acoustic wave biosensor. *Biosens Bioelectron* 23(9):1397–1403
12. Braich RS, Chelyapov N, Johnson C, Rothemund PWK, Adleman L (2002) Solution of a 20-variable 3-SAT problem on a DNA computer. *Science* 296(5567):499–502
13. Branch DW, Brozik SM (2004) Low-level detection of *Bacillus anthracis* simulant using love-wave biosensors on 36°YX LiTaO<sub>3</sub>. *Biosens Bioelectron* 19(8):849–859
14. Breaker RR (1997) DNA enzymes. *Nat Biotechnol* 15(5):427–431
15. Bunde RL, Jarvi EJ, Rosentreter JJ (1998) Piezoelectric quartz crystal biosensors. *Talanta* 46(6):1223–1236
16. Campbell CT, Kim G (2007) SPR microscopy and its applications to high-throughput analyses of biomolecular binding events and their kinetics. *Biomaterials* 28(15):2380–2392
17. Carter J, Balaraman V, Kucharski C, Fraser T, Fraser M (2013) A novel dengue virus detection method that couples DNAAzyme and gold nanoparticle approaches. *Virol J* 10:201
18. Cover TM, Thomas JA (1991) *Elements of information theory*. Wiley, New York
19. Du J, Harding GL, Ogilvy JA, Dencher PR, Lake M (1996) A study of love-wave acoustic sensors. *Sens Actuators, A* 56(3):211–219
20. Duan Y, Wu C, Chowdhury S, Lee MC, Xiong G, Zhang W, Yang R, Cieplak P, Luo R, Lee T, Caldwell J, Wang J, Kollman P (2003) A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *J Comput Chem* 24(16):1999–2012

21. Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32(5):1792–1797
22. Edmonds J (1965) Paths, trees and flowers. *Can J Math* 17:449–467
23. Edwards TL, Harper JC, Polsky R, Lopez DM, Wheeler DR, Allen AC, Brozik SM (2011) A parallel microfluidic channel fixture fabricated using laser ablated plastic laminates for electrochemical and chemiluminescent biodetection of DNA. *Biomicrofluidics* 5:044115
24. Ellington AD, Szostak JW (1990) In vitro selection of RNA molecules that bind specific ligands. *Nature* 346(6287):818–822
25. Eppstein D (2003) Maximum cardinality matching in general graphs
26. Fan CH, Plaxco KW, Heeger AJ (2003) Electrochemical interrogation of conformational changes as a reagentless method for the sequence-specific detection of DNA. *Proc Natl Acad Sci USA* 100(16):9134–9137
27. Ghedin E, Sengamaly NA, Shumway M, Zaborsky J, Feldblyum T, Subbu V, Spiro DJ, Sitz J, Koo H, Bolotov P, Dernovoy D, Tatusova T, Bao Y, St George K, Taylor J, Lipman DJ, Fraser CM, Taubenberger JK, Salzberg SL (2005) Large-scale sequencing of human influenza reveals the dynamic nature of viral genome evolution. *Nature* 437(7062):1162–1166
28. Hall B, Hesselberth J, Ellington A (2007) Computational selection of nucleic acid biosensors via a slip structure model. *Biosens Bioelectron* 22(9):1939–1947
29. Harper JC, Polsky R, Wheeler DW, Dirk SM, Brozik SM (2007) Selective immobilization of DNA and antibody probes on electrode arrays: Simultaneous electrochemical detection of DNA and protein on a single platform. *Langmuir* 23(16):8285–8287
30. He J, Zelikovsky A (2007) Informative SNP selection methods based on SNP prediction. *IEEE Trans Nanobiosci* 6(1):60–67
31. Hockenberry AJ, Jewett MC (2012) Synthetic in vitro circuits. *Curr Opin Chem Biol* 16(3–4):253–259
32. Hoheisel JD (2006) Microarray technology: Beyond transcript profiling and genotype analysis. *Nat Rev Genet* 7(3):200–210
33. Horejsh D, Martini F, Poccia F, Ippolito G, Di Caro A, Capobianchi MR (2005) A molecular beacon, bead-based assay for the detection of nucleic acids by flow cytometry. *Nucleic Acids Res* 33(2):e13
