Structural DNA Nanotechnology: From Bases to Bricks, From Structure to Function

Souvik Modi,† Dhiraj Bhatia,† Friedrich C Simmel,*‡† and Yamuna Krishnan*‡

†National Centre for Biological Sciences, Tata Institute of Fundamental Research, GKV, Bellary Road, Bangalore 560 065, India, and ‡Lehrstuhl für Bioelektronik E14, Physik Department, Technische Universität München, James-Franck-Strasse 1, D-85748 Garching, Germany

ABSTRACT The two fields of structural DNA nanotechnology and functional nucleic acids have been independently coevolving, with the former seeking to arrange and bring about movement of nucleic acid modules precisely and with control in space and the latter producing modules with incredible diversity in effective recognition and function. Here, we track the key developments in structural DNA nanotechnology that reveal a current trend that is seeing the integration of functional nucleic acid modules into their architectures to access a range of new functions. This contribution will seek to provide a perspective for the field of structural DNA nanotechnology where the integration of such functional modules on precisely controlled architectures can uncover phenomena of interest to physical chemists.

DNA has proven to be a powerful material for construction on the nanoscale on the basis of the following properties: (i) the availability of automated synthetic methods and continually dropping costs, (ii) chemical robustness that confers stability on the resultant architectures and their subsequent ability to be functional under a variety of environmental conditions, (iii) the uniformly rod-like nature of the DNA double helix irrespective of its primary sequence, (iv) the specificity of Watson–Crick base pairing, which functions as an easily engineerable, site-specific, molecular-scale glue applicable to any DNA double helix, (v) the periodic nature of the DNA double helix and the predictable nature of sequence-specific thermal stability, both of which predispose it to computational methods to design and fabricate superarchitectures, (vi) the availability of well-characterized biochemical and molecular biological methods to cut, copy, and covalently link B-DNA double helices sequence-specifically, which allows manipulation of the construction material, (vii) the modular nature of the DNA scaffold that allows fabrication of architectures that are complex in terms of both structure and function when multiple modules are appended to each other, and (viii) single-stranded DNA sequences, called functional nucleic acids, which can fold and offer three-dimensional cavities suited to bind with great specificity a range of molecular entities with diverse function.

In 1982, Ned Seeman proposed that DNA, which until then had been thought of as a linear polymer, could be used to make branched architectures by using stable artificial junctions with helical DNA limbs radiating from a central node.¹ These structures were analogous to metastable naturally occurring DNA motifs, such as the replication fork and Holliday junction. “It appears to be possible to generate covalently joined…networks of nucleic acids which are periodic in connectivity and perhaps in space.”¹ This marked the origin of structural DNA nanotechnology that seeks to create defined architectures on the nanoscale using sequences of DNA that self-assemble into rigid rods that are, in turn, connected to form superarchitectures of precise dimensions. In 1999, it was shown that DNA could switch between two forms (the B-form and the Z-form), and this motion could be transduced along a DNA architecture, making it undergo a twisting motion.² Thus began a complementary aspect of structural DNA nanotechnology, of bringing about defined molecular-scale movements of DNA architectures triggered by the addition of input stimuli that are chemical, photonic, thermal, or electrical in nature.

Functional nucleic acids are obtained from a test tube evolution method called SELEX independently conceptualized by the Szostak and Gold groups.³–⁴ It uses molecular biology tools to pick out from a library of ~10¹⁵ different DNA (or RNA) sequences, a subset of sequences based on a given selection criterion and amplify them.⁵ When subjected to the same selection criterion repeatedly with progressively higher stringencies, it is possible to progressively enrich from the library, a pool of DNA (or RNA) sequences with a specific functionality. If the selection criterion is the recognition of a target molecule, then selected single-stranded DNA (ssDNA) sequences are capable of binding to the target with high specificity and affinity. Thus, SELEX has yielded DNA sequences that can bind a huge variety of chemical entities ranging from small molecules to proteins, peptides, transition-state

Received Date: April 17, 2010
Accepted Date: June 4, 2010
Published on Web Date: June 14, 2010
intermediates, and even whole cells. These ssDNA sequences called aptamers (originating from the word aptus, meaning to fit) are generally about 15–40 nucleotides long and can fold up in three dimensions to offer a highly selective cavity into which the target molecule snugly fits. SELEX has also yielded ssDNA (or ssRNA) sequences, called DNAzymes (or RNAzymes), that are capable of functioning as catalysts of a variety of chemical reactions.

