Chapter 2

Programming Protein Patterns on DNA Nanostructures

The text of this chapter was taken in part from two manuscripts co-authored with John Sadowski and Peter Dervan.

Cohen, J. D.; Sadowski, J. P.; Dervan, P. B., *Angew. Chem. Int. Ed.* **2007**, *46*, 7956-7959

Cohen, J. D.; Sadowski, J. P.; Dervan, P. B., *J. Am. Chem. Soc.* **2008**, *130*, 402-403

Abstract

DNA nanotechnology offers the potential for the creation of molecular assemblies with nanometer scale resolution. Although DNA provides an ideal material for the construction of 2-dimensional and 3-dimensional assemblies it lacks the functionality needed for many desired applications. Pyrrole-imidazole polyamides capable of programmable binding to specific DNA sequences provide a unique solution to the challenge of recruiting active biomolecules to a DNA template. Polyamides are capable of binding to a 2-dimensional DNA array and recruiting protein to specific sites on the nanostructure. Multiple polyamides targeted to different sites on a two dimensional array can be used to create unique protein patterns on the same DNA template. Similar efforts to bind to a two-dimensional DNA array directly with the DNA binding protein EGR-1 were done to compare the effectiveness of each approach.

2.1 Introduction

DNA nanotechnology

The ability to create assemblies of proteins with spacing on the nanometer scale has important implications for proteomics, biodetection, and self-assembly. The goal of such systems is to be able to position proteins or other components in distinct patterns with precise spacing. Structural DNA nanotechnology has led to the creation of a variety of nanostructures which should be capable of serving as an addressable template for the creation of complex molecular assemblies.[1-4] These systems take advantage of the well defined structure and spacing of DNA and use these properties to act as a template for secondary components in a bottom-up approach towards self-assembly. The structure of an assembled DNA complex is directly and uniquely determined by the sequence of the DNA bases which can be designed and manipulated. These methods allows for a versatile and programmable way to control the structure and architecture of DNA nanostructures.

Pioneering work by Ned Seeman demonstrated how stable crossovers mimicking a Holliday junction could be used to create a rigid and well-behaved DNA structure known as the double crossover molecule (DX).[5] The DX consists of two helices of DNA that are held together by two stable crossovers. The termini of each of the four helices contained in this molecule have sticky ends which provide an easy way to design interactions between different DX tiles. By creating matching sticky ends on different DX tiles, one gives the molecules the ability to self-assemble into a rigid periodic two-dimensional array of DNA.[6] A variety of complex DNA nanostructures using or extending the DX motif have been designed to create structures such as arrays, nanogrids,

triple crossovers (TX), three- and six-pointed stars, and DNA origami.[1, 7-12] The use of DNA for the creation of 3-dimensional structures has been demonstrated as well.[13, 14]

Although it is an ideal material for the creation of designed nanostructures, DNA itself lacks the functionality desired for applications such as novel sensors, electronic devices, or enzymatic arrays. In order to be useful as an architecture for molecular assembly, a way to recruit molecular components to the DNA motif is necessary so that the self-assembly of the DNA molecules would also serve as a template for the assembly of secondary molecular components.

Previous work in this field has focused on the use of chemical or structural modifications of the DNA template in order to attach or recruit proteins or nanoparticles.[1, 15-21] A variety of methods have been previously used to attach molecules to specific tiles in a DX array. These methods include the incorporation of an additional dsDNA hairpin into specific tiles which projects perpendicular from the plane of the array,[6, 22] the use of biotin-labelled DNA to recruit streptavidin,[6, 8, 23] conjugation of DNA to gold nanoparticles,[15, 24] insertion of the thrombin-binding aptamer and subsequent recruitment of thrombin,[17] insertion of DNA aptamers that can be recognized by single chain antibodies,[16] and the recognition of a DNA-peptide fusion by an antibody.[21] All of these methods however, rely upon the covalent modification of the DNA modules prior to array assembly to allow visualization.

Pyrrole-imidazole polyamides are a class of oligomers that bind with high specificity to the minor groove of DNA.[25, 26] They can be programmed to target a broad repertoire of DNA sequences. Unlike the previously discussed methods used to attach molecules to DNA nanostructures, polyamides do not require any previous modification

of the target DNA. The highly programmable nature of pyrrole-imidazole polyamides make them particularly attractive for targeting specific DNA sequences.[26, 27]

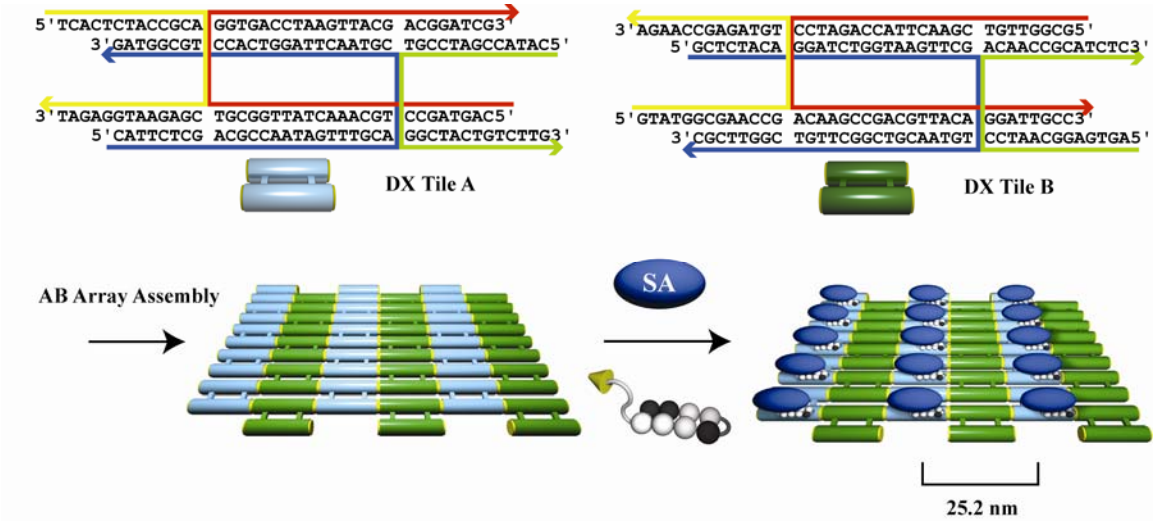

Figure 2.1 A DX **AB** array of DNA modified by a polyamide-biotin conjugate and streptavidin.

The research in this chapter demonstrates that polyamides are capable of accessing and binding to a DNA nanostructure in a sequence-specific manner. In addition, the ability to use different polyamide conjugates to target unique sites on a DX-array made up of *four* tiles and to arrange proteins into distinct spatial patterns is shown. Finally, analogous efforts to the polyamide-based approach were made with DNA binding proteins and are outlined as well.

2.2 Results and Discussion

Affinity Cleavage on DX Molecules

In order to first ascertain whether the DNA in a DX nanostructure was accessible to polyamides, a series of affinity cleavage experiments were done. Polyamide-EDTA-Fe conjugate **1-EDTA** was programmed to target a single DNA sequence 5’ -WGGWCW -

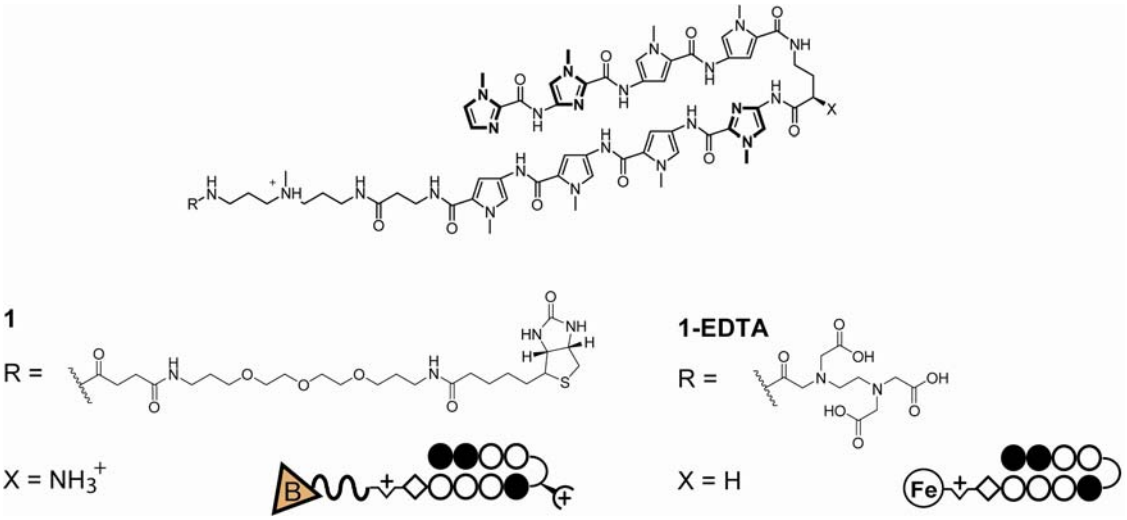

Figure 2.2 Structures and ball-and-stick models of polyamide conjugates **1** and **1-EDTA**. (Filled circle = N-methylimidazole, empty circle = N-methylpyrrole, diamond = β-Ala, half-diamond with plus sign = diamino-N-methyl-dipropylamine, half-circle = γ-aminobutyric acid, half circle with plus sign = NH_3^+, B in an orange triangle = biotin).

3' [28] embedded in a previously characterized DX tile.[6] (Figure 2.2) One DNA strand was radiolabeled and combined with the three unlabelled strands to form tile **A**. The DX tiles formed with over 90% purity when analyzed by gel electrophoresis and were found to be stable under the reaction conditions used for affinity cleavage. (see Experimental Details Figure 2.19) Cleavage occurred at the predicted binding site and in the expected orientation. (Figures 2.4 and 2.5) The observed cleavage pattern has a 2 – 3 bp 3' shift which is characteristic of minor groove binding polyamides.[29, 30] Additional affinity cleavage experiments were done using similarly designed DX tiles where the binding site was placed in different locations and orientations along the DX molecule. (Figures 2.3 – 2.5) In all cases cleavage was observed in the expected location, demonstrating the ability of hairpin polyamides to address a variety of locations along a single DX module. The sole exception was when the binding site spanned the crossover region in DX-Junction. No cleavage was observed, indicating that the structural features present prevent polyamide binding at this location.

