Designed Self-Organization for Molecular Optoelectronics

Michael Norton*, David Neff, Ian Towler, Scott Day, Zachary Grambos, Mikala Shremshock, Heather Butts, Christiaan Meadows, Yuko Samiso, Huan Cao and Mashiur Rahman
Department of Chemistry, 1 John Marshall Drive, Marshall University, Huntington, WV 25755

ABSTRACT

The convergence of terahertz spectroscopy and single molecule experimentation offer significant promise of enhancement in sensitivity and selectivity in molecular recognition, identification and quantitation germane to military and security applications. This presentation reports the results of experiments which address fundamental barriers to the integration of large, patterned bio-compatible molecular opto-electronic systems with silicon based microelectronic systems. The central thrust of this approach is sequential epitaxy on surface bound single stranded DNA one-dimensional substrates. The challenge of producing highly structured macromolecular substrates, which are necessary in order to implement molecular nanolithography, has been addressed by combining "designer" synthetic DNA with biosynthetically derived plasmid components. By design, these one dimensional templates are composed of domains which contain sites which are recognized, and therefore addressable by either complementary DNA sequences and/or selected enzymes. Such design is necessary in order to access the nominal 2 nm linewidth potential resolution of nanolithography on these one-dimensional substrates. The recognition and binding properties of DNA ensure that the lithographic process is intrinsically self-organizing, and therefore self-aligning, a necessity for assembly processes at the requisite resolution. Another requirement of this molecular epitaxy approach is that the substrate must be immobilized. The challenge of robust surface immobilization is being addressed via the production of the equivalent of molecular tube sockets. In this application, multi-valent core-shell fluorescent quantum dots provide a mechanism to prepare surface attachment sites with a pre-determined 1:1 attachment site : substrate (DNA) molecule ratio.

Keywords: optical, electronic, self-organizing, DNA, lithography, molecular, self-aligning, dimensionality, immobilization.

1. INTRODUCTION

One impediment to the advancement of molecular electronics is the lack of reliable, reproducible single molecule testbeds. Such testbeds would be capable of positioning any selected molecule at any given location on a substrate, with angstrom precision. Furthermore, the molecule of interest could be presented with any orientation selected by the experimenter, in an environment specifically designed for the study. It would be preferred that this environment would provide optical and/or electronic and/or mechanical (vibronic) input and output connections to the target molecule. Eventually, this testbed should be capable of providing this level of control over many molecules simultaneously. It is from these general criteria that the project discussed in this paper arose and has developed.

1.1 Testbed Building Blocks:

The best model for a testbed with such dimensional control was provided in a paper by Winfree et al in 1998 (1). Although the central theme of that publication was the production of infinite two dimensional crystals of DNA, an important breakthrough was the production of well defined, tightly packed, rigid block structures, which can be used as molecular building blocks bearing a strong resemblance to the construction materials called LEGO blocks. A molecular model of a part of one of these building blocks is displayed in Fig. 1 below. The model was generated using Hyperchem software.

Norton@marshall.edu; phone 304-696-6627;fax 304-696-3243

Terahertz for Military and Security Applications IV, edited by
Dwight L. Woolard, R. Jennifer Hwu, Mark J. Rosker, and James O. Jensen,
Proc. of SPIE Vol. 6212, 621203, (2006) · 0277-786X/06/$15 · doi: 10.1117/12.665297

Major properties of these blocks, which render them most suitable for the fabrication of a molecular testbed device is that they are composed of a considerable number of DNA bases, and that they have multiple connectivity, in the form of 4 unique tethers, which bind them to other blocks. In terminology abstracted from molecular biology, these uniquely coded, floppy single stranded ends are called sticky ends. The fact that these ends are most stable when paired with their complements enables these features to endow a block with a unique addressing code. The blocks will be most stable, and therefore are most likely to assemble into, an environment where the ends at each of the 4 corners are complemented, ie are bound (paired) to their complementary bases on another block.

