14

Construction of Ordered Protein Arrays

Jarrod Clark, Taras Shevchuk, Piotr M. Swiderski, Rajesh Dabur, Laura E. Crocitto, Yaroslav I. Buryanov, and Steven S. Smith

Summary

Artificially ordered protein arrays provide a facile approach to a variety of problems in biology and nanoscience. Current demonstration systems use either nucleic acid tethers or methyltransferase fusions in order to target proteins or peptides of interest to nucleic acid scaffolds. These demonstrations point to the large number of useful devices and assemblies that can be envisioned using this approach, including smart biological probes and drug delivery systems. In principle, these systems are now capable of imitating the earliest forms of prebiotic organisms and can be expected to reach the complexity of a small virus in the near future. Third-generation methyltransferase inhibitors provide an example of a smart chemotherapeutics that can be constructed with this approach. We describe the use of mechanistic enzymology, computer-aided design, and microfluidic chip–based capillary electrophoresis in assessing the final assembly and testing of designs of this type.

Key Words: Bionanotechnology; DNA methyltransferase; protein ordering; protein targeting; DNA scaffold; inhibitor design; inhibitor delivery; smart chemotherapeutics; macromolecular carcerand; microfluidics.

1. Introduction

Bionanotechnology is an emerging field placed squarely at the interface between biotechnology and nanotechnology. Like other nanotechnologies, its goal is the production of useful materials and devices. Much of the technology in this area is biomimetic, meaning that it seeks to mimic biological processes. Most of the approaches taken to date are a form of macromolecular natural product chemistry aimed at producing macromolecular machines from nucleic acids and proteins. Biosensors ***(1)***, two-dimensional crystals ***(2)***, electronic circuitry ***(3)***, and a fueled oscillator ***(4)*** have all been described.

From: *Methods in Molecular Biology, vol. 300:*
Protein Nanotechnology, Protocols, Instrumentation, and Applications
Edited by: T. Vo-Dinh © Humana Press Inc., Totowa, NJ

The technology is now capable of imitating prebiotic molecular evolution. According to current thinking, the earliest form of organization in prebiology may have generated supramolecular aggregates called metabolosomes ***(5)***, which ultimately gave rise to translation ***(6)***. Prior to the appearance of translation, we think the RNA world was an evolving system in which RNA played the only structural and catalytic roles. Metabolosomes are viewed as assemblies in which nucleic acid base-pairing rules order different functions along a nucleic acid scaffold. Base pairing between a short region of the nucleic acid scaffold and a short region of the functionalized adapter RNAs is thought to have generated the appropriate alignment of the functionalities ***(5)***. Niemeyer et al. ***(7,8)*** have successfully utilized this approach in developing a bionanotechnology of ordered supramolecular assemblies.

In their initial experiments, they ordered the thermostable protein streptavidin along the screw axis of a DNA helix ***(8)*** by using biotinylated DNA as the adapter. The streptavidin protein binds to the end of the DNA owing to the innate capacity of streptavidin to bind extremely tightly to biotin. A DNA or an RNA molecule was then used as a scaffold to sequester and align the tethered streptavidin molecules ***(8)*** much as originally envisioned ***(5)*** for the metabolosome. When DNA is used as the organizing molecule, the streptavidin moieties are expected to be arrayed around the screw axis of the DNA molecule (36°/bp); however, when RNA is used, this orientation is expected to be in accordance with the screw axis of the RNA·DNA hybrid (30°/bp).

Streptavidin aggregates also provide the opportunity for secondary assembly. For example, by adding biotin to both of the 5' ends of complementary oligodeoxynucleotides, it has been possible to form ring structures or networks of DNA linked by two, three, or four biotin contacts per streptavidin residue ***(7)***. Moreover, the streptavidin molecule binds biotin with a stoichiometry of four biotin residues per protein moiety. This leaves three unused sites on the DNA-streptavidin conjugate. These unused sites have been used to attach biotinylated antibodies to DNA-tethered streptavidin ***(8)***. This permits the derivatization of microstructured surface arrays of DNA using the base-pairing complementarity of the array so as to convert it into a protein array of different antigen specificities ***(8)***.

Biotinylated proteins can be linked to avidin and then targeted to DNA- or RNA-organizing molecules provided that they can withstand the extended exposure to temperatures above 50°C required for the annealing process. Alternatively, biotinylated proteins can be sequestered without regard to order along an organizing DNA or RNA molecule by incubating them at low temperature with streptavidin molecules tethered to DNA or RNA. Using these methods, the effects of protein proximity have also been studied with this system. In an example of this approach, the NADPH oxidoreductase was placed

adjacent to luciferase in order to improve the properties of light emission by the system that is normally coupled by these enzymes. A measurable improvement was observed ***(9)***. In summary, the tethered adapter technology studied by Niemeyer et al. has generated several interesting and potentially useful nanoscale assemblies containing ordered protein arrays.