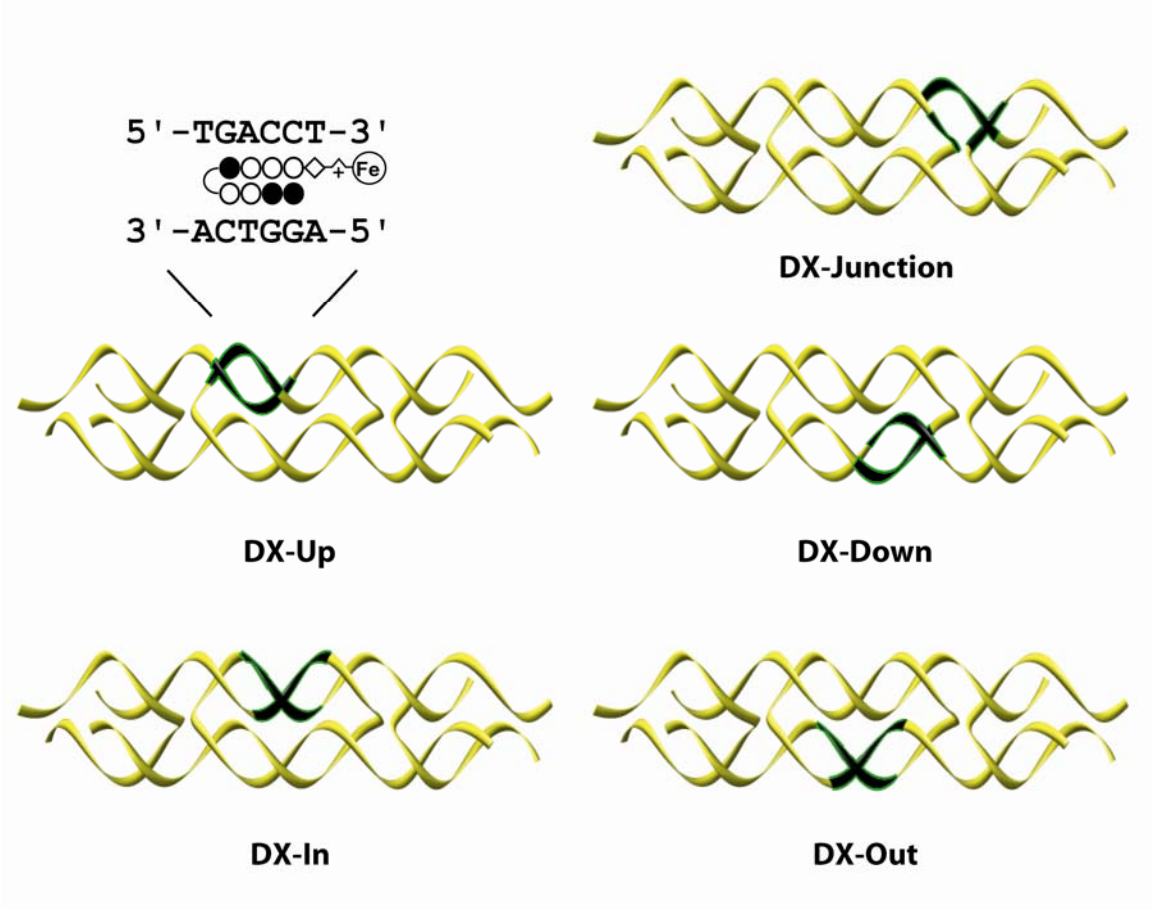

Figure 2.3 Models of five DX-tiles used to study various binding orientations. Each DX consists of four individual DNA strands shown in yellow with the 6 bp polyamide binding site shown in black. The designations of each molecule refer to how the minor groove of each binding site, and thus the polyamide are situated.

(A)

Labelled Strand

DX-Up (Tile A)	5' - CAT TCT CGA CGC CAA TAG TTT GCA CGT AAC TTA GGT CAC CTG CGG TAG - 3'
DX-In	5' - CAT TCT CGA CGC CAT AAC TAA GCA CGT ATA GGT CAT TGC CTG CGG TAG - 3'
DX-Down	5' - CAT TCT CGA CGC CTT TAG GTC ACA CGT AAC TTC ATT TGC CTG CGG TAG - 3'
DX-Out	5' - CAT TCT CGA CGC TAG GTC ACA GCA CGT AAC TTC ATT TGC CTG CGG TAG - 3'

(B)

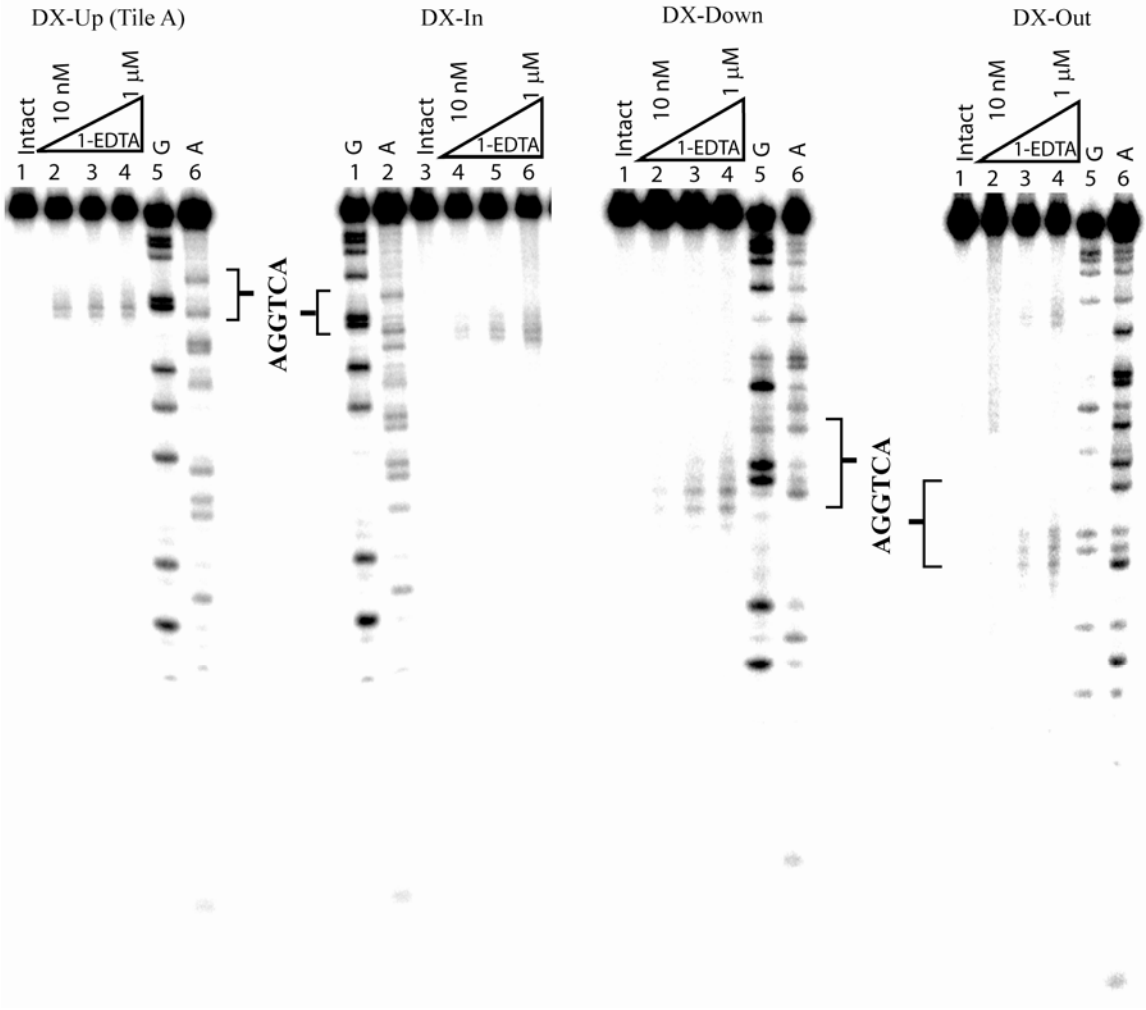

Figure 2.4 Affinity cleavage gels. (a) List of DNA sequences for the labeled strand in each DX tile. The binding site for **1-EDTA** is highlighted in red. (b) Affinity Cleavage on each DX tile. Strand 2 of each DX was ^{32}P-radiolabeled and affinity cleavage was performed using polyamide **1-EDTA**; the complex was then denatured and visualized by gel electrophoresis. The target polyamide binding site is indicated by brackets. A and G sequencing lanes are shown for each gel. Polyamide **1-EDTA** was added at concentrations of 10 nM, 100 nM, and 1 µM. Intact lanes were not incubated with **1-EDTA**.

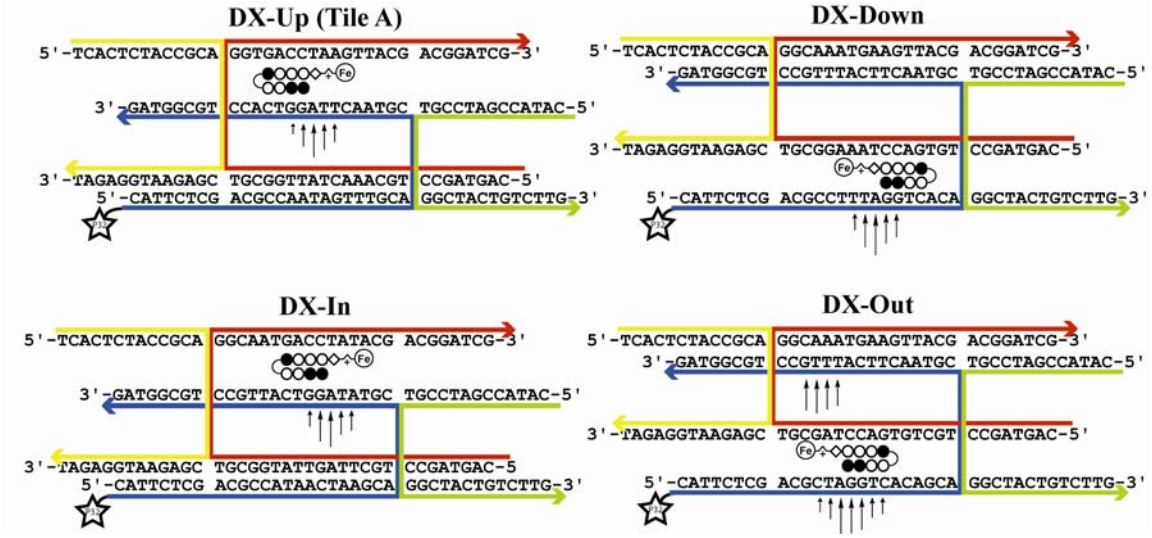

Figure 2.5 Affinity cleavage mapping. Representations of each DX tile with polyamide **1-EDTA**. Arrows represent the extent of cleavage at the indicated base position.

Biophysical Characterization of Polyamide Biotin-Conjugates

Having shown the ability of polyamides to bind to individual DX tiles, we next addressed whether a polyamide conjugated to biotin would be capable of recruiting streptavidin to DNA and therefore could be used as a labelling agent in AFM studies. DNAse I footprinting was used to ensure that attachment of a biotin did not interfere with a polyamide's ability to bind to DNA. (Figure 2.6) The ability of the polyamide-biotin conjugate to bind to streptavidin was then tested using an electrophoretic mobility shift assay (EMSA). DNA duplex was incubated with polyamide-biotin conjugate **1** as well as 10 μM streptavidin. (Figure 2.7) The streptavidin concentration was chosen to ensure a 10-fold excess of the protein compared to polyamide-biotin conjugate **1**. As the concentration of **1** is increased to 30 nM the naked DNA is shifted to a band of lower mobility corresponding to the tertiary complex containing polyamide-biotin, streptavidin and DNA.