Here, the selection criterion is based on whether the reaction is a bond-forming or bond-breaking reaction and makes use of the fact that DNA (or RNA) sequences with the desired property have undergone a chemical change such as losing a segment in a bond-breaking reaction or gaining a segment in a bond-making reaction. If the segment that is lost or gained incorporates a molecular tag (such as biotin), then it is facile to enrich the pool by separating the molecules that have the tag from those that do not, and the desired pool may be taken on for further enrichment. Thus, nucleic acid enzymes are known that can catalyze several reactions such as Diels–Alder, Michael, aldol, acylation, phosphotransferases, esterases, and ligase activities, to name a few.

The two fields of structural DNA nanotechnology and functional nucleic acids have been independently coevolving (Figure 1), with the former seeking to arrange and bring about movement of nucleic acid modules precisely and with control in space and the latter producing modules with incredible diversity in effective recognition and function. Here, we track the key developments in structural DNA nanotechnology that reveal a current trend that is seeing the integration of functional nucleic acid modules into their architectures to access a range of new functions. This contribution will seek to provide a perspective for the field where the integration of such functional modules on precisely controlled architectures can uncover phenomena of interest to physical chemists.

Rigid Architectures. In 1991, Seeman demonstrated the first simple topological architecture made from DNA, which resembled a cube (Figure 2a). However, it was soon realized that in order to construct more complex, rigid architectures, stronger duplex DNA motifs were required. Inspired by biological crossover DNA motifs, Seeman designed structures such as double crossover (DX) and paranemic crossover (PX)
junctons in which two helices are conjoined along their long axes by ss-DNA strands that traverse alternately between both of the helices (Figure 2b).  

The advantage with these motifs is the ease of modulating the dimensions of a tile (see later) while maintaining junction geometry, important for constructing higher-order structures. DX motifs could be made to combine to form arrays in 2D via programming their overhang sequences (Figure 3). However, the use of unique sequences in overhangs demanded different types of DX tiles that, in turn, required a larger number of sequences, reducing yields. Mao et al. exploited the symmetry inherent in a DX tile to make DX tiles with a reduced number of sequences per tile (Figure 3a) and higher yields. In an important step forward, Yan et al. devised a way to make 2D sheets of defined dimensions. They preserved the symmetry of the tile junction but made the tile overhangs asymmetric. Now, since the overhangs are asymmetric, the number of unique tiles required to create a finite-sized tile is determined by the axis of symmetry in the final tile. Thus, using tiles where all four overhangs are asymmetric to make a finite-sized 2D sheet comprising $N$ tiles would require the association of $N$ distinct types of such tiles. With Yan’s approach, if the tiles have $C_m$ symmetry (where $m = 2, 3, 4, 6$), the number of unique tiles needed will be only $\frac{N}{m}$ (Figure 3b). Controlled growth of 2D patterns was also achieved by Erik Winfree using a different strategy of algorithmic self-assembly (see later). An example is illustrated using Sierpinski triangles, which are generated using DX tiles whose sticky ends represent logical 0 and 1 that act as inputs or outputs for the algorithm (or a set of instructions). These tiles assemble into various 2D shapes using molecular logic. For example, the association of DX tiles according to XOR logic leads to the formation of a Sierpinski triangle (Figure 3c).