34. Hutchins CJ, Rathjen PD, Forster AC, Symons RH (1986) Self-cleavage of plus and minus RNA transcripts of avocado sunblotch viroid. *Nucleic Acids Res* 14(9):3627–3640
35. Iliuk AB, Hu L, Tao WA (2011) Aptamer in bioanalytical applications. *Anal Chem* 83(12):4440–4452
36. Janshoff A, Galla HJ, Steinem C (2000) Piezoelectric mass-sensing devices as biosensors—An alternative to optical biosensors? *Angewandte Chemie-International Edition* 39(22):4004–4032
37. Jhaveri S, Rajendran M, Ellington AD (2000) In vitro selection of signaling aptamers. *Nat Biotechnol* 18(12):1293–1297
38. Johnston DH, Glasgow KC, Thorp HH (1995) Electrochemical measurement of solvent accessibility of nucleobases using electron transfer between DNA and metal complexes. *J Am Chem Soc* 117(35):8933–8938
39. Jung C, Ellington AD (2014) Diagnostic applications of nucleic acid circuits. *Acc Chem Res* 47(6):1825–1835
40. Karamollaoglu I, Oktem HA, Mutlu M (2009) QCM-based DNA biosensor for detection of genetically modified organisms (GMOs). *Biochem Eng J* 44(2–3):142–150
41. Kari L, Kari J, Landweber LF (1999) Reversible molecular computation in ciliates. In: Karhumaki J, Maurer H, Paun G, Rozenberg G (eds) *Jewels are Forever. Contributions on theoretical computer science in honor of Arto Salomaa*. Springer, Berlin, pp 353–363
42. Katz E, Privman V (2010) Enzyme-based logic systems for information processing. *Chem Soc Rev* 39(5):1835–1857



43. Kawde AN, Wang J (2004) Amplified electrical transduction of DNA hybridization based on polymeric beads loaded with multiple gold nanoparticle tags. *Electroanalysis* 16 (1–2):101–107
44. Krishnan Y, Simmel FC (2011) Nucleic acid based molecular devices. *Angewandte Chemie-International Edition* 50(14):3124–3156
45. Kudelski A (2008) Analytical applications of Raman spectroscopy. *Talanta* 76(1):1–8
46. Landweber LF, Kari L (1999) The evolution of cellular computing: nature's solution to a computational problem. *Biosystems* 52(1–3):3–13
47. Long N, Gianola D, Rosa G, Weigel K, Avendano S (2009) Comparison of classification methods for detecting associations between SNPs and chick mortality. *Gen Sel Evol* 41:18
48. Li X, Liao B, Cai L, Cao Z, Zhu W (2013) Informative SNPs Selection Based on Two-Locus and Multilocus Linkage Disequilibrium: Criteria of Max-Correlation and Min-Redundancy. *IEEE/ACM Trans Comput Biol Bioinf* 10(3):688–695
49. Li D, Shlyahovsky B, Elbaz J, Willner I (2007) Amplified analysis of low-molecular-weight substrates or proteins by the self-assembly of DNAzyme-aptamer conjugates. *J Am Chem Soc* 129(18):5804–5805
50. Lin SY, Probert W, Lo M, Desmond E (2004) Rapid detection of isoniazid and rifampin resistance mutations in *Mycobacterium tuberculosis* complex from cultures or smear-positive sputa by use of molecular beacons. *J Clin Microbiol* 42(9):4204–4208
51. Liu X, Farmerie W, Schuster S, Tan W (2000) Molecular beacons for DNA biosensors with micrometer to submicrometer dimensions. *Anal Biochem* 283:56–63
52. Liu Y, Tuleouva N, Ramanculov E, Revzin A (2010) Aptamer-based electrochemical biosensor for interferon gamma detection. *Anal Chem* 82(19):8131–8136
53. Lund K, Manzo A, Dabby N, Michelotti N, Johnson-Buck A, Nangreave J, Taylor S, Pei R, Stojanovic M, Walter N, Winfree E, Yan H (2010) Molecular robots guided by prescriptive landscapes. *Nature* 465:206–210