We have developed an alternative approach to the construction of ordered protein arrays. This approach is analogous to a later stage in prebiotic evolution after nucleic acid complementarity had allowed metabolosomes to evolve translation. The advent of translation permitted prebiotic systems to produce peptides and proteins that provided additional flexibility to supramolecular aggregates. The flexibility provided by sequence–specific interaction between nucleic acids and proteins is thought to have allowed these supramolecular aggregates to move past the RNA world into the modern molecular biology that forms the current basis of the living state. The bionanotechnology that we have developed for the construction of ordered protein arrays is based on protein–nucleic acid interaction ***(10,11)***. The technology's current state suggests that it will permit the construction of programmable devices with a degree of sophistication equal to that of the most advanced prebiotic metabolosomes and perhaps that of small viruses.

2. Ordered Assembly of a Three-Address Array

The fundamental principle of ordered assembly of a three-address array is the capacity of DNA methyltransferases to form covalent linkages with DNA in the reaction depicted in **Fig. 1**. This property, coupled with their selectivity for defined nucleic acid sequences, allows DNA methyltransferases to serve as targeting agents for fusion proteins directed to specific sites on a DNA scaffold ***(11)***. In what follows, the decoration of a Y-Junction recombination intermediate serves as an example of the methodology employed in ordering proteins using this technology.

2.1. Molecular Models

Although not always possible or absolutely necessary, we have found it useful in most cases to prepare at least preliminary molecular models of the assemblies under consideration. In this case (**Fig. 2**), molecular models of the Y-Junction recombination intermediate were constructed in Insight II (Accelrys, San Diego, CA). The DNA builder was used to extend the DNA present in the three-dimensional structure of the M·*Hha*I protein–DNA complex ***(12)*** obtained from the Protein Data Bank ***(13)*** with PDB ID:1MTH. The arms were linked at the center to form a model of the Y-Junction having an accurate sequence. Assisted model building with energy refinement molecular force field ***(14)*** was used to minimize the structure while constraining the DNA and protein struc-

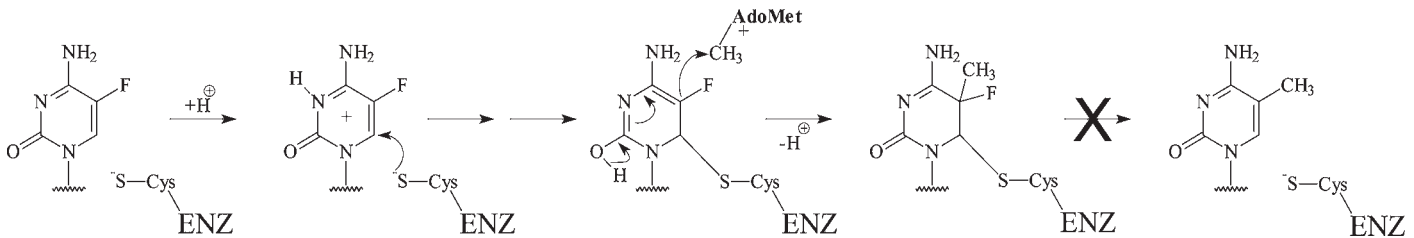

Fig. 1. DNA methyltransferase mechanism of action. Methyltransferases make a nucleophilic attack on C6 of cytosine or 5-fluorocytosine. This breaks the 5-6 double bond in the ring and activates C5 for methyltransfer. After the methyl group is transferred from *S*-adenosyl-[L]-methionine (AdoMet) to C5 of cytosine, the hydrogen at C5 and the enzyme nucleophile at C6 are removed by β-elimination. When fluorine is present at C5, this cannot occur because of the strength of the fluorine-carbon bond. Thus, the progress of the reaction is blocked, leaving a covalent linkage between the enzyme and the cytosine ring targeted by the enzyme. In the case of M•*Eco*RII, it is the second cytosine in the CCWGG recognition sequence that is attacked, in which W stands for A or T. M•*Hha*I attacks the first cytosine in the GCGC sequence, and human Dnmt1 (hDnmt1) attacks the cytosine in the CG dinucleotide.