Extended 2D sheets of DNA can curve onto themselves such that sticky edges meet, giving rise to DNA nanotubes with varying dimensions. LaBean and Reif designed a strategy to create nanotubes using DNA helix bundles, which made it possible to control the circumference of these nanotubes. This is achieved by first creating 2D sheets of varying lengths where the first and the last domains in these sheets are complementary. Therefore, the simple pairing between domains in the same sheet results in monodisperse nanotubes whose circumference is dictated by the width of the 2D sheet. Nanotubes, like 2D tiles, can also act as scaffolds to precisely position gold nanoparticles along their lengths. Alternatively, the surface of DNA nanotubes can be coated with a thin metal film to form nanowires. DNA nanotubes provide ideal platforms to create predefined molecular tracks for molecular motors like myosin or kinesin by positioning precise footholds for the latter.

In 2006, Paul Rothemund introduced a paradigm change by using a viral genome to fold DNA helices into any 2D shape by adding small staple strands. This involved the folding of a long viral genomic DNA by using many small staple DNA strands which could control the local folding of the long DNA, resulting in it adopting the desired 2D shape. This is called DNA origami (Figure 4). Shih and co-workers then showed the analogous folding and twisting of DNA helix bundles into 3D structures, called 3D origami (Figure 5). Origami-based nanoconstructions, though simple to form, suffer from the drawback that they require hundreds of strands to be mixed, which considerably increases the complexity of the systems. In this respect, other approaches toward nanconstruction using only small numbers of strands can be favorable.

The tetrahedron is likely to emerge as a powerful model system to test the controlled manipulation of 3D structures.

William Shih designed an octahedron that folded using a 1.7 kb long single-stranded piece of DNA. In 2005, a DNA tetrahedron was made by Turberfield and co-workers in very high yields by mixing of four strands of DNA in solution.
This system is likely to emerge as a powerful model system to test the controlled manipulation of 3D structures due to its ease of production and clean characteristics. Krishnan et al. showed that a DNA icosahedron could encapsulate other...
nanoscale objects from solution.\textsuperscript{24} This system also has a potential as a test bed to study the behavior of various biomolecules under nanoscale confinement. Biochemists have always used bacteria to produce large amounts of specific DNA sequences for their assays. In an important proof of concept, Yan and Seeman produced simple DNA nanostructures inside bacteria that could be isolated and purified on larger scales.\textsuperscript{25} The in vivo replication of DNA

Figure 5. 3D DNA origami. Top panel: A long strand of DNA (shown in gray) is first folded into a 2D sheet using staple strands (orange and blue). Selected portions of the 2D sheet can then be joined together in space using other sequences of DNA which protrude out from the plane of the DNA sheet (blue and white), folding the sheet up into a 3D object. Middle panel: Projections of desired 3D objects, where each DNA double helix is represented by a cylinder. Bottom panel: Relevant projections of the EM images of DNA folded by 3D origami. Scale bar: 20 nm. Reprinted with permission from Macmillan Publishers Ltd., Nature 2009, 459, 414, copyright 2009, Nature Publishing Group.

Figure 6. (a) First DNA nanomechanical device comprising a circular DNA with a cruciform at the center. This device changes its state from a maximally extruded relaxed position (right) to a minimally extruded strained position (left), based on the degree of supercoiling. Adapted from Trends Biochem. Sci. 2005, 30, 119. (b) Device based on the B/Z transition. A DNA sequence which acts as a shaft (yellow) links two DX motifs. The shaft either adopts the B-DNA conformation, positioning two fluorophores in close proximity, or adopts the Z-DNA conformation in the presence of $\left[\text{Co(NH}_3)_6\right]^{3+}$, which rotates the assembly to place the fluorophores distally. Reprinted with permission from Macmillan Publishers Ltd., Nature 1999, 397, 144, copyright 1999, Nature Publishing Group.
DNA motifs can also exist in a topoisomeric form called the JX-handed Z-DNA, resulting in a twisting motion. However, the first device with more well-defined conformational DNA strands—the first DNA tweezers, operated by the sequential addition of complementary DNA strands bearing a toehold of a DNA tetrahedron by the sequential addition of complementary sequences—was not observed. The first real-time transition between two states and triggered by the addition of a set of three DNA strands that controls PX→JX2 transition and a reporter hairpin whose position is spatially altered depending on whether the device is in the PX topology or the JX2 topology.