54. Macdonald J, Li Y, Sutovic M, Lederman H, Pendri K, Lu WH, Andrews BL, Stefanovic D, Stojanovic MN (2006) Medium scale integration of molecular logic gates in an automaton. *Nano Lett* 6(11):2598–2603
55. MacKerell AD Jr, Banavali N, Foloppe N (2001) Development and current status of the CHARMM force field for nucleic acids. *Biopolymers* 56(4):257–265
56. Maley CC (1998) DNA computation: theory, practice, and prospects. *Evol Comput* 6 (3):201–229
57. Mannelli I, Minunni M, Tombelli S, Mascini M (2003) Quartz crystal microbalance (QCM) affinity biosensor for genetically modified organisms (GMOs) detection. *Biosens Bioelectron* 18(2–3):129–140
58. Mannelli I, Minunni M, Tombelli S, Wang R, Spiriti MM, Mascini M (2005) Direct immobilization of DNA probes for the development of affinity biosensors. *Bioelectrochemistry* 66(1–2):129–138
59. Manor O, Segal E (2013) Disease Risk Using Bootstrap Ranking and Classification Algorithms. *PLoS Comput Biol* 9(8):e1003200
60. Margulies D, Hamilton AD (2009) Digital analysis of protein properties by an ensemble of DNA quadruplexes. *J Am Chem Soc* 131(26):9142–1943
61. Markert H, Ring J, Campbell N, Grates K (2011) A comparison of four commercially available portable Raman spectrometers. National Forensic Science Technology Center. [http://www.nfstc.org/?dl\\_id=214](http://www.nfstc.org/?dl_id=214)
62. May E, Vouk M, Bitzer D, Rosnick D (2004) Coding theory based models for protein translation initiation in prokaryotic organisms. *BioSyst J* 76(1–3):249–260
63. May E, Vouk M, Bitzer D (2006) An Error-Control Coding Model For Classification of *Escherichia coli* K-12 ribosome binding sites. *IEEE EMB Mag* 25(1):90–97
64. May EE, Dolan PL, Crozier PS, Brozik SM, Manginell M (2008) Towards de novo design of deoxyribozyme biosensors for GMO detection. *IEEE Sens J* 8(6):1011–1019

65. May E, Lee M, Dolan P, Crozier P, Brozik S, Manginell M (2008b) Computational sensing and in vitro classification of GMOs and biomolecular events. In: Proceedings of the 26th army science conference
66. Mehlmann M, Dawson ED, Townsend MB, Smagala JA, Moore CL, Smith CB, Cox NJ, Kuchta RD, Rowlen KL (2006) Robust sequence selection method used to develop the FluChip diagnostic microarray for influenza virus. *J Clin Microbiol* 44(8):2857–2862
67. Mhlanga M, Malmberg L (2001) Using molecular beacons to detect single-nucleotide polymorphisms with real-time PCR. *Methods* 25(4):463–471
68. Monroy-Contreras R, Vaca L (2011) Molecular beacons: powerful tools for imaging RNA in living cells. *J Nucleic Acids Article ID* 741723, 15 pp
69. Motornov M, Zhou J, Pita M, Gopishetty V, Tokarev I, Katz E, Minko S (2008) “Chemical transformers” from nanoparticle ensembles operated with logic. *Nano Lett* 8(9):2993–2997
70. Motornov M, Zhou J, Pita M, Tokarev I, Gopishetty V, Katz E, Minko S (2009) An integrated multifunctional nanosystem from command nanoparticles and enzymes. *Small* 5(7):817–820
71. Obenauer JC, Denson J, Mehta PK, Su X, Mukatira S, Finkelstein DB, Xu X, Wang J, Ma J, Fan Y, Rakestraw KM, Webster RG, Hoffmann E, Krauss S, Zheng J, Zhang Z, Naeye CW (2006) Large-scale sequence analysis of avian influenza isolates. *Science* 311(5767):1576–1580
72. Palese P, Young JF (1982) Variation of influenza A, B, and C viruses. *Science* 215(4539):1468–1474
73. Pattnaik P (2005) Surface plasmon resonance—applications in understanding receptor-ligand interaction. *Appl Biochem Biotechnol* 126(2):79–92