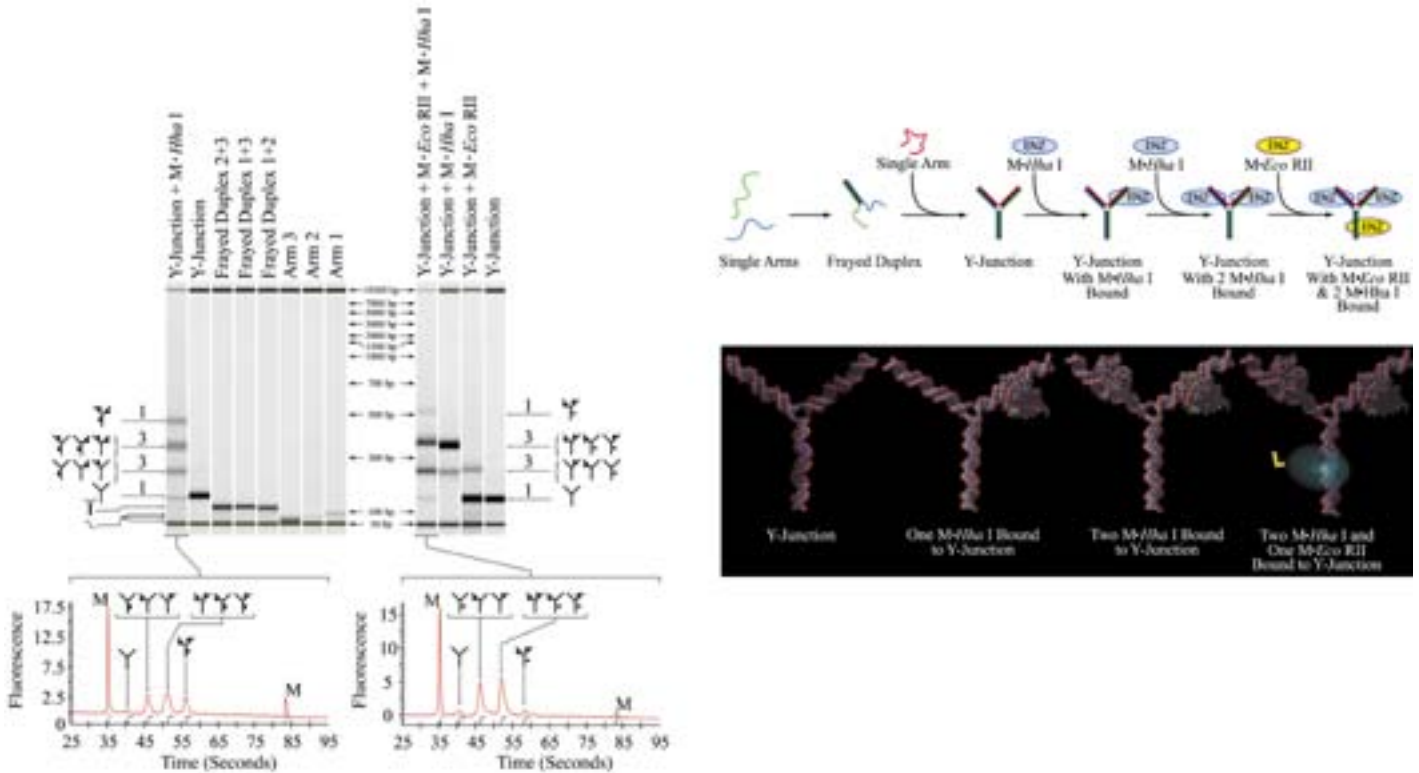

Fig. 2. Assembly of a three-address nucleoprotein array. A schematic of the final assembly process for the machine is depicted in the upper right. The black panel in the lower right depicts molecular models of the predicted intermediates in the assembly of the three-address array once the duplex Y-Junction itself is formed. In the last frame of the molecular model panel, NLS-M•*Eco*RII is depicted as an ovoid with a crooked appendage depicting the NLS peptide. *(Figure 2 caption continued on next page.)*

tures in the arms beyond the 10 bp at the center of the junction. The energies of the resulting structures were minimized with respect to molecular mechanics until the RMS force was <0.1 (kcal/mol)/Å. Following molecular mechanics, each structure was further minimized with 1000 steps of simple dynamics. We did not attempt to prepare a molecular model of the nuclear localization signal (NLS)-M•*Eco*RII fusion protein because a crystal structure for M•*Eco*RII was not available. Current information on this protein ***(15)*** suggested that a 30-bp arm on the Y-Junction could accommodate the fusion protein in the environment created by the linking of M•*Hha*I proteins to the adjacent arms of the Y-Junction.