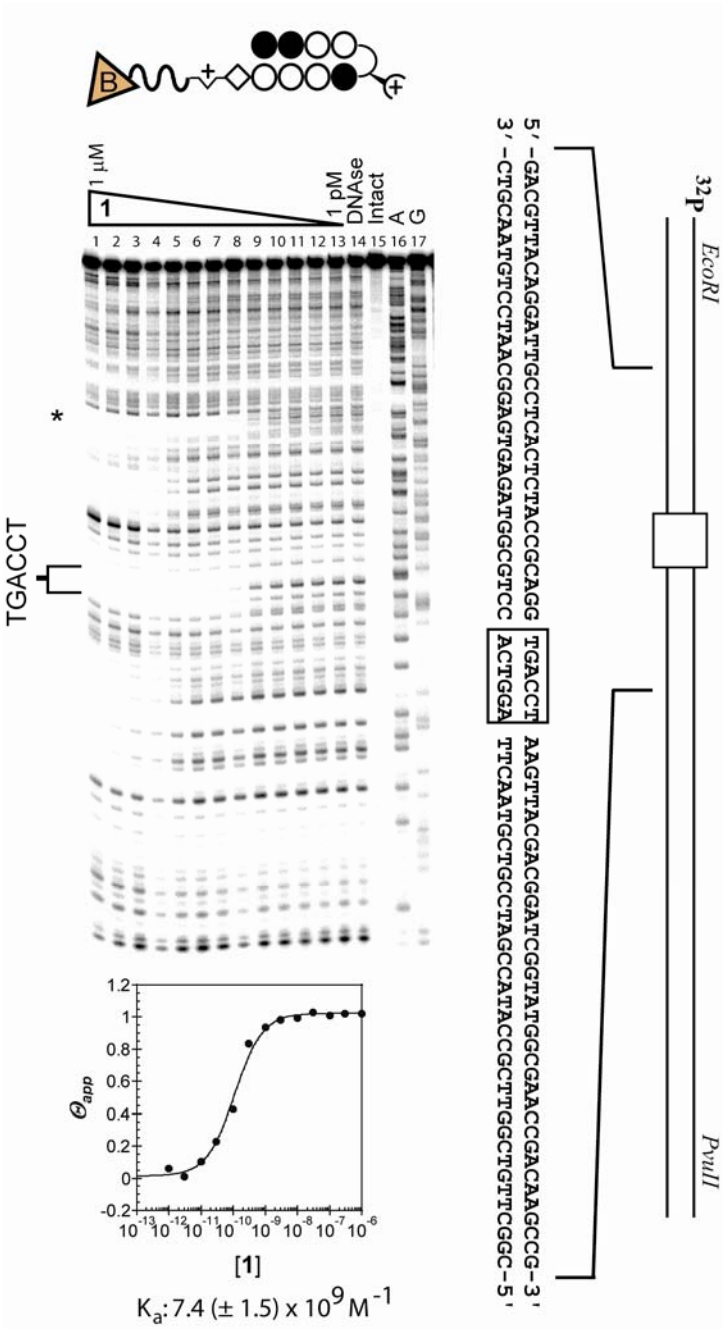

Figure 2.6 DNAse I quantitative footprinting with polyamide **1**. The 75 bp insert and schematic of the plasmid are shown with the binding site boxed. Lanes 1-13: 1 μM, 300 nM, 100 nM, 30 nM, 10 nM, 3 nM, 1 nM, 300 pM, 100 pM, 30 pM, 10 pM, 3 pM, and 1 pM of polyamide. Lane 14: DNAse standard. Lane 15: Intact DNA. Lane 16: A-sequencing lane. Lane 17: G-sequencing lane. Representative isotherm for polyamide binding and the calculated Ka value are shown. An additional binding site consisting of the sequence 5'-TGGTCA-3' is contained in the pUC19 plasmid and indicated by the asterisk above.

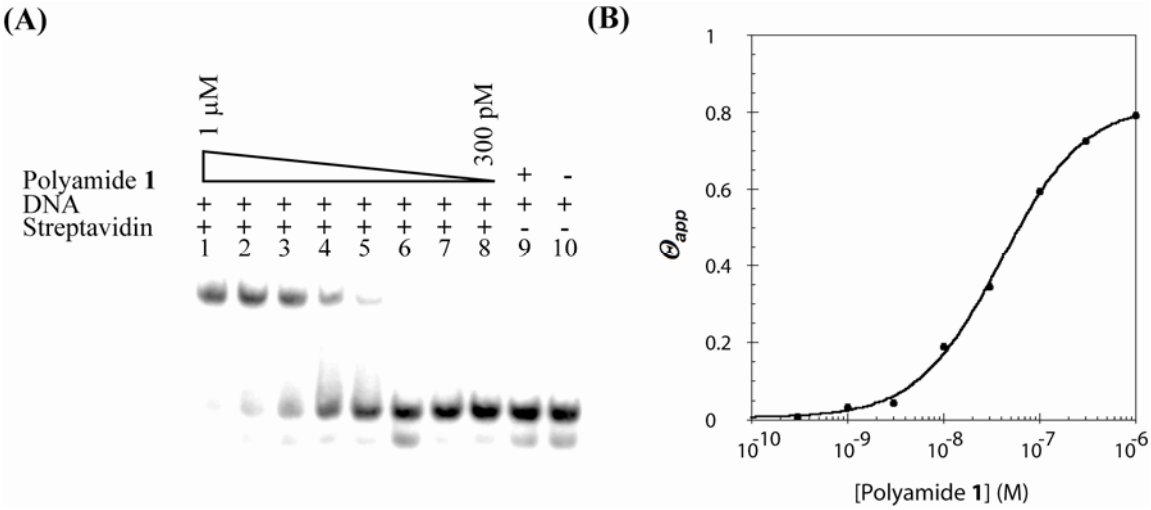

Figure 2.7 Electrophoretic Mobility Shift Assay of **1.** (a) EMSA for polyamide **1** in the presence and absence of 10 μM streptavidin. Lanes 1-8: 1 μM, 300 nM, 100 nM, 30 nM, 10 nM, 3 nM, 1 nM, 300 pM polyamide **1**. Lane 9: Control containing 1 μM polyamide **1** and DNA**.** Lane 10: Control containing only DNA. (b) Isotherm for streptavidin recruitment by polyamide **1**.

Atomic Force Microscopy of DX arrays with Polyamide-Biotin conjugates and Streptavidin

We then examined whether these complexes would be stable and visible using atomic force microscopy in the context of an **AB** type DX array. A new DX tile **B** which does not contain a match site for the polyamide was designed to form an **AB**-type array with the previously studied DX tile **A**. In our experiments we combined tile **A** and tile **B** to form an **AB** array in which every other tile has a polyamide binding site. As shown in Figure 2.8a, the DX tiles were capable of forming well-defined arrays when visualized by AFM.

Addition of polyamide **1** alone or streptavidin alone did not affect the formation or appearance of the DNA arrays. However, addition of polyamide-biotin conjugate **1** and streptavidin led to recruitment of the streptavidin to the **AB** DX array. (Figure 2.8cd) Incubation of polyamide **1** with tile **A** prior to array formation as opposed to after array formation, led to identical results indicating that order of addition is not critical. The

streptavidin molecules align with a regular spacing occurring between them as expected. The distance between alternating tile **A** in the **AB** array is expected to be 25.2 nm. (Figure 2.1) A section analysis shows that the average spacing between individual streptavidin molecules observed is 24.1 (± 1.6) nm, in agreement with the expected distance. (Figure 2.8e) For the purposes of this study it appears polyamide binding is relatively non-disruptive to array formation and stability.

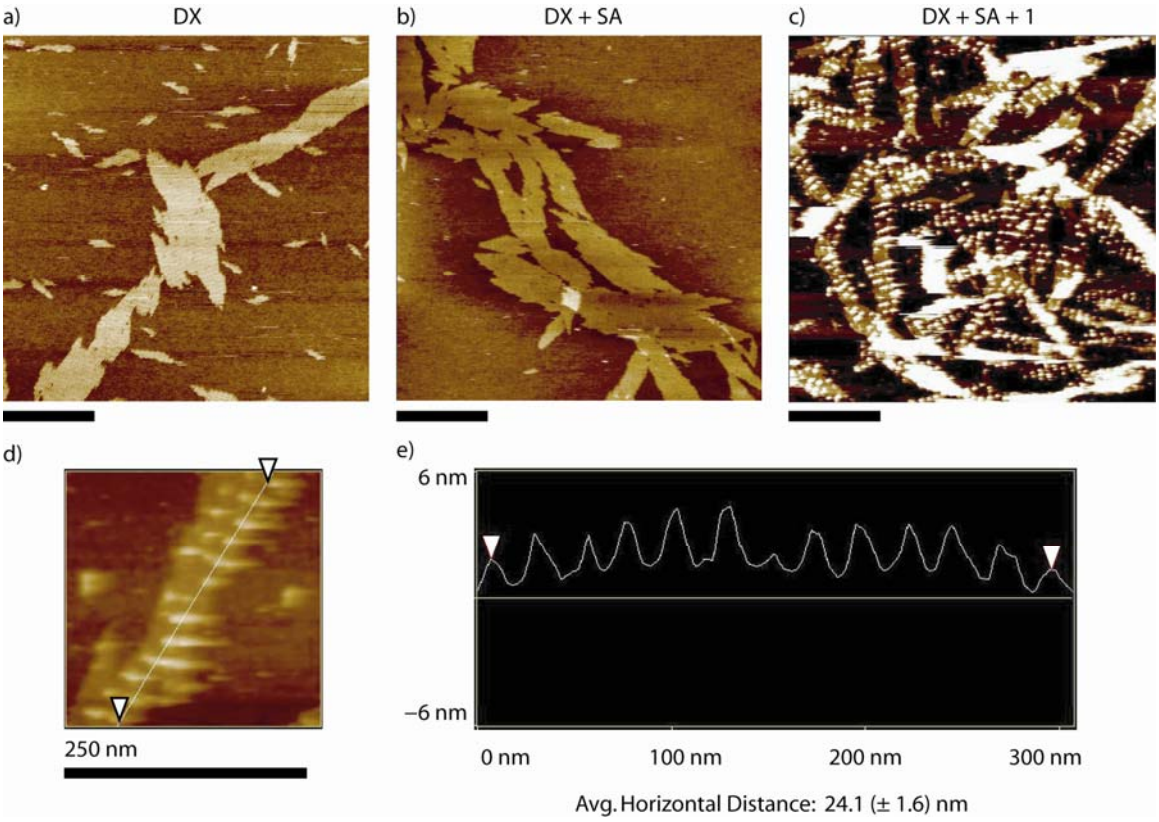

Figure 2.8 AFM images of combined DX tiles **A** and **B**. Scale bar is equal to 250 nm length for all images. (a) 100 nM DX array. (b) 100 nM DX array + 200 nM streptavidin. (c) 100 nM DX array + 200 nM streptavidin + 100 nM **1**. (d) Close up image of DX array with **1**. (e) Height profile along the path indicated in d. The average horizontal distance between peaks is shown.