In this molecular model a central double stranded region can be observed, the light blue (central) strand is complementary, or binds to the red (uppermost) strand in some locations, and the light blue strand binds to the dark blue (lower) strand at other locations. The central region bounded by the light blue, circular strand is 21 base pairs in length, which is approximately two full rotations of the double stranded helix of DNA. At both extrema of the light blue central strand, this polymeric strand of DNA "crosses over" from pairing with one strand of DNA, for example the top strand, to being paired with another strand, the bottom strand. Winfree et al termed this a double crossover structure, and they designed it this way not only to achieve close packing for the assembly, but also to induce rigidity and therefore dimensional control. The effect on rigidity is similar to that observed upon crosslinking other polymers.

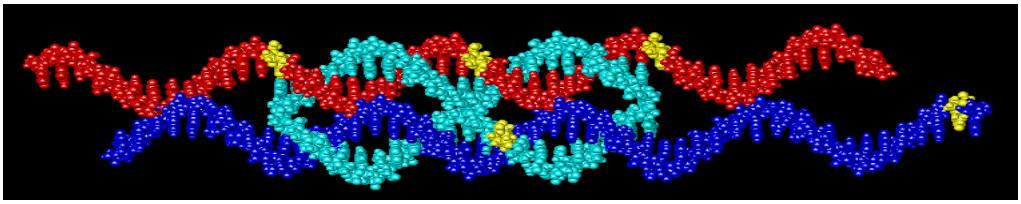

Figure 1: Molecular model of Double Crossover "Block" Structure.

The ability to place multiple molecules into one of these blocks is also represented in Fig. 1. Each of the light yellow (roughly circular white spots) represesents a location where fluorescent molecules (fluorescein) have been "installed" into the structure, for diagnostic purposes, in these studies. It should be noted that any conceivable (water stable) molecule could be similarly bound to this block system. The ability to emplace molecules with sub-nanometer precision is only possible in systems composed of high information density molecules such as DNA. First, the molecule must be capable of precise self alignment, which is a characteristic of DNA and other biopolymers which use specific patterns of hydrogen bonding to bind in an aligned form. Secondly, the polymer must be, at least locally, synthetic, since the fluorophor, or other molecule of interest, must be spliced into the polymer backbone at the precise location designated for the particular experiment. RNA, PNA and protein (peptides) are amenable to this same type of molecular engineering. They are listed in ascending order of cost. There is a price in thermodynamic stability and in structural rigidity to be paid for each substitution of a test molecule for a base. The loss in stability is due to the fact that each substitution removes one of the bases from one of the DNA polymer strands, which lowers the number of hydrogen bonds which participate in binding together the block system. The loss of stability due to the loss of one such pairing per block may, to a first approximation, be considered negligible. Of the bases in DNA, the A-T base pair contributes 2 hydrogen bonds to the stability of the system, and the G-C pair contributes 3 hydrogen bonds. For the 21 base pair region in the center of the block, then, assuming an even A-T-G-C distribution in the composition, the loss due to substitution of a fluorophor for a G (a "worst case" calculation) would be ca 3 parts in 52.5. In terms of melting point, a method of determining and describing stability of a structure, the decrease in melting point would be on the order of 3 degrees C, unimportant for this particular construct. End modifications would have an even lower impact on block stability, however the ends of DNA molecules posses the highest mobilities of any component of the constructs, and therefore substitutions at these sites are the least well determined geometrically. The loss in local rigidity is due to the fact that after the substitution two bases (the substituted base and the complementary base) now have a widely expanded range of conformational freedom, since they are not constrained to the restricted geometry imposed by the hydrogen bonding pattern of normal base pairs. Due to the low scattering power of these block constructs, and the fact

that they will not form well ordered three dimensional crystals, structural determinations at atomic resolution are currently lacking, and molecular modeling is the only approach to molecular design at this scale.