74. Penchovsky R, Breaker RR (2005) Computational design and experimental validation of oligonucleotide-sensing allosteric ribozymes. *Nat Biotechnol* 23(11):1424–1433
75. Penchovsky R (2012) Engineering integrated digital circuits with allosteric ribozymes for scaling up molecular computation and diagnostics. *ACS Synthet Biol* 1(10):471–482
76. Peyret N, Seneviratne PA, Allawi HT, SantaLucia J Jr (1999) Nearest-neighbor thermodynamics and NMR of DNA sequences with internal AA, CC, GG, and TT mismatches. *Biochemistry* 38(12):3468–3477
77. Piatek A, Tyagi S, Pol A, Telenti A, Miller L, Kramer F, Alland D (1998) Molecular beacon sequence analysis for detecting drug resistance in *Mycobacterium tuberculosis*. *Nat Biotechnol* 16(4):359–363
78. Pita M, Zhou J, Manesh KM, Halamek J, Katz E, Wang J (2009) Enzyme logic gates for assessing physiological conditions during an injury: towards digital sensors and actuators. *Sens Actuators B-Chem* 139(2):631–636
79. Polsky R, Harper JC, Brozik SM (2012) Nanomaterial-based electrochemical DNA detection. In: Ozsoz M (ed) *Electrochemical DNA Biosensors*. Pan Stanford Publishing, Singapore, pp 427–480
80. Prody GA, Bakos JT, Buzayan JM, Schneider IR, Bruening G (1986) Autolytic processing of dimeric plant-virus satellite RNA. *Science* 231(4745):1577–1580
81. Privman M, Tam TK, Bocharova V, Halamek J, Wang J, Katz E (2011) Responsive interface switchable by logically processed physiological signals: Toward “smart” actuators for signal amplification and drug delivery. *ACS Appl Mater Interfaces* 3(5):1620–1623
82. Privman V, Pedrosa V, Melnikov D, Pita M, Simonian A, Katz E (2009) Enzymatic AND-gate based on electrode-immobilized glucose-6-phosphate dehydrogenase: Towards digital biosensors and biochemical logic systems with low noise. *Biosens Bioelectron* 25(4):695–701
83. Privman M, Tam TK, Pita M, Katz E (2009) Switchable electrode controlled by enzyme logic network system: approaching physiologically regulated bioelectronics. *J Am Chem Soc* 131(3):1314–1321
84. Rabbee N, Speed T (2006) A genotype calling algorithm for affymetrix SNP arrays. *Bioinformatics* 22(1):7–12

85. Ramlan E, Zauner K-P (2013) In-silico design of computational nucleic acids for molecular information processing. *J Cheminform* 5:22
86. Robertson DL, Joyce GF (1990) Selection in-vitro of an RNA enzyme that specifically cleaves single-stranded-DNA. *Nature* 344(6265):467–468
87. Rowe AA, Miller EA, Plaxco KW (2010) Reagent less measurement of aminoglycoside antibiotics in blood serum via an electrochemical, ribonucleic acid aptamer-based biosensor. *Anal Chem* 82(17):7090–7095
88. SantaLucia J Jr (1998) A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. *Proc Natl Acad Sci USA* 95(4):1460–1465
89. Santoro SW, Joyce GF (1997) A general purpose RNA-cleaving DNA enzyme. *Proc Natl Acad Sci USA* 94(9):4262–4266
90. Schluecker S (2009) SERS microscopy: nanoparticle probes and biomedical applications. *ChemPhysChem* 10(9–10):1344–1354
91. Shinya K, Hamm S, Hatta M, Ito H, Ito T, Kawaoka Y (2004) PB2 amino acid at position 627 affects, replicative efficiency, but not cell tropism, of Hong Kong H5N1 influenza A viruses in mice. *Virology* 320(2):258–266
92. Shlyahovsky B, Di L, Weizmann Y, Nowarski R, Kotler M, Willner I (2007) Spotlighting of cocaine by an autonomous aptamer-based machine. *J Am Chem Soc* 129(13):3814–3815
93. Simmel FC (2007) Towards biomedical applications for nucleic acid nanodevices. *Nanomedicine* 2(6):817–830