Programmable Binding to Unique sites on an ABCD DX Array

To demonstrate how different polyamides could be used to target separate sequences in a DNA nanostructure, a new periodic array consisting of four DX tiles was designed as shown in Figure 2.10. Tiles **A**, **C**, and **D** were designed to contain a single

polyamide binding site embedded in them while tile **B** does not contain any match sites. Three polyamide-biotin conjugates **1**–**3** were synthesized with each polyamide core designed to address a unique tile in the array. (Figure 2.9) The binding affinity and specificity of each of these polyamide cores has been previously established. Polyamide **1**

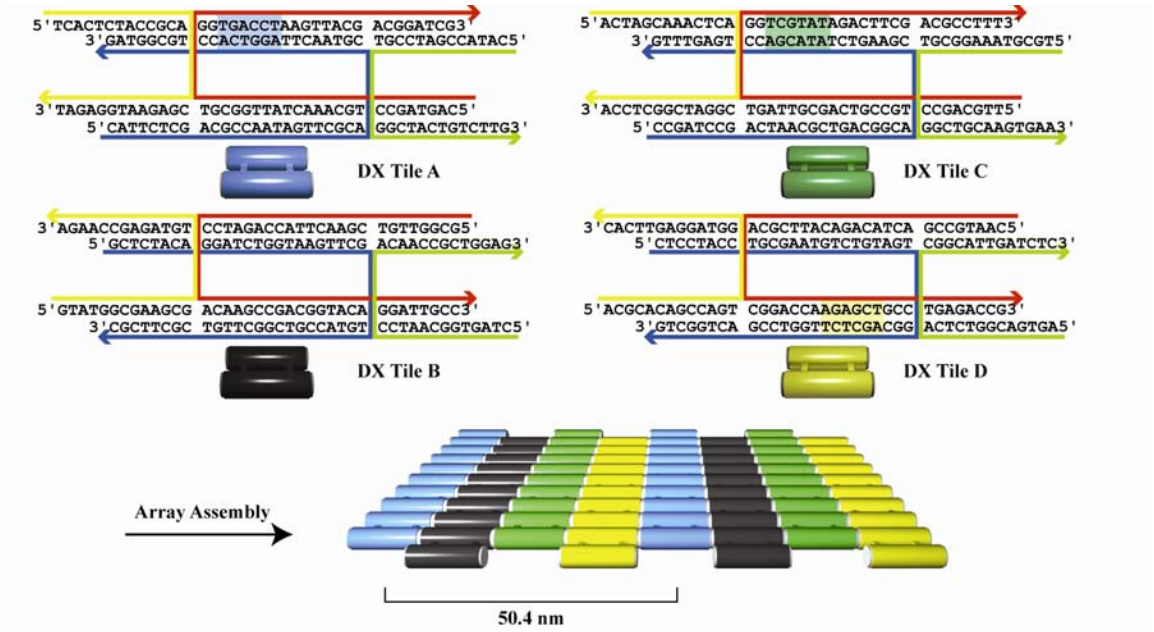

Figure 2.9 A DX array consisting of four tiles. Tiles **A**, **C**, and **D** contain a binding site for a different polyamide while tile **B** does not contain any binding sites. The predicted spacing between tiles in the array is shown.

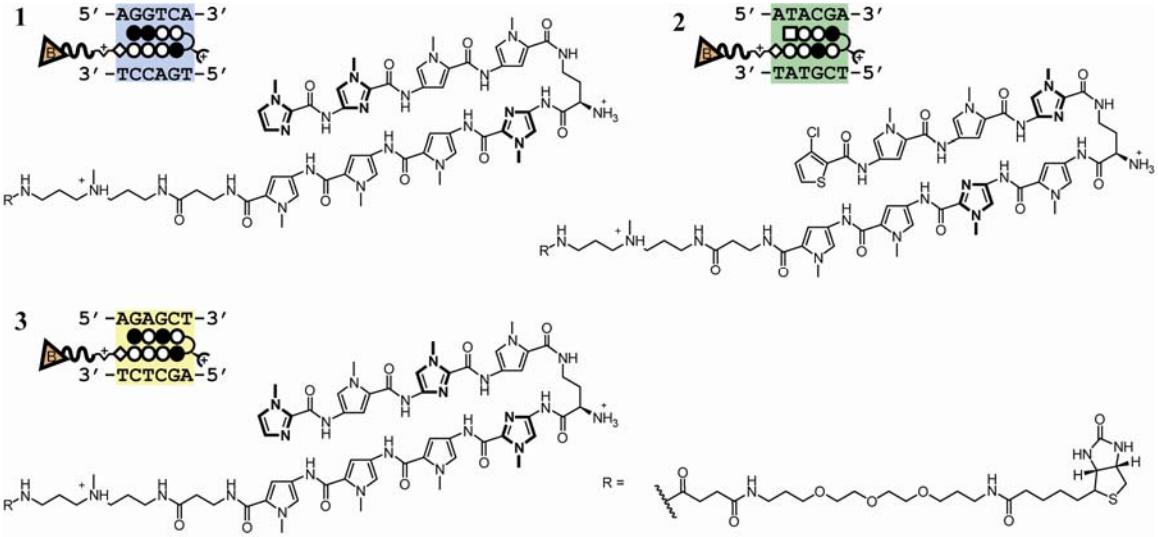

Figure 2.10 Structure and ball-and-stick models for polyamide-biotin conjugates **1**, **2**, and **3**. The DNA sequence that each is targeted to is shown. The colors correspond to the tile in the array that contains the target sequence as in Figure 2.9.

as discussed in the previous sections is designed to target 5'-WGGWCW-3',[28] polyamide **2** targets 5'-WTWCGW-3',[31] and polyamide **3** targets 5'-WGWGCW-3'.[27]

The four tile DX-**ABCD** was assembled and found by AFM to give 2-D arrays several μm long. Polyamides **1**, **2**, and **3** were then individually incubated with **ABCD** DNA array for 1 hr prior to AFM imaging. The concentration of DNA was 100 nM and the concentration of polyamide used was 200 nM for **1** and **3**, and 150 nM for **2** which showed a propensity to bind at additional sites at higher concentrations. Samples were diluted in half, 5 μL was spotted on mica and allowed to absorb for 1 min, followed by the addition of 2 μL of streptavidin (1 μM) for 1 min. As shown in Figure 2.11,

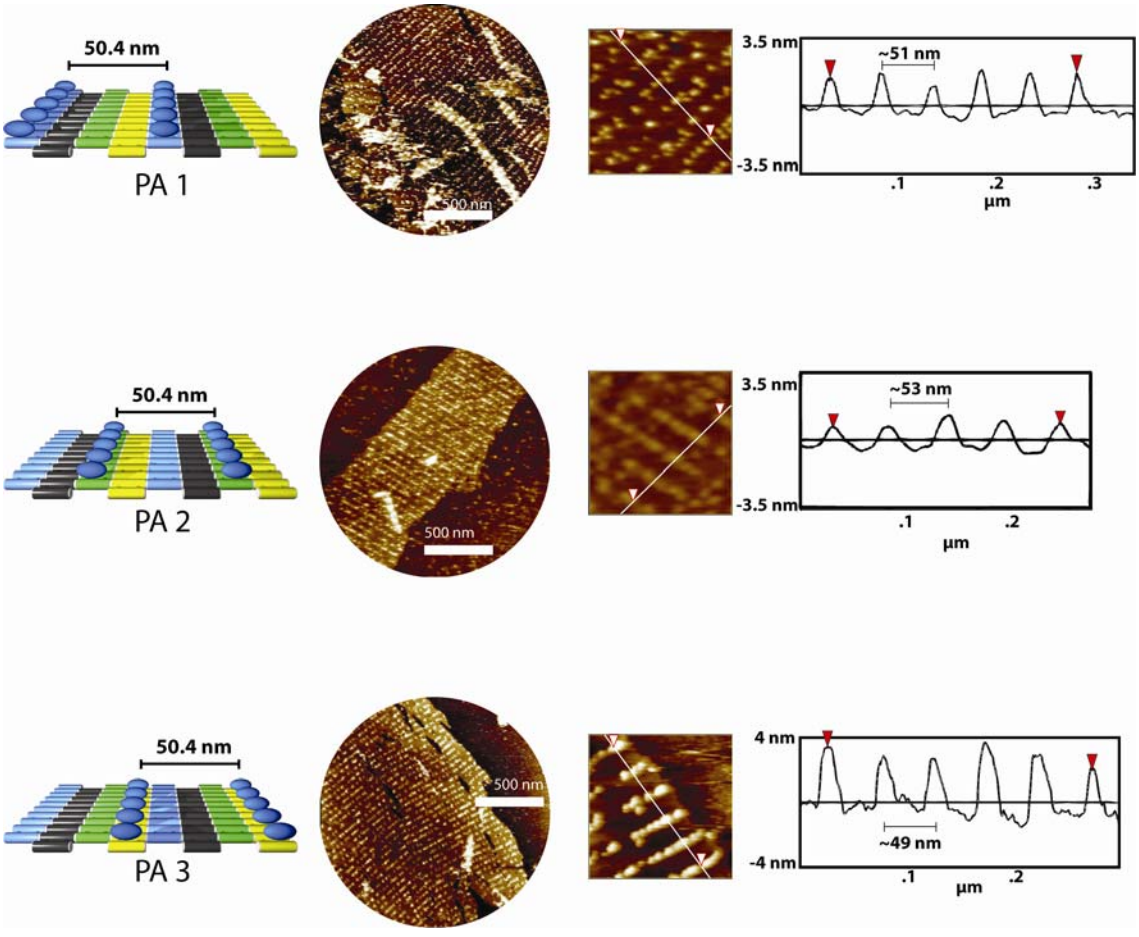

Figure 2.11 AFM images for individual polyamides incubated with **DX-ABCD** and streptavidin. Section analysis shows the height along the indicated path in the square image.

streptavidin recruitment is observed in all three cases. The average spacing for **1**, **2**, and **3** is ~51 nm, ~53 nm, and ~49 nm respectively, corresponding to a spacing of four tiles as expected in each case. As shown, these conjugates are capable of targeting unique locations on the DX array with high specificity and affinity.

Creating Unique Protein Patterns on an Individual DNA Template

In order to demonstrate the ability to target multiple sites on the array simultaneously, we incubated the array with both polyamides **1** and **3**. Since these polyamides target adjacent tiles **A** and **D** on the array, they should give "double-wide" columns of streptavidin. As shown in Figure 2.12, this is what we observe. The average spacing from the center of two adjacent peaks to the next two peaks is ~53 nm, similar to what was observed in the case of the individual polyamides and in agreement with what we would predict. We next incubated our array with polyamides **1** and **2** as before. In this case, every other tile in the array is targeted, so we expect to see stripes with a spacing of exactly half that previously observed for the individual polyamides. We observe a spacing of ~25 nm. As a final experiment, we incubated our arrays with a combination of polyamides **1**, **2**, and **3**. We observe binding at three sites, with a single unbound site between them as would be predicted. The average spacing between unoccupied sites is ~47 nm corresponding to a spacing of every four tiles. This shows the ability to not only simultaneously target multiple DNA sequences in an array, but also to create unique protein patterns on the same DNA template.