Although a high level of control at the subnanometer size regime appears to be entirely possible utilizing this block structure (which has approximate dimensions of 4 nm X 32 nm [1] a problem then becomes how to uniquely localize such a very small individual block within a larger structure, to form systems and architectures. Winfree's original work emphasized the crystal like properties of two dimensional self assembling constructs of these blocks. Crystals measuring multiple square microns in area were constructed. Although highly repetitive arrays do have an appropriate place in system design, unique addressing requires broken symmetries inconsistent with the regularities of growth of crystals from identical building blocks.

1.2 One Dimensional Asymmetry in two dimensional DNA systems:

The problem of imposing lowered symmetry in a two dimensional system has been well addressed by the Reif group in the form of "Barcode Lattices". In this excellent approach, a "scaffold" strand is encoded with sequence information which enables preferential binding of blocks with one type of binding sequence to one segment of the scaffold strand and binding of another subset of blocks to other regions of the strand [2]. In this model, identical blocks stack on top of each other, forming a barcode type pattern. A schematic representation of the modulation provided in a barcoded block system is provided in Fig. 2. The constructs are constrained by the length of the scaffold strand, and are therefore finite in the barcoded (horizontal) direction (the low symmetry dimension), and are (in theory) infinite in the second, high symmetry (vertical) dimension. These constructs have been assembled in solution phase.

Figure 2: Schematic representation of a barcoded 2-D structure, asymmetric in horizontal direction.

While these systems clearly only demonstrate unique addressing in one dimension, the work represents a definitive step forward towards achieving unique addressing in two dimensions. There are two main approaches to extending this symmetry reduction further, to provide unique identity and therefore unique addressability to every block. One approach uses the concept of cellular automata realized using DNA constructs and is well represented by Winfree's more recent work [3]. In this approach the unique two dimensional addressing required to localize a particular block at a particular location is provided by the preceeding steps in the assembly program. The previous molecular computational steps generate an address in the form of a resultant combination of sticky end sequences, which call for a particular next group of sequences, ie a block. The cellular automata approach is attractive for many reasons, and is in principle amenable to fabrication in solution phase. Another approach, which we term directed sequential self assembly, is the approach we are developing. This approach has been described [4] and is briefly reviewed in the next section.

1.3 Two dimensional asymmetry in two dimensional DNA systems: Directed Sequential Self Assembly

If one could grow, row by row, new layers of blocks on a one dimensionally indexed substrate, then one could combine the one dimensional spatial indexing provided by Reif's [2] scaffold strand with a time (or step) dependent assembly index to provide each block with a unique identity or address. In the directed sequential self assembly approach, a one dimensional director strand of DNA acts as a one dimensional substrate. This substrate must be immobilized on a surface in order to sequentially expose the substrate to a solution containing a mixture of blocks which are to be incorporated into the next one dimensional layer on the one dimensional substrate in an epitaxial process. The reaction must be self terminating such than only one additional layer of blocks, and no more, can attach to the one dimensional "surface" or frontier of the assembly. In this approach, two of the 4 sticky ends on each block are used to ensure integration of the blocks into the correct location through the address encoding capability of DNA. The other two sticky ends are not complementary to any other strands in the solution, and therefore addition of a block terminates the reaction by presenting two non-complementary ends to the solution. Because it is impossible, at this time, to meter out precisely the correct number of blocks into the solution to ensure that no extra blocks remain after the reacton has gone

to completion (ie one row of blocks have added), the solution must be replaced to clear out unreacted blocks which might interfere in subsequent reaction cycles, either by integrating into the growing structures of by reacting with blocks which are components of a later reaction. It is the necessity of this rinse step, which is an intrinsic part of the sequential growth method, which leads to the requirement that the substrate strand must be immobilized on a surface. The sequential process is illustrated in Fig. 3 below.