Dynamic Architectures. The first nanomechanical device was constructed from a circular DNA containing a cruciform which could undergo branch migration depending on the supercoiled status of the circular DNA. It was shown to change from the maximally extruded relaxed position to a minimally extruded strained position by the addition of ethidium bromide, which altered the degree of supercoiling. However, the first device with more well-defined conformational states achieved on an artificial DNA assembly was based on the transition from right-handed B-DNA to left-handed Z-DNA, resulting in a twisting motion.

However, the real-time transition between these two states was not observed. The first real-time transition between two switch states was observed by Yurke and colleagues, who made the first DNA tweezers, operated by the sequential addition of DNA strands. The first robust switchable device triggered by strand hybridization was the PX→JX2 device of Seeman et al. The parameric crossover (PX) DNA motif can also exist in a topoisomeric form called the JX2 form, shown in Figure 7b. The PX→JX2 device uses two strands that hybridize to the PX form and stabilize the tile in its JX2 topology. These strands are removed as duplexes to restore the device to the PX form. This principle was used to construct a DNA robotic arm. This device consists of three domains: one is an attachment site for incorporation into a 2D array, whereas the other two domains contain (i) a region that incorporates a rotary PX→JX2 module that is activated by set strands that control PX→JX2 transition and (ii) a reporter hairpin whose position is spatially altered depending on whether the device is in the PX topology or the JX2 topology.

Cells use molecular motors to transport molecular entities on the nanoscale to carry out processes at designated locations. Consequently, one of the major challenges in nanotechnology is to design scaffolds that unidirectionally transport a nanoscale object from a specific location to a precise destination along a predesignated path. The same year that the kinesin molecular motor’s Hand-Over-Hand walking mechanism was established also saw DNA devices that walked on a
molecular track also made from DNA. Sherman and Seeman made a DNA walker consisting of a DNA-based footpath, two legs connected by single-stranded feet, and footholds on the footpaths. The resting and moving states were achieved by using two sets of DNA fuel strands designated as unset and set strands (Figure 8a). Shin and Pierce demonstrated a processive bipedal DNA motor that moved by advancing the trailing foot to the leader foot at each step. Yan and co-workers also developed an autonomous, unidirectional DNA walker. However, the coordinated, synchronous movement of walker legs remained a challenge until very recently. Omabegho et al. showed a DNA bipedal walker where such coordinated motion was achieved by hybridization of a set of metastable DNA strands. Interestingly, cross-linking studies showed that the walker functioned as a Brownian motor completing a full walking cycle on a track of variable lengths. Such artificial DNA motors could be used as reduced model systems to understand general aspects of the physics of molecular motors.

DNA nanodevices powered by motifs other than B-DNA have been also characterized. Such devices undergo conformational changes in response to ions, small molecules, proteins, as well as other DNA or RNA strands. In 2002, the Merzlyak and Tan groups simultaneously reported G-quadruplex-based nanomachines where one of the states was a G-quadruplex that could be stretched out into a duplex by the sequential addition of fuel and antifuel DNA strands (Figure 8b). Sen et al. used a modified strategy employing a duplex with a mismatched internal G-rich region that, in the presence of Sr$^{2+}$, produced quadruplex that resulted in a pinched duplex that could be relieved by the addition of EDTA. This pinching motion providing a contractile force could be the basis of a force sensor to probe molecular mechanical processes. Similarly, C-rich strands form a compact structure called an I-motif under acidic conditions. The Balasubramanian group constructed a similar device which used protons as a toggle between an extended state (at neutral pH) and a compacted I-motif state (at acidic pH). Simmel et al. demonstrated that such proton-fueled DNA devices could be driven autonomously by pH oscillations. Recently, Modi et al. showed that an I-motif-based DNA nanomachine could function as an ultrasensitive pH sensor, reporting on pH changes inside of endosomes of living cells in real time, illustrating the potential of such DNA devices in biology (Figure 8c).