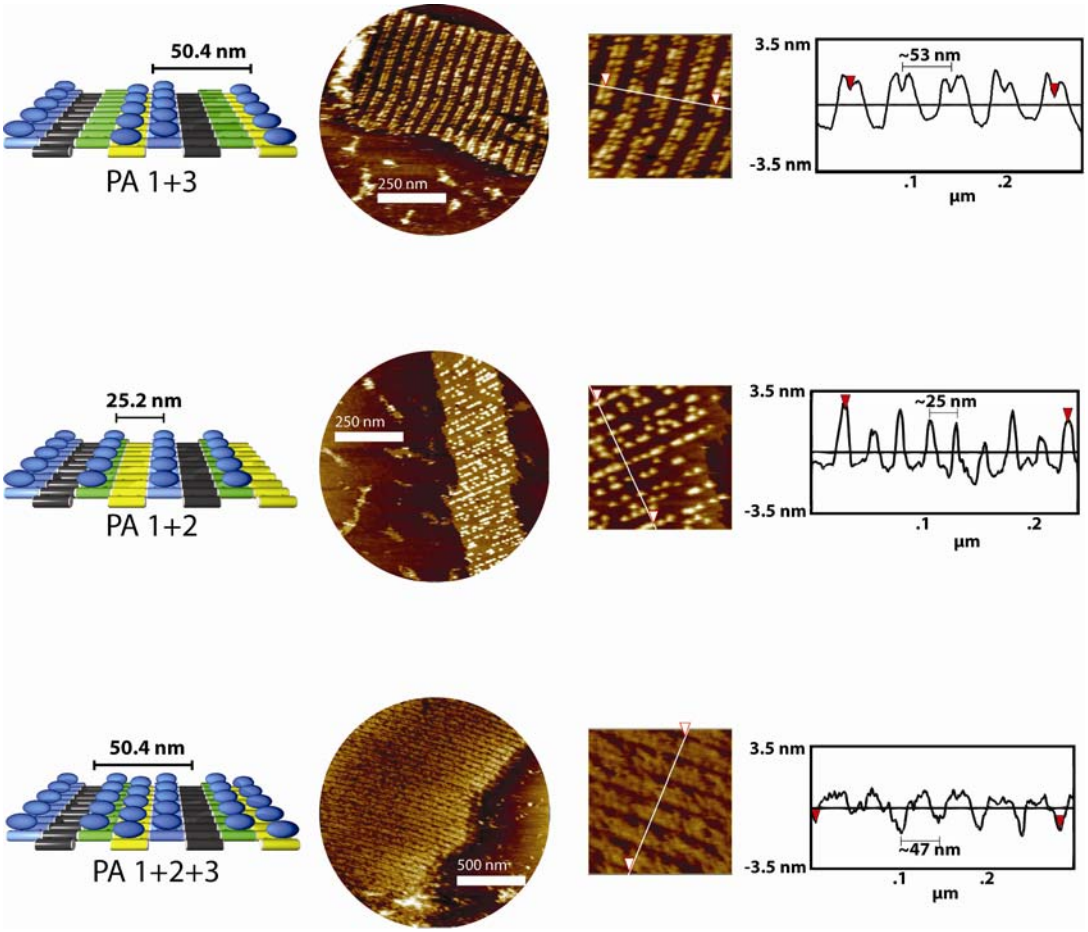

Figure 2.12 AFM images of **DX-ABCD** incubated with combinations of polyamides **1**, **2**, and **3**. Section analysis shows the height along the indicated path in the square image.

Targeting DNA Nanostructures with DNA-binding Proteins

The previous sections demonstrated the ability of polyamides to effectively bind to specific sequences in DNA nanostructures and to create novel protein patterns on a DNA template. A natural question that arose was whether DNA nanostructures could be targeted directly via a DNA-binding protein. AFM imaging of proteins bound to DNA has been demonstrated for several proteins including MutS and Sp1.[32-35] As previously discussed, the incorporation of DNA aptamers into DNA nanostructures has been used to recruit proteins to the assembled surface.[36] In this case, the molecular recognition is programmed into the DNA, and not performed by the protein. To date, one of the only

examples of recognition of a 2-dimensional DNA array by a protein has been the use of RuvA which naturally binds to Holliday junctions.[37] It was shown that upon addition of RuvA to an array made up of four-armed immobile Holliday Junctions, the structure of the array switched from a Kagome lattice into a square-planar lattice, consistent with the known architectural role of RuvA during branch migration.

Zinc finger proteins are well-known for their ability to bind to DNA,[38-42] and have affinities and specificities similar to polyamides.[43] We decided to examine whether the zinc finger binding protein EGR-1 would be capable of binding to a DX array that contained its target binding sequence. EGR-1 is a 543 amino acid, 57.5 kDa protein that contains three zinc finger domains and binds to the 9 bp sequence 5'-GCGTGGGCG-3'.

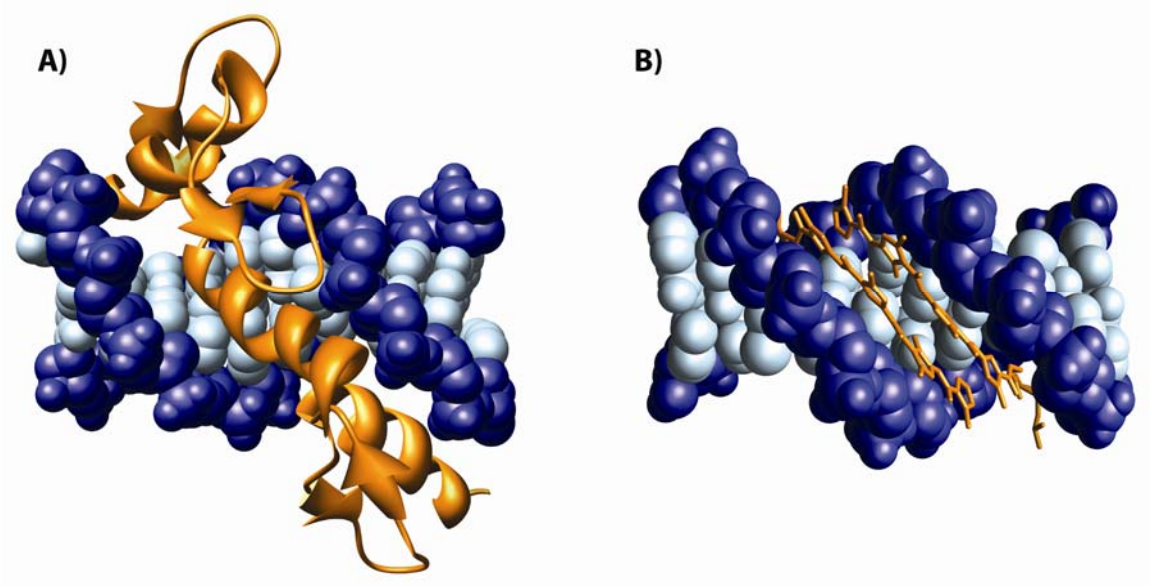

Figure 2.13 Crystal structures of zinc finger protein Zif268 and a polyamide bound to DNA. A) The zinc finger protein Zif268 binds in the major groove of DNA. Zif268 is colored orange. B) A polyamide bound in the minor groove of DNA. The polyamide is colored orange. (PDB codes 1AAY and 1CVY)

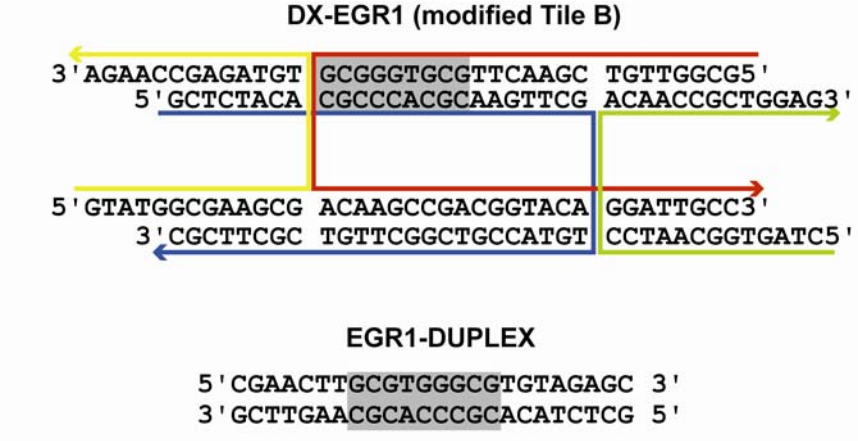

Figure 2.14 Tile DX-EGR1 and EGR1-DUPLEX. The tile was designed by modifying tile **B** of the **ABCD** array. A duplex containing the same sequence as DX-EGR1 was also constructed as a positive control for gel shift experiments.

To examine whether EGR-1 could bind to a DX-array, a new tile containing its 9-bp binding site was created. As shown in Figure 2.14, the binding site was placed in a location where the major groove was deemed most likely to be accessible for binding. The binding site 5'-GCGTGGGCG-3' was incorporated into Tile **B** from the **ABCD** DX array so that the ability to perform recruitment on tiles **A**, **B** and **D** with polyamides would be retained. A duplex containing the identical binding sequence was also used as a positive control for gel shift experiments.

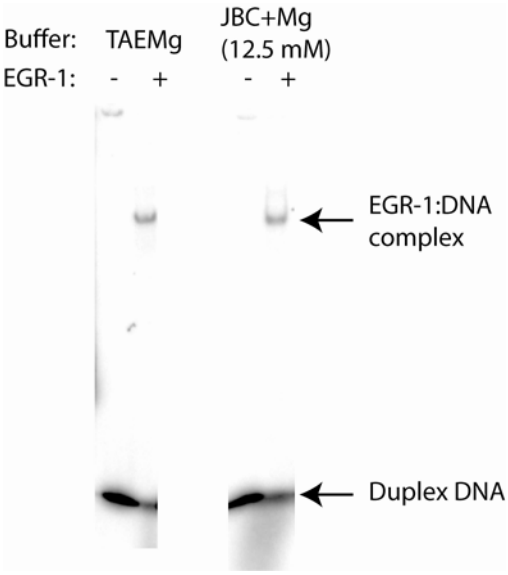

Figure 2.15 Gel shift assay for EGR-1 binding to EGR1-DUPLEX DNA. A shift in mobility showing the formation of the EGR-1:DNA complex is observed.

To first examine whether EGR-1 could bind to a DX-array, a series of gel shift experiments were done. EGR-1 was incubated with radiolabeled DNA EGR1-DUPLEX for 1

hr. TAEMg buffer (40 mM Tris pH 7.5, 1 mM EDTA, 12.5 mM $Mg(OAc)_2$, 20 mM acetic acid), used in the previous sections for the formation of DX arrays, as well as a set of published gel shift buffer conditions (JBC Buffer: 10 mM Tris pH 7.5, 1 mM EDTA, 50 mM NaCl)[44] were used in these assays. The JBC buffer when noted was also supplemented with 12.5 mM $Mg(OAc)_2$ as magnesium is considered necessary for DNA nanostructure formation. 1 mM DTT, 20 μM $ZnSO_4$, and 5% glycerol were also added. After incubation, the sample was run on a .5× TBE gel. The results are shown in Figure 2.15. As shown, the addition of EGR-1 to the samples resulted in a clear shift indicative of the protein:DNA complex forming. Analogous experiments using a 1x TAEMg gel for electrophoresis gave similar results.

Gel shifts were next attempted on the DX-EGR1 tile. The gel shifts were performed as before except the labeled DX tile was used instead of the duplex. These results are shown in Figure 2.16. As shown, a clear shift was observed when JBC buffer was used. The addition of magnesium to the buffer resulted in diminished binding although some complex was still observed.

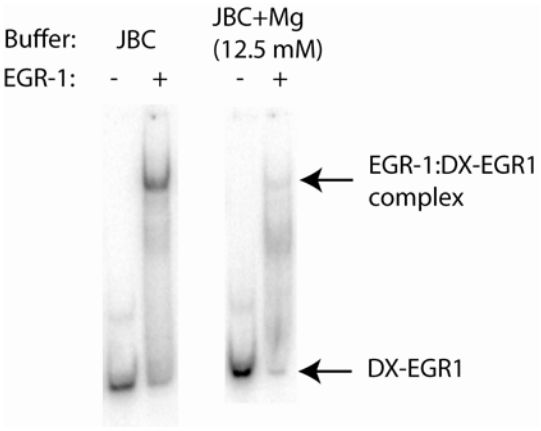

Figure 2.16 Gel shift for EGR-1 binding to DX-EGR1. The formation of the protein:DNA complex is partially inhibited by the addition of 12.5 mM $Mg(OAc)_2$.