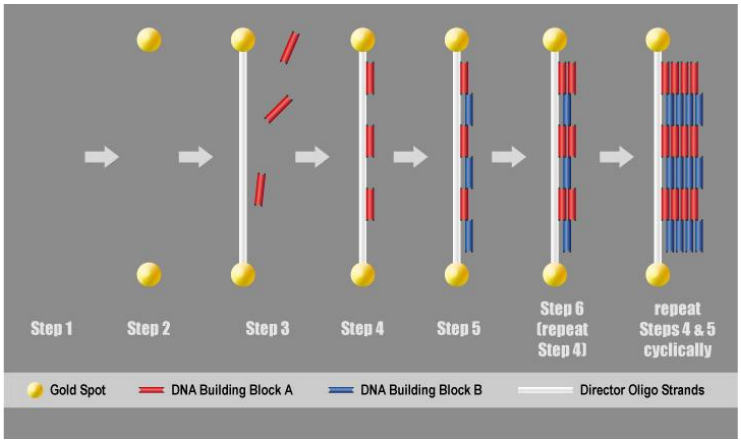

Figure 3: Representation of stepwise or sequential assembly on a 1-D substrate

This sequential method provides the designer with the option of re-using the block addressing sequences (the sticky ends defined above) in successive additions. The configuration of molecular species within a block could be changed even while maintaining the same indexing sequences or sticky ends. Implementation of this time encoding of block index is a focus of our efforts. The remainder of this paper outlines the design parameters and progress toward generation of a construct that will not only demonstrate feasibility of Directed Sequential Self Assembly as an approach to full sub-nanometer addressability on a two dimensional surface, but which can also be developed into an experimental testbed which will be useful for a wide variety of researchers in the nanoscale sciences.

2. METHODOLOGY

2.1 One Dimensional Substrates for Single Molecule Lithography:

The design elements of the targeted substrate architecture are represented schematically in Fig 4 below.

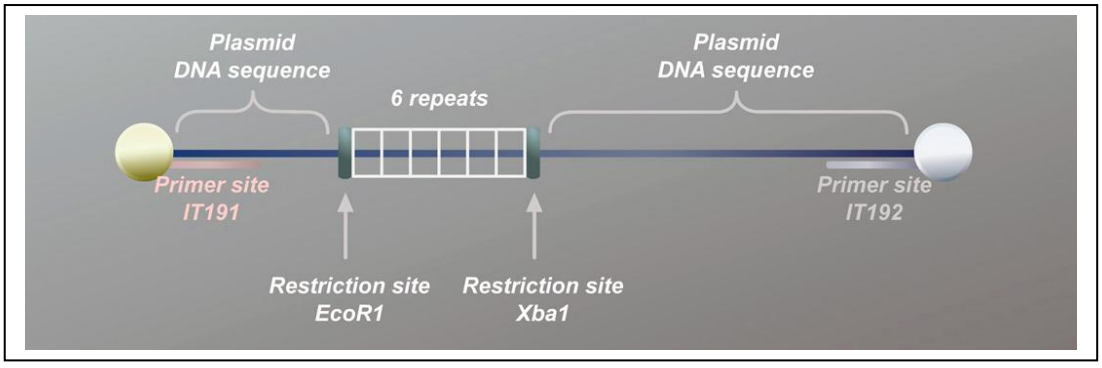

Figure 4: Components of single molecule substrate for use in molecular lithography.

There are 7 different functional domains within this 1 micron (3,167 base) long single stranded substrate. Starting from the left hand side of the molecule these modular regions are 1) an initial binding moiety, 2) a single stranded spacer (left), 3) a restriction enzyme site (left), 4) the active assembly domain, 5) a second restriction enzyme site (right), 6) a second single stranded spacer (right) and finally, 7) a terminal binding moiety. Each of these domains is discussed in greater detail below.