Despite a variety of proof-of-concept demonstrations, it was at least 6 years before a functional dynamic DNA device was reported. Thrombin is a coagulation protein in the blood, and an aptamer that binds it (thrombin binding aptamer, TBA) is known to exist as a G-quadruplex. Taking advantage of this, Simmel et al. made a G-quadruplex-based DNA device, incorporating the TBA to control binding and release of thrombin coordinated with the opening and closing of the DNA device (Figure 9a). This opened up a new vista of functional applications for such movable DNA devices in vitro. In a separate study, the Willner group used DNA strips as a rigid scaffold which had regions that could immobilize DNA strands conjugated to proteins. By using DNA conjugated to glucose oxidase (GOx, which converts glucose to gluconic acid and generates H$_2$O$_2$) and horseradish peroxidase...
(HRP, breaks down H$_2$O$_2$) these enzymes could be accurately positioned in close proximity. The relative positions of GOx and HRP could be precisely tuned on the DNA scaffold, thus resulting in an apparent increase in the local concentration of the enzyme. Thus, the addition of glucose to the assembly generates H$_2$O$_2$, which is the substrate for the proximally positioned HRP, resulting in an enzyme cascade.$^{28}$ Subsequently, this positioning could be achieved using a molecular trigger such as cocaine by integrating a cocaine aptamer that facilitates self-assembly of the DNA scaffold (Figure 9b).$^{52}$

The integration of many other aptamers as functional modules into both rigid and dynamic scaffolds is emerging. For example, Yingfu Li’s group was the first to show that the fluorescently labeled adenosine aptamer, coupled with the sequential addition of adenosine and adenosine deaminase, led to a promising molecular switch.$^{53}$ This was used as a probe for high-throughput small-molecule screening for inhibitors of adenosine deaminase.$^{54}$ The Willner group extended this idea to make tweezers based on this design strategy that could transduce a conformational change onto another set of tweezers (Figure 9c).$^{55}$ Very often, cell signaling that originates at the cell membrane is accompanied by the orchestrated clustering of specific proteins, the molecular arrangements of which are as yet unknown. 2D arrays incorporating aptamers to key proteins could be used as reduced systems to understand the physics of signal propagation in terms of protein densities and arrangements.

The strategic integration of quantum dots (QDs) into dynamic devices would enable tracking device operation over extended periods of time. Conjugated polymers are light sources with positive charges and thus might not even require covalent linkage with the DNA scaffold. By combining QD-functionalized DNA and conjugated polymers, one could make photonic cascades that couple multiple molecular logic gates to achieve more complex logical operations.$^{56}$ An underutilized possibility to control DNA nanomachines is the use of light. Recent work shows that azobenzene-functionalized nucleobases are a promising light-based deactivator of DNA structure.$^{57}$ Thus, one can envisage phototriggered control of DNA switches in vitro and in cellulo as a key advance in making next-generation sensors where function is elicited with spatial and temporal control. Importantly, a unifying observation in such dynamic architectures is a defined operation (motion) that occurs upon the introduction of a defined input or molecular stimulus.
This has resulted in many of these assemblies being used as molecular logic gates in DNA-based computation.

**Computation with DNA.** A simple chemical reaction may be viewed as a computation where reactants are the inputs, the product is the output, and the reaction is the processor of that input. DNA can also be used as a substrate for computing since it can store information in the form of its bases and there is a rich set of biochemical processes available to it. In 1994, Len Adleman showed the first proof-of-concept that DNA could be used for computation in vitro using a set of DNA strands to solve the Hamiltonian path problem, which is a special case of the traveling salesman problem. This pioneered a cascade of other NP-complete problems that were solved using DNA scaffolds such as the satisfiability problem (SAT) and the maximum clique problem. With the help of restriction enzymes such as the endonuclease FokI, DNA has also been used to construct simple computational units belonging to the class of finite-state machines. When DNA is used to compute solutions to hard NP-complete problems, the output is read using a sequencing paradigm. However, silicon-based computation is achieved using binary logic, where the input and output are in a binary format. Thus, there has also been interest in exploring DNA’s ability to compute within the same paradigm of binary logic, where the output is also in a binary format. As a testament to DNA’s applicability to computation, this was achieved in two ways, (i) construction of logic gates using DNA and (ii) logical self-assembly of DNA into specific superarchitectures.