94. Simpson ML, Sayler GS, Fleming JT, Applegate B (2001) Whole-cell biocomputing. *Trends Biotechnol* 19(8):317–323
95. Sistare MF, Holmberg RC, Thorp HH (1999) Electrochemical studies of polynucleotide binding and oxidation by metal complexes: effects of scan rate, concentration, and sequence. *J Phys Chem B* 103(48):10718–10728
96. Sivan S, Tuchman S, Lotan N (2003) A biochemical logic gate using an enzyme and its inhibitor. Part II: The logic gate. *Biosystems* 70(1):21–33
97. Smith GJD, Fan XH, Wang J, Li KS, Qin K, Zhang JX, Vijaykrishna D, Cheung CL, Huang K, Rayner JM, Peiris JSM, Chen H, Webster RG, Guan Y (2006) Emergence and predominance of an H5N1 influenza variant in China. *Proc Natl Acad Sci USA* 103(45):16936–16941
98. Stojanovic MN, de Prada P, Landry DW (2001) Aptamer-based folding fluorescent sensor for cocaine. *J Am Chem Soc* 123(21):4928–4931
99. Stojanovic MN, de Prada P (2001) D. W. Catalytic molecular beacons. *ChemBioChem* 2(6):411–415
100. Stojanovic MN, Mitchell TE, Stefanovic D (2002) Deoxyribozyme-based logic gates. *J Am Chem Soc* 124(14):3555–3561
101. Stojanovic MN, Semova S, Kolpashchikov D, Macdonald J, Morgan C, Stefanovic D (2005) Deoxyribozyme-based ligase logic gates and their initial circuits. *J Am Chem Soc* 127(19):6914–6915
102. Stojanovic MN, Stefanovic D (2003) A deoxyribozyme-based molecular automaton. *Nat Biotechnol* 21(9):1069–1074
103. Stojanovic MN, Stefanovic D (2003) Deoxyribozyme-based half-adder. *J Am Chem Soc* 125(22):6673–6676
104. Stojanovic MN, Stefanovic D, LaBean T, Yan H (2005) In: Willner I, Katz E (eds) *Bioelectronics: from theory to applications*. Wiley-VCH, Weinheim, pp 427–455
105. Stojanovic MN, Stefanovic D, Rudchenko S (2014) Exercises in Molecular Computing. *Acc Chem Res* 47(6):1845–1852
106. Strack G, Pita M, Ornatska M, Katz E (2008) Boolean logic gates that use enzymes as input signals. *ChemBioChem* 9(8):1260–1266
107. Tan L, Li Y, Drake TJ, Moroz L, Wang K, Li J, Munteanu A, Yang CJ, Martinez K, Tan W (2005) Molecular beacons for bioanalytical applications. *Analyst* 130(7):1002–1005

108. Taylor SK, Pei R, Moon BC, Damera S, Shen A, Stojanovic MN (2009) Triggered release of an active peptide conjugate from a DNA device by an orally administrable small molecule. *Angewandte Chemie-International Edition* 48(24):4394–4397
109. Tichoniuk M, Ligaj M, Filipiak M (2008) Application of DNA hybridization biosensor as a screening method for the detection of genetically modified food components. *Sensors* 8(4):2118–2135
110. Tokarev I, Gopishetty V, Zhou J, Pita M, Motornov M, Katz E, Minko S (2009) Stimuli-responsive hydrogel membranes coupled with biocatalytic processes. *ACS Appl Mater Interfaces* 1(3):532–536
111. Tuerk C, Gold L (1990) Systematic evolution of ligands by exponential enrichment—RNA ligands to bacteriophage-T4 DNA-polymerase. *Science* 249(4968):505–510
112. Tyagi S, Kramer FR (1996) Molecular beacons: probes that fluoresce upon hybridization. *Nat Biotechnol* 14(3):303–308
113. Unger R, Moulton J (2006) Towards computing with proteins. *Proteins-Structure Function and Bioinformatics* 63(1):53–64
114. Vo-Dinh T, Wang H-N, Scaffidi J (2010) Plasmonic nanoprobe for SERS biosensing and bioimaging. *J Biophotonics* 3(1–2):89–102
115. Von Maltzahn G, Harris TJ, Park J-H, Min D-H, Schmidt AJ, Sailor MJ, Bhatia SN (2007) Nanoparticle self-assembly gated by logical proteolytic triggers. *J Am Chem Soc* 129(19):6064–6065