2.1.1 Initial Binding Moiety:

As discussed above in the description of the sequential assembly process, it is essential that the substrate molecule must be immobilized on a surface. All DNA molecules have two ends, an initial end, termed the 5' end and a terminal end, called the 3' end. The dot at the left end of the molecular substrate is meant to represent a domain with a specific attachment chemistry which binds the 5' end of the molecular substrate to the macroscopic substrate. This chemistry should be different than the chemistry used on the other end of the molecular substrate if directional (vectorial) binding of the template to the solid support (substrate) is planned. Because the Polymerase Chain Reaction (PCR) is used to produce the molecular substrate, the 5' end of the molecule is particularly easy to label. In the polymerase reaction a short (often 20 base long) sequence of DNA is extended enzymatically to generate the eventual molecular substrate. This initiator, or primer is indicated as Primer Site IT191 in Fig. 4 above. Several companies manufacture these short sequences chemically and these companies offer a variety of modifications on the end which is not to be extended, the 5' end. At this time, use of biotin is preferred. Biotin can react with a surface immobilized protein molecule of streptavidin with a binding strength which is a significant fraction of a bond strength under very mild conditions. Many other binding chemistries could be employed.

2.1.2 Single Stranded Spacer (Left)

The next domain is a short (ca 198 nm, 582 bp) single stranded spacer constructed of plasmid DNA which places the active assembly region of the molecule away from the first binding site. This region must be masked with a 582 bp masking strand in order to convert it to double stranded DNA. This mask performs several functions. Masking prevents growth of adventitious or erroneous structures on this segment of the sequence, the mask increases the mechanical strength and chemical stability of the construct and this double stranded region increases the dimensional stability of the construct, since the bend-persistence length of double stranded DNA is 50 nm, much greater than that of single stranded DNA. Although single stranded DNA can extend to lengths twice the 0.34 nm average base to base distance observed for double stranded DNA, it is more common to observe that single stranded DNA, in room temperature buffer, exists in a form resembling the molten globular form of denatured proteins. The single stranded "masking strand" is produced using asymmetric PCR. In this technique, PCR is applied to make copies of only one of the two strands, rather than its normal application in amplifying the construct strand and its complement.

This relatively rigid, reinforced tether physically isolates the active region from the binding sites. There are 5 important design reasons to geometrically isolate the active growth region from the binding regions : (A) the binding region is likely to be topographically locally a high z (tall) feature which could interfere with AFM or NSOM imaging/characterization of the active region, (B) such isolation provides the option of electronic addressing (gating) using a buried gate [5], (C) any possible luminescence from the attachment site can be readily discriminated against, (D) quenching of luminescence caused by the composition (eg metallic nature) of the attachment points will be minimized, (E) in cases where the attachment sites are metallic and therefore opaque, the active site can be suspended over an electron transparent or optically transparent window.

It is important to note that the length of the spacer is not fixed, and that it can be modified relatively easily. In order to preserve the active area of the constructs, the sequence for the active area has been inserted, into plasmids, relatively short circular double stranded extranuclear bacterial components. These plasmid constructs are then inserted into bacteria, Ecoli. These host cells can be readily stored at low temperatures, then grown to large numbers relatively rapidly. This biofiling of nanostructures, which enables researchers worldwide to easily transfer "live nanostructure carriers" has been described previously [6]. The nanostructure is extracted from the bacteria by amplifying the target sequence using PCR. The spacing distance is determined by the DNA PCR primer selected by the investigator. The asymmetry of the construct described here was selected in order to provide a method to readily verify, using AFM imaging, that the construct has been immobilized with the correct polarity (5' bound).

2.1.3 Restriciton Enzyme Site: EcoR1

The left tether region is terminated by a DNA sequence which is recognized by the restriction enzyme EcoR1. The sequence G^AATTC of this region of double stranded DNA is cut, or cleaved between the G and A on both strands. This site is commonly used for cloning in bacteria, and therefore allows extraction of the assembly region and its transport into other biological systems. Because the enzyme also cleaves through both strands, this also provides an method to separate the active region from the tether, and therefore from the 5' binding site.

2.1.4 Active Assembly Domain

The active assembly domain is a 194 nm long (570 bp) region of single stranded DNA composed of six concatenated or tandem repeats of 95 base pairs. Production of long single stranded concatemers of this 95 bp sequence using rolling circle synthesis has been reported [7]. In this prototype system, sequence symmetry is maintained along the director strand coordinate, ie in one dimension, the strand dimension, the addressing coordinate is not unique. The unique addressing dimension is designed to be in the growth direction, or the sequential assembly dimension. Structures will be grown in this region of the single molecule substrate through a process of base pairing, in which the single stranded DNA of this region of the substrate will be incorporated into any growing nanoarchitecture, and will thereby be converted to double stranded DNA. The architectures designed for this template thus far bind three different single strands of DNA to each of these 95 bp regions. The attachment/integration strategy for one of the 95 bp regions is illustrated in Fig. 5. This would correspond to the completion of Step 4 in Fig. 3 above.