Next, AFM experiments were done to examine whether EGR-1 would be capable of binding to its site in the context of a 2-dimensional array. Tile-EGR1 was incubated

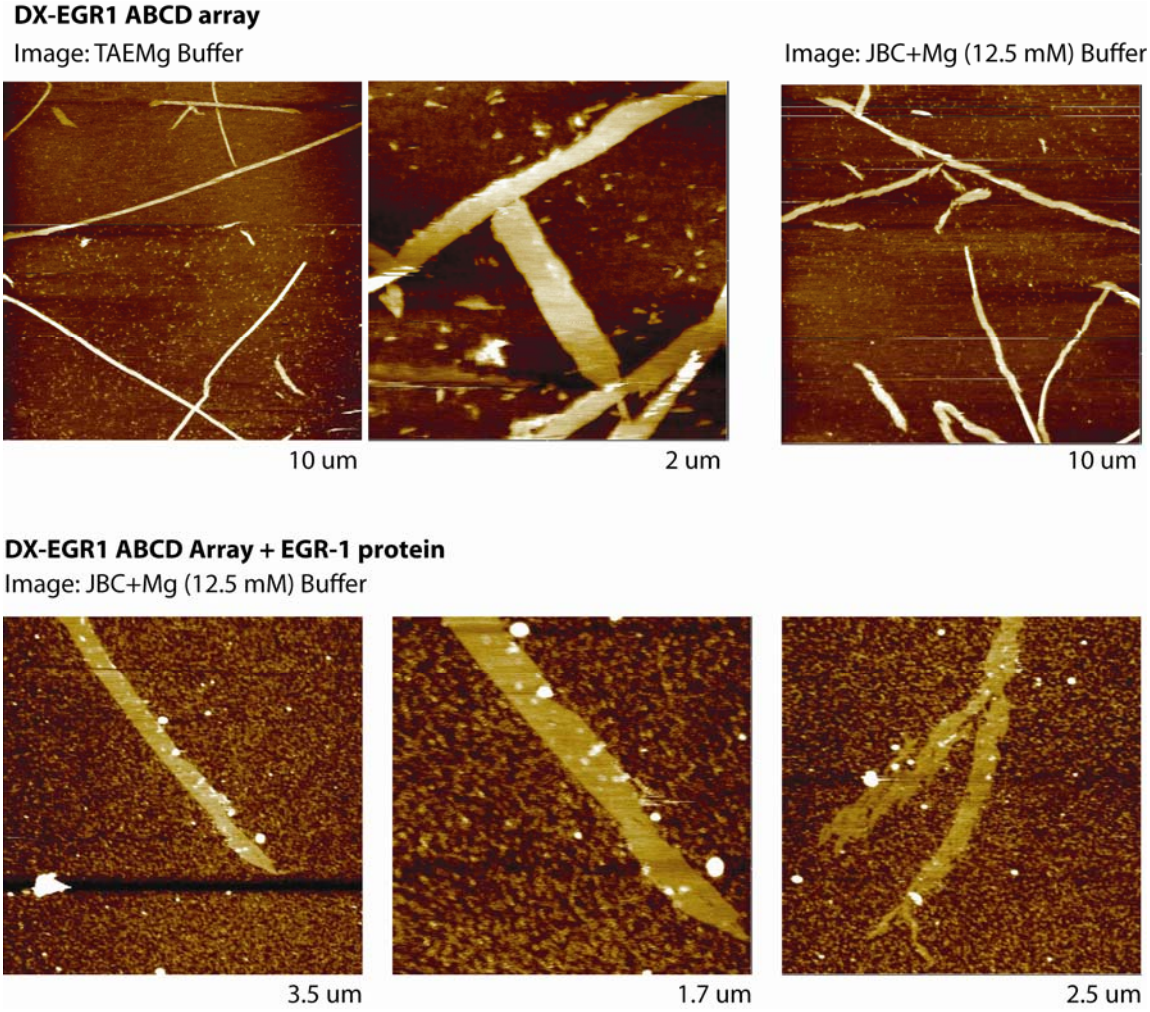

Figure 2.17 AFM images of EGR-1 and DX arrays. Incubation with EGR-1 protein led to very low levels of possible recruitment on the arrays.

with tiles **A**, **C**, and **D** to form a four tile array. The arrays formed in both TAEMg and JBC+Mg buffer, although the long narrow nature of the observed arrays is associated with the formation of nanotubes. The formation of nanotubes is thought to be a result of slight differences in the overall curvature of the DX tiles. As expected, JBC buffer without the addition of magnesium did not result in array formation. The arrays were incubated with EGR-1 and the results are shown in Figure 2.17. As shown, there was very little recruitment of EGR-1 to the arrays. In addition, a large amount of background

protein binding was observed. This led us to re-examine the source of EGR-1 protein which had been used for these studies. The EGR-1 sample purchased from Axxora was a partially purified protein extract. To determine the purity of the protein, non-denaturing polyacrylamide gel electrophoresis followed by Coomassie Blue staining was done. As seen in Figure 2.18, a large number of additional proteins are present in the sample. The presence of so many additional proteins is the probable cause for the nonspecific protein binding to the mica during the AFM experiments.

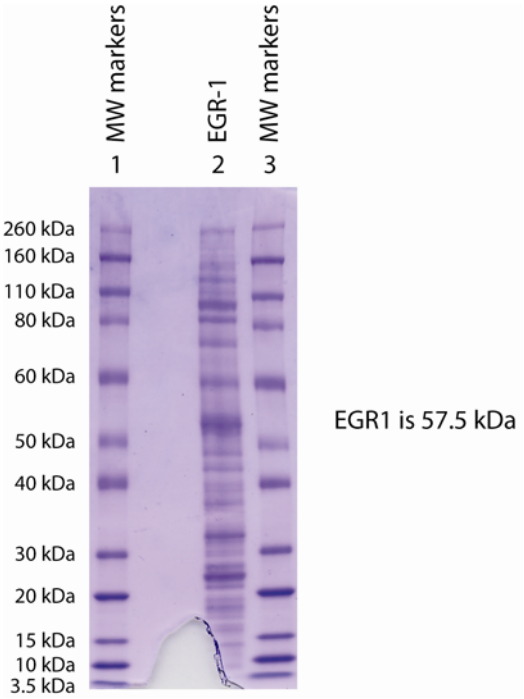

Figure 2.18 Non-denaturing Polyacrylamide Gel Electrophoresis of EGR-1 Protein Sample. Lanes 1 and 3: Molecular Weight Markers. Lane 2: EGR-1. A large number of proteins are present in the sample.

As a final set of experiments, a purified fusion protein, containing the 100 amino acid C-terminal DNA-binding portion of EGR-1 fused to Glutathione-S-transferase (GST) was purchased from Abnova. However, the fusion protein failed to function in either the AFM or gel shift assays. As such, it remains unclear whether the fusion protein is still functional for DNA binding.

2.3 Conclusions

In our experiments, a polyamide-biotin conjugate was able to recruit streptavidin molecules to one tile (**A**) of an **AB** DX array, effectively labeling individual tiles. In addition, we have demonstrated that polyamide-biotin conjugates are capable of

addressing specific elements in a multi-component DX array. We are able to organize streptavidin molecules into distinctly different patterns on the same DNA template, solely by changing the polyamide cores that are used. The synthetic ease in creating polyamides, and the existence of a library of well characterized solutions to target a wide variety of DNA sequences makes these conjugates ideal for arranging proteins at a variety of sites. In addition, the ability to address an array without requiring prior covalent modification increases the flexibility and usefulness of DNA templates, as they could serve to make a variety of increasingly complex assemblies. Nanoparticles, proteins, or other biomolecules could be conjugated to or recruited by polyamides, allowing them to be targeted to specific tiles in an array. Polyamide conjugates would act as a sequence specific glue or molecular staple that allows the self-assembly properties of DNA to be transferred to complicated functional molecular assemblies.

The work using the DNA-binding protein EGR-1 illustrates several of the challenges involved in creating complex assemblies. The successful gel shift experiments using a single DX tile show that EGR-1 can recognize and bind to its target sequence when embedded in a DNA nanostructure. The small amount of protein recruitment observed in the AFM experiments, and the successful gel shift indicate that this approach towards targeting DNA nanostructures has merit. Future experiments will be needed to determine whether the protein can still access its binding site in the context of the fully formed 2-dimensional array. It is worth noting that the Yan group, in addressing sites on a DNA origami nanostructure found distinct differences in binding affinity at different locations on the nanostructure's surface.[45] Locations nearest the edges were significantly better for binding than sites located towards the center of the structure. It will be

interesting to determine whether this is the case for proteins binding to DNA arrays as well. Given the relatively large size of proteins when compared to a DNA oligomer or a polyamide, it is likely that proteins may have even greater positional preferences. One hypothesis that could explain the lack of binding in the AFM studies is that the protein is too sterically hindered to access to the surface of the DX when it is incorporated into a 2-dimensional array but is not restricted in the gel shift as it is only binding to a single tile in solution.

Future experiments are needed to ascertain whether the impurities in the sample were responsible for the lack of binding in the AFM experiments. It is possible that the presence of a large number of other proteins could have interfered with binding. In addition, the concentration of EGR-1 in the samples is not known. It may be that the protein is not present in sufficiently high concentrations to bind to all of the sites present in the array. Future experiments using purified EGR-1 would eliminate the difficulties that occur when using a cellular extract and have the advantage of knowing the amount of protein being added to each sample.

It is interesting to compare the successful use of polyamides for targeting DNA nanostructures when compared to the difficulties experienced with EGR-1. The purity issues that were encountered with EGR-1 are non-existent for polyamides which can be rapidly synthesized using solid phase methods and purified by HPLC. The ability to rapidly generate polyamides to target virtually any desired DNA sequence is also a considerable advantage. While a large number of zinc finger motifs to target a variety of sequences are known,[38, 39, 42] the expression and purification of these constructs is non-trivial. Polyamides were functional in all of the buffers examined including those used for

DNase footprinting, affinity cleavage, and AFM studies. In contrast, many proteins are highly sensitive on buffer conditions for proper folding and function. The need for high magnesium concentrations to stabilize DX and other nanostructures may prove to be problematic for the use of DNA-binding proteins as the addition of 12.5 mM $Mg(OAc)_2$ in the gel shift buffer was detrimental to EGR-1 binding. Finally, it should be noted that the smaller size of polyamides when compared to proteins may allow them to bind to their target sites in a 2-dimensional array whereas proteins may be unable to access more sterically restricted sites. DNA nanostructures, such as nanogrids, that contain significantly more spacing between the DNA elements in the array, may be superior to DX arrays, which are tightly packed in comparison, for these purposes.

In conclusion, the experiments with EGR-1 highlight both the challenges involved in creating complex assemblies at the molecular level, and also may of the advantages of our polyamide based approach. Future work towards creating assemblies should benefit from the approaches investigated here for the molecular recognition of DNA nanostructures.

.

2.4 Experimental Details

Abbreviations. *N,N*-dimethylformamide (DMF), *N,N*-diisopropylethylamine (DIEA), *rac*-dithiothreitol (DTT), *N*-[2-Hydroxyethyl]piperazine-N'-[2-ethanesulfonic acid] (HEPES), 2-(1H-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU), trifluoroacetic acid (TFA), Trishydroxymethylaminomethane (TRIS), ethylenediaminetetraacetic acid (EDTA).