(i) Logic gates using DNA: Small units which control the processing of information according to a set of operations are called logic gates. A molecular logic gate senses one or more inputs and, based on some intrinsic/designed information processing, produces an output. Here, DNA assemblies contain fluorescent tags that are either activated or deactivated (the output) based on a trigger (the input) because the underlying DNA scaffold undergoes a molecular transformation in response to the input strands. A DNAzyme consists of two components, a catalytic core and an internal single-stranded loop. When a single-stranded DNA substrate binds to the internal loop of the DNAzyme, it is cleaved at a specific location by the DNAzyme. Stojanovic and colleagues designed methods to use DNAzymes as molecular logic gates. Thus, different sequences of ssDNA that activate the DNAzyme are considered as inputs, cleavage of the target DNA is the operation, and the properties of the cleaved products are the output. The most commonly used output measurement is fluorescence detection. Figure 10 shows the operation of such an AND gate based on a DNAzyme. This is a simple logic gate in which the output is 1 or YES only if both inputs are present. In the absence of at least one input, the output is 0 or NO. In a major advance, Stojanovic et al. combined different logic gates to create a platform for multiple autonomous operations. This setup, called a molecular automaton, performs all of the downstream actions once the operation is triggered by MgCl2. The automaton uses a combination of YES and ANDANDNOT gates using an array of DNAzymes and is called MAYA (molecular array of YES and ANDANDNOT gates). The ANDANDNOT operation can actually be used to implement the well-known tic-tac-toe game. An operator plays against the molecular computer by adding input strands to the 3 x 3 = 9 wells of MAYA, similar to making crosses in the classic tic-tac-toe game. MAYA then autonomously computes its corresponding move. Stojanovic and colleagues have also used other combinations of DNAzyme-based logic gates to create a half adder, full adder, and logic gates with more than two inputs.

(ii) Logical assembly of DNA tiles: Another approach, pioneered by Erik Winfree, demonstrates DNA’s capability of binary logic. It uses as its input combinations of different DNA tiles that further self-assemble according to the input logic into different kinds of superarchitectures. Here, DNA tiles are programmed with single-stranded overhangs and are considered as inputs. Combinations of these DNA tiles self-assemble via these overhangs into larger superstructures in 2D. Such programmed assembly of DNA is called algorithmic self assembly (outlined already), where the positional information of the input tiles on the resultant superarchitecture, according to binary logic, encodes the output. However, this is limited by high error rates in tile assembly, and a major challenge is to design programmed overhangs that minimize self-assembly error rates. An error reduction strategy developed by the Winfree lab adopts proof-reading tile sets that utilize cooperative binding effects.

With DNA computation, we need not be limited to binary logic. It is possible to construct assemblies that have multiple inputs and multiple operational states and perform higher-order logical computation. DNA computation is certainly finding niche applications in biology. One of these is detection and diagnostics in various biological samples. Winfree and colleagues recently demonstrated that DNA-based logic gates could be used to detect microRNAs in vitro. One could envisage coupling smart DNA and RNA units in stages to create molecular cascades or circuits, and thus, the emergence of a parallel field of RNA computation is not surprising. Recently, Smolke and colleagues were able to couple RNAzyme function and gene activity in cellulo to perform a logical operation. The input was the activation of an RNAzyme module engineered onto an mRNA scaffold, and the output was a change in the protein expression levels encoded by the mRNA.
Biographies

Souvik Modi obtained his M.Sc. (Chemistry) in 2005 from Jadavpur University, Kolkata, and has been working towards his Ph.D. since 2005 with Dr. Yamuna Krishnan. His interests are in DNA-based nanomachines, biosensors for in cellulo applications, methods in microscopy, and single-molecule studies.

Dhiraj Bhatia obtained his Masters in Organic Chemistry from the University of Pune (2007). He has been a graduate student with Dr. Yamuna Krishnan since 2007 and is currently working on DNA-based devices that are many-fold greater in complexity both structurally and functionally should become possible in the future.

REFERENCES