Figure 5: Example "epitaxial" structure. Top line represents 116 bp of the director strand.

Such a design provides 18 rodlike segments within the active growth region. The sites at the joins of these segments are mechanically equivalent to single stranded breaks in genomic DNA. Two of these joins per 95 bp repeat are bridged via the block structures shown in Fig. 5. The remaining discontinuities, 1 per one dimensional unit cell, will provide a site where the structure can bend at much higher angles than would be the case for a continuously double stranded DNA structure. The next step in sequential assembly, Step 5 in the scheme illustrated in Fig. 3, bridges these joins, reinforces the structure, and produces a rigid framework for further stepwise assembly. These early growth steps represent critical, and difficult to characterize elements of the assembly process.

2.1.5 Restriction Enzyme Site: Xba1

The active assembly domain or building module is bounded on the right (3') side by a release domain which contains the sequence for another restriction enzyme, Xba1. This enzyme restricts, or cuts the sequence between the T and the C in the sequence T^CTAGA. One or both of the tether sites can be cut by using one or both of these restriction enzymes. Using both of the enzymes would free the nanostructure from the substrate and into solution, possibly for incorporation into larger systems. Xba1 is a high efficiency restriction enzyme which is commonly used in cloning experiments, since when combined with enzymes like Ecor1 it provides sticky ends which ensure insertion in the correct direction.

2.1.6 Single Stranded Spacer (Right)

The next domain is the second single stranded spacer and it is approximately 685 nm (2014 bp) in length. The asymmetry of the length of the spacers is meant to allow ready visualization and confirmation of the orientation of the molecular assembly after it is installed on the support. The end of this long spacer is the 3' end of the DNA molecule.

2.1.7 Terminal Binding Moiety

The last domain of the single molecule template represents a region for a second type of chemical attachment. Use of two different attachment strategies allows "orthogonal chemistry" [8], or enables two different, non-interfering attachment chemistries to be utilized simultaneously. There are two mechanisms which may be used to attach binding chemistries to the 3' end of a DNA molecule. A ligase can be employed to directly, covalently attach such a moiety. Alternatively, the primer used to generate the masking strand for the right side spacer can be functionalized before it is incorporated into the masking strand via asymmetric PCR. If a masking strand is to be used, this approach is preferred for its high efficiency.

3. RESULTS AND DISCUSSION

The 1 micron long director strand is most easily visualized via AFM in its double stranded form, as seen in Fig. 6 A. The viability of the biotin attachment chemistry is demonstrated by the AFM in Fig. 6 B, which is an AFM micrograph of double stranded director strand single molecule templates associated with QDots via biotin/streptavidin linker chemistry. One of the major challenges in the production of isolated molecular substrates is the production of single molecule support junctions. Although ebeam lithography is readily capable of producing arrays of attachment sites of ca 30 nm diameter, smaller metallic features are less reproducibly generated. This produces a "footprint" problem in achieving the desired one molecule to one surface attachment site ratio. In order to create a useful testbed, a low error rate method of achieving this ratio must be developed. One solution to this problem is the production of an adapter element, capable of joining one 30 nm feature to one 2 nm diameter DNA molecule. Quantum dots, which have diameters approaching 15 nm may limit access to the surface sites to one attached QDot via steric interactions. Although current QDots are multivalent, and capable of binding multiple substrate strands, preparative gel electrophoresis can be used to identify and separate out QDots bound to only one single molecule substrate strand.