Materials. Boc-*β*-Ala-PAM resin was purchased from Peptides International. Trifluoroacetic acid (TFA) was purchased from Halocarbon. Methylene Chloride (DCM) was obtained from Fisher Scientific and *N,N*-dimethylformamide (DMF) from EMD. Ethylenediaminetetraacetic dianhydride was purchased from Aldrich. EZ-Link TFP-PEO_3-Biotin was purchased from Pierce. Streptavidin was purchased from Rockland. EGR-1 protein extract was purchased from Axxora. The recombinant protein EGR-1-GST was purchased from Abnova. All DNA oligonucleotides were purchased from Integrated DNA Technologies. DTT and Ultra Pure TRIS were purchased from ICN. Magnesium Acetate 4-hydrate was obtained from J.T. Baker. Ferrous Ammonium Sulfate was purchased from Fisher Scientific. Magnesium Chloride 6-Hydrate was obtained from Mallinckrodt. [γ-^{32}P]-adenosine-5'-triphosphate ($\geq$ 7000 Ci/mmol) was obtained from MP Biomedicals. Calf thymus DNA was from Amersham and all enzymes were obtained from Roche. Water (18 MΩ) was purified using a aquaMAX-Ultra water purification system. Biological experiments were performed using Ultrapure Water (DNase/RNase free) purchased from USB. The pH of buffers was adjusted using a Thermo Orion 310 PerpHect Meter. All buffers were sterilized by filtration through a Nalgene 0.2 μm cellulose nitrate filtration.

Polyamide Synthesis. Polyamide monomers were prepared as described previously.[46] Synthesis was performed using established protocols and all polyamides were characterized by MALDI-TOF and analytical HPLC.

1-EDTA: (MALDI-TOF-MS) $[M+H]^+$ calc. for $C_{69}H_{90}N_{25}O_{17}^+$ 1540.7, observed 1540.6

1: (MALDI-TOF-MS) $[M+H]^+$ calc. for $C_{83}H_{117}N_{28}O_{17}S^+$ 1809.9, observed 1810.0

2: (MALDI-TOF-MS) $[M+H]^+$ calc. for $C_{83}H_{114}ClN_{26}O_{17}S_2^+$ 1846.8, observed 1846.5

3: (MALDI-TOF-MS) $[M+H]^+$ calc. for $C_{83}H_{117}N_{28}O_{17}S^+$ 1809.9, observed 1810.2

Preparation of Labeled DX Tiles. In all cases 5' radiolabeling of 60 pmol of a single DNA strand was done using Polynucleotide Kinase. The labeled strand was then added to the three unlabelled strands in Affinity Cleavage Buffer (20 mM HEPES, 60 mM NaCl, 62.5 mM $MgCl_2$, pH 7.3) and a total volume of 20 μL. Several samples were made to titrate the labeled stand against the unlabelled strands in order to maximize formation of the four strand DX complex. The strands were heated to 95°C for 10 min and then allowed to cool slowly to room temperature over several hours. Purity was assessed by running 1 μL of the sample on a 6% polyacrylamide gel. The gel was then dried and exposed to a phosphor screen which was visualized using a Molecular Dynamics 400S Phosphorimager. In all cases the DX tile that was used was greater than 90% pure.

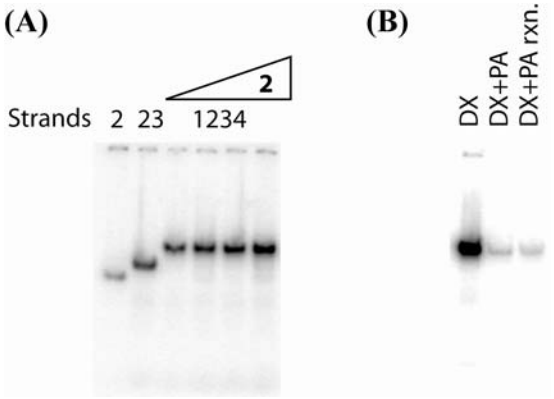

Figure 2.19 Formation and stability of DXs. (a) Representative native gel showing the formation of the four stranded DX-In tile. Lane 1: strand 2. Lane 2: strands 2 and 3. Lanes 3-6: .7, .9, 1.0, and 1.2 equivalents of strand 2 with 1 equivalent of strands 1, 3, and 4 (b) Representative native gel showing stability of the fully formed complex. Lane 1: DX-In. Lane 2: DX-In and 1 μM polyamide **1-EDTA** in affinity cleavage buffer. Lane 3: DX-Up and 1 μM polyamide **1-EDTA** in affinity cleavage buffer after the cleavage reaction has occurred.

Affinity Cleavage on DX complexes. Affinity Cleavage experiments were done following previously established protocols.[47] The total reaction volume was 50 μL. The polyamide was allowed to equilibrate with DX (12,000 cpm / lane) for one hour in Affinity Cleavage Buffer (see above). A solution of freshly prepared ferrous ammonium sulfate was added to a final concentration of 1 μM and allowed to sit for 30min. To initiate the reaction, DTT was added to a final concentration of 5 mM. The reaction was allowed to proceed for 40min and stopped by the addition of 1.25 μL of precipitation buffer (Glycogen 2.8mg / ml, 140 μM bp calf thymus DNA). The DNA was then isolated by ethanol precipitation and run on an 8% polyacrylamide denaturing gel.

Electrophoretic Mobility Shift Assay (EMSA). For the polyamide:streptavidin EMSA experiments 60 pmol of DNA strands DUPLEX1 and DUPLEX2 were first annealed by heating to 95°C for 10 min and allowing them to cool slowly to room temperature over several hours. Using Polynucleotide Kinase they were then 5' radiolabeled. The labeled DNA (3,000 cpm/lane) was then incubated with polyamide for 1 hr in TAEMg buffer followed by incubation with streptavidin for 30 min. The total volume for the reaction was 10 μl which was then run on a 6% polyacrylamide gel and imaged. The EGR-1 experiments were performed using a modified version of a reported protocol.[44] First, DX-EGR1 and EGR1-DUPLEX were radiolabeled as described above. The labeled DNA (5,000 cpm/lane) was then incubated with 1 band forming unit of EGR-1 protein extract, 1 mM DTT, 20 μM $ZnSO_4$, and 5% glycerol for 1 hour in either TAEMg or JBC (10 mM Tris pH 7.5, 1 mM EDTA, 50 mM NaCl) buffer. In certain noted cases 12.5 mM $Mg(OAc)_2$ was also added to the JBC buffer. The reaction volume was 20 μL. After

incubation, 10 μL of each sample was run on a 6% acrylamide .5X TBE gel for 1.5 hr at 180 V and imaged.

Affinity determination by quantitative DNase I footprinting. Reactions were carried out in a volume of 400 μL in aqueous TKMC buffer according to published protocols.[47] Standard molecular biology techniques were used to insert a 75bp DNA sequence into the *Bam*HI/*Hind*III restriction site of pUC19.[48] This plasmid was used to generate a 5' ^{32}P labeled 283bp *Eco*RI/*Pvu*II restriction fragment which was used for all footprinting experiments. Developed gels were imaged using storage phosphor autoradiography using a Molecular Dynamics 400S Phosphorimager. Equilibrium association constants were determined as previously described.[47]

AFM Sample Preparation. Individual DX tiles **A** and **B** were first annealed by mixing equimolar amounts of each of four strands in TAEMg Buffer (40 mM Tris-HCl (pH 8.0), 20 mM acetic acid, 1 mM EDTA, 12.5 mM magnesium acetate) and heating to 95°C for 10 min at a concentration of 200 nM. They were then allowed to cool slowly over several hours to room temperature. Equal amounts of DX tile **A** and **B** were then mixed at a final concentration of 100 nM. Samples were heated to 50°C and allowed to cool slowly over 10 – 12 hr to room temperature. Polyamide **1** was added to the solution of DX arrays and allowed to equilibrate for 1 hr at a concentration of 100 nM, followed by addition of streptavidin to a final concentration of 200 nM which was incubated with the sample for 30 min. In certain noted cases the polyamide was incubated with a single tile prior to the 50°C annealing step, rather than with the already formed array. 5 μL of the sample was

spotted on freshly cleaved mica and allowed to absorb for 1 min. Imaging was done using a DI Multimode Atomic Force Microscope.

For **DX-ABCD** the array was formed in two steps. Individual tiles **A**, **B**, **C**, and **D** were first annealed by mixing equimolar amounts of each of the four input strands at 1 μM concentration in TAEMg Buffer. Each sample was heated to 95°C for 10 min and allowed to cool slowly over several hours to room temperature. The four tiles were then combined, heated to 45°C and allowed to cool to RT over approximately 12 hrs. For all of the experiments, polyamide was incubated with DX-**ABCD** (100 nM) for 1 hr prior to imaging. The polyamide concentration was 200 nM for **1** and **3**, and 150 nM for **2**. After the incubation, the sample was diluted in half which led to cleaner AFM images. 5 μL of sample was spotted on freshly cleaved mica and allowed to absorb for 1 min. 2 μL of 1 μM Streptavidin was then added to the sample for 1 min. Calibration for the distance measurements was done by averaging four measurements of individual tiles in an untreated ABCD array. The reported spacings for each of the polyamides with streptavidin were calculated as the average distance between peaks in the graphs shown in Figures 11 and 12.

The **DX-EGR1** tile was formed as described above. It was incorporated into the **ABCD** array in the same manner with the sole exception being that Tile **B** was replaced with **DX-EGR1**. The annealed array (67.5 nM) was incubated in either TAEMg or JBC+Mg buffer with 20 μM $ZnSO_4$ and 1 band forming units of EGR-1. The sample was then diluted five fold as the protein concentration was too high to obtain good images. 2.5 μL of the diluted sample was spotted on the mica and imaged.

Denaturing PAGE. All buffers and reagents used were purchased from Invitrogen and the manufacturer's protocols were followed. Briefly, the sample was prepared using 4 μL of EGR-1, 1 μL of NuPAGE Reducing Agent, 2.5 μL of 4X NuPAGE LDS Sample Buffer and 2.5 μL of water. Novex Sharp Protein Standard (prestained) 3.5 kDa – 260 kDa molecular weight standards from Invitrogen (VXLC5800) were used. The sample was heated at 70°C for 10 min prior to loading and then run on a Bis-Tris 4-12% Gel with MOPS running buffer. 500 μL of NuPAGE Antioxidant was added to the inner buffer chamber and the gel was run at 200 V for 25 min. Coomassie staining was done using a staining solution of .1% Coomassie R-250 in 40% ethanol and 10% acetic acid. The gel was placed in 100 mL of staining solution and microwaved for 1 min and then placed on a shaker for 15 min. The gel was then rinsed with water and placed in 100 mL of destain solution containing 10% ethanol and 7.5% acetic acid. It was again placed in the microwave for 1 min and then on a shaker for 15 min. The destain procedure was repeated a second time using fresh destain solution to obtain contrast.

Additional DNA Oligomers. The majority of the DNA oligomers used are described in Figures 2.1, 2.5, 2.9 and 2.14. The sequences used for DX-Junction which did not show any binding in the affinity cleavage assays, and for the duplex used in the EMSA experiments with polyamide **1** and streptavidin are shown below:

DX-Junction-1: 5' - TCACTCTACCGCACGAGAATGGAGAT - 3'

DX-Junction-2: 5' – CATTCTCGACGCCAATAGTTTGCACCTAACTTCATGTGCCT GCGGTAG - 3'

DX-Junction-3: 5'-CTGTAGCCTGCAAACTATTGGCGTGGCACATGAAGTTAGGA CAGATCG - 3'

DX-Junction-4: 5' - CATACCGATCTGTGGCTACAGTCTTG - 3'

DUPLEX1: 5' - CATTCTCGACGCTAGGTCACAGCAGGCTACTGTCTTG - 3'

DUPLEX2: 5' - CAGTAGCCTGCTGTGACCTAGCGTCGAGAATGGAGAT - 3'

Acknowledgement. We thank Prof. Erik Winfree and Dr. Sung Ha Park for useful discussions and assistance with the AFM experiments. We are grateful for support from the Ralph M Parsons Foundation and W. Webster.