116. Wang C, Trau D (2011) A portable generic DNA bioassay system based on in situ oligonucleotide synthesis and hybridization detection. *Biosens Bioelectron* 26(5):2436–2441
117. Wang G, Wang Y, Chen L, Choo J (2010) Nanomaterial-assisted aptamers for optical sensing. *Biosens Bioelectron* 25(8):1859–1868
118. Wang J (2000) From DNA biosensors to gene chips. *Nucleic Acids Res* 28(16):3011–3016
119. Wang J, Katz E (2011) Digital biosensors with built-in logic for biomedical applications. *Isr J Chem* 51(1):141–150
120. Wang J, Katz E (2010) Digital biosensors with built-in logic for biomedical applications—biosensors based on a biocomputing concept. *Anal Bioanal Chem* 398(4):1591–1603
121. Waterman MS, Eggert M (1987) A new algorithm for best subsequence alignments with application to tRNA-rRNA comparisons. *J Mol Biol* 197(4):723–728
122. Willner I, Shlyahovsky B, Zayats M, Willner B (2008) DNAzymes for sensing, nanobiotechnology and logic gate applications. *Chem Soc Rev* 37(6):1153–1165
123. Wilson DS, Szostak JW (1999) In vitro selection of functional nucleic acids. *Annu Rev Biochem* 68:611–647
124. Win MN, Smolke CD (2008) Higher-order cellular information processing with synthetic RNA devices. *Science* 322(5900):456–460
125. Weizmann Y, Beissenhirtz MK, Cheglakov Z, Nowarski R, Kotler M, Willner I (2006) A virus spotlighted by an autonomous DNA machine. *Angewandte Chemie-International Edition* 45(44):7384–7388
126. Weizmann Y, Cheglakov Z, Pavlov V, Willner I (2006) Autonomous fueled mechanical replication of nucleic acid templates for the amplified optical detection of DNA. *Angewandte Chemie-International Edition* 45(14):2238–2242
127. Wu Z-S, Jiang J-H, Shen G-L, Yu R-Q (2007) Highly sensitive DNA detection and point mutation identification: An electrochemical approach based on the combined use of ligase and reverse molecular beacon. *Hum Mutat* 28(6):630–637
128. Xiao Y, Lubin AA, Heeger AJ, Plaxco KW (2005) Label-free electronic detection of thrombin in blood serum by using an aptamer-based sensor. *Angewandte Chemie-International Edition* 44(34):5456–5459
129. Yamada S, Suzuki Y, Suzuki T, Le MQ, Nidom CA, Sakai-Tagawa Y, Muramoto Y, Ito M, Kiso M, Horimoto T, Shinya K, Sawada T, Kiso M, Usui T, Murata T, Lin YP, Hay A, Haire LF, Stevens DJ, Russell RJ, Gamblin SJ, Skehel JJ, Kawaoka Y (2006)

- Haemagglutinin mutations responsible for the binding of H5N1 influenza A viruses to human-type receptors. *Nature* 444(7117):378–382
130. Yashin R, Rudchenko S, Stojanovic MN (2007) Networking particles over distance using oligonucleotide-based devices. *J Am Chem Soc* 129(50):15581–15584
  131. Yeh H-Y, Yates M, Mulchandania A, Chen W (2010) Molecular beacon–quantum dot–Au nanoparticle hybrid nanoprobes for visualizing virus replication in living cells. *Chem Commun* 46:3914–3916
  132. Zhou J, Arugula MA, Halamek J, Pita M, Katz E (2009) Enzyme-based NAND and NOR logic gates with modular design. *J Phys Chem B* 113(49):16065–16070
  133. Zhou J, Halamek J, Bocharova V, Wang J, Katz E (2011) Bio-logic analysis of injury biomarker patterns in human serum samples. *Talanta* 83(3):955–959
  134. Zhou J, Tam TK, Pita M, Ornatska M, Minko S, Katz E (2009) Bioelectrocatalytic system coupled with enzyme-based biocomputing ensembles performing boolean logic operations: Approaching “smart” physiologically controlled biointerfaces. *ACS Appl Mater Interfaces* 1(1):144–149