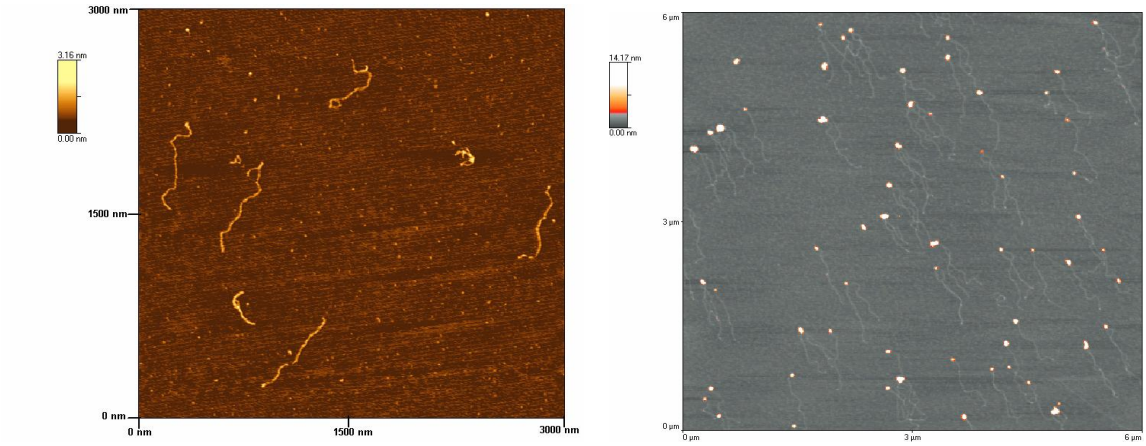

Figure 6: A (left) 1 micron single molecule substrates (scale: 3 microns X and Y; 3 nm Z) B (right) Substrates end bound to 10 nm dia QDots (scale : 6 microns X and Y ; 14 nm Z)

4. CONCLUSIONS AND FUTURE OUTLOOK

This paper only describes in detail one possible path to achieving complete control of surface chemistry and composition. The advent of such control will enable experiments in near field optics, including plasmonics and Forster wire production which will help define the limits of optical physics and computation. As an enabling technology for multiply connected molecules in molecular architectures, one may anticipate that such work will usher in the next generation of optical, electronic and vibronic systems.

5. ACKNOWLEDGMENTS

The authors recognize support for this work from the ARO, WV-EPSCoR and NSF-EPSCoR programs.

6. REFERENCES

1. E. Winfree, F. Liu, L.A. Wenzler and N.C. Seeman; Design and Self-Assembly of Two-Dimensional DNA Crystals, *Nature* 394, 539-544 (1998).

2. H. Yan, T.H. LaBean, L. Feng and J.H. Reif, Directed nucleation assembly of DNA tile complexes for barcode-patterned lattices, PNAS, 100, 14, 8103-8108, 2003.

3 P.W.K. Rothemund, N. Papadakis and E. Winfree, Algorithmic self-assembly of DAN sierpinski triangles, PLoS Biology, 2, 12, 2:e424, 2004.

4 M.Norton, A. Barhoumi and D. Neff, Templates for sequential assembly of DNA based nanostructures, Proceedings 2005 5th IEEE Conference on Nanotechnology, TU-P1-4, 2005.

5 D. Kinneret, R.S. Berman, E. Buchstab, U. Sivan and E. Braun, DNA templated carbon nanotube field effect transistor, Science, 302, 1380-1382, 2003.

6 M.Norton, A. Barhoumi and D. Neff, Templates for sequential assembly of DNA based nanostructures, Proceedings 2005 5th IEEE Conference on Nanotechnology, TU-P1-4, 2005.

7 M. Norton, A. Barhoumi and D. Neff, Toward Large Nanostructures, Proceedings 2003 3rd IEEE Conference on Nanotechnology, 432-436, 2003.

8 T.D. Clark, J. Tien, D.C. Duffy, D.E. Paul and G.M. Whitesides, Self assembly of 10 micron sized objects into ordered three-dimensional arrays, J. Am. Chem. Soc., 123, 7677-7682, 2001.