2.5 References

1. Lin, C. X.; Liu, Y.; Rinker, S.; Yan, H., DNA tile based self-assembly: Building complex nanoarchitectures *ChemPhysChem* **2006,** 7, 1641-1647.

2. Seeman, N. C.; Lukeman, P. S., Nucleic acid nanostructures: Bottom-up control of geometry on the nanoscale *Rep. Prog. Phys.* **2005,** 68, 237-270.

3. Aldaye, F. A.; Palmer, A. L.; Sleiman, H. F., Assembling materials with DNA as the guide *Science* **2008,** 321, 1795-1799.

4. Seeman, N. C., DNA in a material world *Nature* **2003,** 421, 427-431.

5. Fu, T. J.; Seeman, N. C., DNA double-crossover molecules *Biochemistry* **1993,** 32, 3211-3220.

6. Winfree, E.; Liu, F. R.; Wenzler, L. A.; Seeman, N. C., Design and self-assembly of two-dimensional DNA crystals *Nature* **1998,** 394, 539-544.

7. Constantinou, P. E.; Wang, T.; Kopatsch, J.; Israel, L. B.; Zhang, X. P.; Ding, B. Q.; Sherman, W. B.; Wang, X.; Zheng, J. P.; Sha, R. J.; Seeman, N. C., Double cohesion in structural DNA nanotechnology *Org. Biomol. Chem.* **2006,** 4, 3414-3419.

8. Park, S. H.; Pistol, C.; Ahn, S. J.; Reif, J. H.; Lebeck, A. R.; Dwyer, C.; LaBean, T. H., Finite-size, fully addressable DNA tile lattices formed by hierarchical

assembly procedures (vol 45, pg 735, 2006) *Angew. Chem., Int. Ed.* **2006,** 45, 6607-6607.

9. Lund, K.; Liu, Y.; Yan, H., Combinatorial self-assembly of DNA nanostructures *Org. Biomol. Chem.* **2006,** 4, 3402-3403.

10. Rothemund, P. W. K., Folding DNA to create nanoscale shapes and patterns *Nature* **2006,** 440, 297-302.

11. He, Y.; Chen, Y.; Liu, H. P.; Ribbe, A. E.; Mao, C. D., Self-assembly of hexagonal DNA two-dimensional (2D) arrays *J. Am. Chem. Soc.* **2005,** 127, 12202-12203.

12. He, Y.; Tian, Y.; Ribbe, A. E.; Mao, C. D., Highly connected two-dimensional crystals of DNA six-point-stars *J. Am. Chem. Soc.* **2006,** 128, 15978-15979.

13. Aldaye, F. A.; Sleiman, H. F., Modular access to structurally switchable 3D discrete DNA assemblies *J. Am. Chem. Soc.* **2007,** 129, 13376-13377.

14. He, Y.; Ye, T.; Su, M.; Zhang, C.; Ribbe, A. E.; Jiang, W.; Mao, C. D., Hierarchical self-assembly of DNA into symmetric supramolecular polyhedra *Nature* **2008,** 452, 198-201.

15. Le, J. D.; Pinto, Y.; Seeman, N. C.; Musier-Forsyth, K.; Taton, T. A.; Kiehl, R. A., DNA-templated self-assembly of metallic nanocomponent arrays on a surface *Nano Lett.* **2004,** 4, 2343-2347.

16. Li, H. Y.; LaBean, T. H.; Kenan, D. J., Single-chain antibodies against DNA aptamers for use as adapter molecules on DNA tile arrays in nanoscale materials organization *Org. Biomol. Chem.* **2006,** 4, 3420-3426.

17. Liu, Y.; Lin, C. X.; Li, H. Y.; Yan, H., Protein nanoarrays - Aptamer-directed self-assembly of protein arrays on a DNA nanostructure *Angew. Chem., Int. Ed.* **2005,** 44, 4333-4338.

18. Park, S. H.; Yin, P.; Liu, Y.; Reif, J. H.; LaBean, T. H.; Yan, H., Programmable DNA self-assemblies for nanoscale organization of ligands and proteins *Nano Lett.* **2005,** 5, 729-733.

19. Chhabra, R.; Sharma, J.; Liu, Y.; Yan, H., Addressable molecular tweezers for DNA-templated coupling reactions *Nano Lett.* **2006,** 6, 978-983.

20. Zhang, J. P.; Liu, Y.; Ke, Y. G.; Yan, H., Periodic square-like gold nanoparticle arrays templated by self-assembled 2D DNA nanogrids on a surface *Nano Lett.* **2006,** 6, 248-251.

21. Williams, B. A. R.; Lund, K.; Liu, Y.; Yan, H.; Chaput, J. C., Self-assembled peptide nanoarrays: An approach to studying protein-protein interactions *Angew. Chem., Int. Ed.* **2007,** 46, 3051-3054.

22. Liu, F. R.; Sha, R. J.; Seeman, N. C., Modifying the surface features of two-dimensional DNA crystals *J. Am. Chem. Soc.* **1999,** 121, 917-922.

23. Yan, H.; Park, S. H.; Finkelstein, G.; Reif, J. H.; LaBean, T. H., DNA-templated self-assembly of protein arrays and highly conductive nanowires *Science* **2003,** 301, 1882-1884.

24. Sharma, J.; Chhabra, R.; Liu, Y.; Ke, Y. G.; Yan, H., DNA-templated self-assembly of two-dimensional and periodical gold nanoparticle arrays *Angew. Chem., Int. Ed.* **2006,** 45, 730-735.

25. Dervan, P. B., Molecular recognition of DNA by small molecules *Bioorg. Med. Chem.* **2001,** 9, 2215-2235.

26. Dervan, P. B.; Edelson, B. S., Recognition of the DNA minor groove by pyrrole-imidazole polyamides *Curr. Opin. Struct. Biol.* **2003,** 13, 284-299.

27. Hsu, C. F.; Phillips, J. W.; Trauger, J. W.; Farkas, M. E.; Belitsky, J. M.; Heckel, A.; Olenyuk, B. Z.; Puckett, J. W.; Wang, C. C. C.; Dervan, P. B., Completion of a programmable DNA-binding small molecule library *Tetrahedron* **2007,** 63, 6146-6151.

28. White, S.; Szewczyk, J. W.; Turner, J. M.; Baird, E. E.; Dervan, P. B., Recognition of the four Watson-Crick base pairs in the DNA minor groove by synthetic ligands *Nature* **1998,** 391, 468-471.

29. Dervan, P. B., Design of sequence-specific DNA-binding molecules *Science* **1986,** 232, 464-471.

30. Taylor, J. S.; Schultz, P. G.; Dervan, P. B., DNA affinity cleaving - Sequence specific cleavage of DNA by distamycin-EDTA.Fe(II) and EDTA-distamycin.Fe(II) *Tetrahedron* **1984,** 40, 457-465.

31. Olenyuk, B. Z.; Zhang, G. J.; Klco, J. M.; Nickols, N. G.; Kaelin, W. G.; Dervan, P. B., Inhibition of vascular endothelial growth factor with a sequence-specific hypoxia response element antagonist *Proc. Natl. Acad. Sci. U. S. A.* **2004,** 101, 16768-16773.

32. Tessmer, I.; Yang, Y.; Sass, L.; Hsieh, P.; Erie, D. A., AFM and single-molecule fluorescence studies of MutS-DNA interactions reveal the mechanism of mismatch recognition *Environ. Mol. Mutagen.* **2006,** 47, 416-416.

33. Wang, H.; Yang, Y.; Schofield, M. J.; Du, C. W.; Fridman, Y.; Lee, S. D.; Larson, E. D.; Drummond, J. T.; Alani, E.; Hsieh, P.; Erie, D. A., DNA bending and unbending by MutS govern mismatch recognition and specificity *Proc. Natl. Acad. Sci. U. S. A.* **2003,** 100, 14822-14827.

34. Yang, Y.; Sass, L. E.; Du, C. W.; Hsieh, P.; Erie, D. A., Determination of protein-DNA binding constants and specificities from statistical analyses of single molecules: MutS-DNA interactions *Nucleic Acids Res.* **2005,** 33, 4322-4334.

35. Nettikadan, S.; Tokumasu, F.; Takeyasu, K., Quantitative analysis of the transcription factor AP2 binding to DNA by atomic force microscopy *Biochem. Biophys. Res. Commun.* **1996,** 226, 645-649.

36. Lin, C. X.; Katilius, E.; Liu, Y.; Zhang, J. P.; Yan, H., Self-assembled signaling aptamer DNA arrays for protein detection *Angew. Chem.-Int. Edit.* **2006,** 45, 5296-5301.

37. Malo, J.; Mitchell, J. C.; Venien-Bryan, C.; Harris, J. R.; Wille, H.; Sherratt, D. J.; Turberfield, A. J., Engineering a 2D protein-DNA crystal *Angew. Chem., Int. Ed.* **2005,** 44, 3057-3061.

38. Desjarlais, J. R.; Berg, J. M., Use of a zinc-finger consensus sequence framework and specificity rules to design specific DNA-binding proteins *Proc. Natl. Acad. Sci. U. S. A.* **1993,** 90, 2256-2260.

39. Mandell, J. G.; Barbas, C. F., Zinc finger tools: custom DNA-binding domains for transcription factors and nucleases *Nucleic Acids Res.* **2006,** 34, W516-W523.

40. Pavletich, N. P.; Pabo, C. O., Zinc finger DNA recognition - Crystal-structure of a Zif268-DNA complex at 2.1-A *Science* **1991,** 252, 809-817.

41. Shi, Y. G.; Berg, J. M., A direct comparison of the properties of natural and designed zinc-finger proteins *Chem. Biol.* **1995,** 2, 83-89.

42. Wright, D. A.; Thibodeau-Beganny, S.; Sander, J. D.; Winfrey, R. J.; Hirsh, A. S.; Eichtinger, M.; Fu, F.; Porteus, M. H.; Dobbs, D.; Voytas, D. F.; Joung, J. K., Standardized reagents and protocols for engineering zinc finger nucleases by modular assembly *Nat. Protoc.* **2006,** 1, 1637-1652.

43. Nguyen-Hackley, D. H.; Ramm, E.; Taylor, C. M.; Joung, J. K.; Dervan, P. B.; Pabo, C. O., Allosteric inhibition of zinc-finger binding in the major groove of DNA by minor-groove binding ligands *Biochemistry* **2004,** 43, 3880-3890.

44. Skerka, C.; Decker, E. L.; Zipfel, P. F., A regulatory element in the human interleukin-2 gene promoter is a binding-site for the zinc-finger proteins Sp1 and Egr-1 *J. Biol. Chem.* **1995,** 270, 22500-22506.

45. Ke, Y. G.; Lindsay, S.; Chang, Y.; Liu, Y.; Yan, H., Self-assembled water-soluble nucleic acid probe tiles for label-free RNA hybridization assays *Science* **2008,** 319, 180-183.

46. Baird, E. E.; Dervan, P. B., Solid phase synthesis of polyamides containing imidazole and pyrrole amino acids *J. Am. Chem. Soc.* **1996,** 118, 6141-6146.

47. Trauger, J. W.; Dervan, P. B., Footprinting methods for analysis of pyrrole-imidazole polyamide/DNA complexes *Methods Enzymol.* **2001**, 340, 450-466.

48. Sambrook, J.; Fritsh, E. F.; Maniatis, T., *Molecular Cloning: Standard Protocols for DNA Manipulation. A laboratory Manual 2nd ed.* Cold Spring Harbor Laboratory: Plainview, NY, 1989.