2.2. Preparation of Oligodeoxynucleotide

5-Fluorodeoxycytidine was introduced as the TMP-F-dU-CE convertible phosphoramidite (Glen Research, Sterling, VA). All other bases were introduced with standard phosphoramidite chemistry. Concentrations were measured by absorbance spectroscopy at 260 nm. Duplexes and Y-Junctions were formed by mixing equimolar amounts of oligodeoxynucleotides in 10 m*M* Tris-HCl (pH 7.2), 1 m*M* EDTA, and 100 m*M* NaCl, so as to give a final concentration of 6 µ*M*. Annealing was carried out at 95°C for 5 min, 50°C for 60 min, room temperature for 10 min, and then put on ice for 10 min. The Y-Junction oligodeoxynucleotide sequences employed in this study were as follows:

Fig. 2. *(continued from previous page)* Mobility shift detection of nucleoprotein complexes was carried out using a newly developed microfluidics chip method ***(20)***. This method employs the Bioanalyzer system (Agilent, Palo Alto, CA) and LabChip® microfluidics chips (Caliper Technologies, Mountain View, CA) normally employed in nucleic acid detection and characterization ***(56,57)***. DNA 7500, chips designed for the separation of DNA duplexes over the range of 15 to 7500 bp, were used in the analysis. Each of the three single strands migrate at their expected lengths. Frayed duplexes formed by each of the combinations of two strands migrate with an apparent mobility corresponding to about 110-bp duplex markers. The Y-Junction forms readily when all three strands are annealed (rightmost lane in the DNA 500 chip virtual gel panel). It has a mobility corresponding to that of a 150-bp duplex in the 7500 chip. The Y-Junction bound to the fusion protein exhibits a mobility corresponding to that of a 270-bp duplex. When M•*Hha*I was introduced into the reaction, the expected set of intermediates was observed when three addresses for M•*Hha*I were present (leftmost lane in the DNA chip virtual gel panel) and inset mobility scan depiction, lower left). Alternatively, when one address for NLS-M•*Eco*RII and two addresses for M•*Hha*I were present on the junction, a different set of intermediate forms was observed (right lane of the DNA chip virtual gel panel and the inset mobility scan depiction, lower right).

M F
1. 5'-GCTGGCTATGCCACCAGGTGAGGCTAACTGAGGTAGCACGACCGCGCTACAATCTGGTGA-3'

M F
2. 5'-TCACCAGATTGTAGCGCGGTCGTGCTACCTGGTTCCACCAGATGCGCGTGACCTGTAGTT-3'

M F
3. 5'-AACTACAGGTCACGCGCATCTGGTGGAACCCAGTTAGCCTCACCTGGTGGCATAGCCAGC-3'

in which M indicates 5-methyl and F indicates a 5-fluoro moiety on cytosine.

2.3. Monitoring Y-Junction Assembly With Microfluidics Chips

In the past, we have used polyacrylamide gel electrophoresis (PAGE) to monitor assembly of DNA scaffolds ***(11,16)***; however, this methodology is time-consuming and, in general, requires the use of ^{32}P end labeling. More recently, we have used microfluidics chip systems for this analysis because it is very rapid, requires negligible amounts of sample, and does not require the use of radiolabeled DNA.

The DNA 500 LabChip or the DNA 7500 LabChip can be used for the detection of retardation products. Nine microliters of gel dye mix was added to the appropriate well, and the chip was primed with 0.8 mL of air for 30 s. When the DNA 500 LabChip was employed, 5 µL of the gel dye mix was added to each well, and 1 µL of the 500 chip ladder was added to the ladder well. When the DNA 7500 LabChip was employed, 5 µL of the internal marker mix and 1 µL of the 7500 chip ladder were added to the ladder well. Nine microliters of gel dye mix was then added to each of the two gel reservoir wells. One microliter from each reaction was added to 1 of the 12 sample wells. The chip was then vortexed at 2400 rpm for 1 min. Finally, the chip was placed in an Agilent 2100 Bioanalyzer and processed within 5 min.

A schematic of the steps in the assembly of the Y-Junction is given in **Fig. 2** (first three steps in the upper right panel). Representative results of an analysis of this type are also depicted in **Fig. 2** (leftmost virtual gel, lanes 1–7 right to left).

2.4. Cloning and Expression of the NLS-M•EcoRII Fusion

The construction and expression of fusion proteins using recombinant DNA techniques is an essential component of this technology. Each fusion is different and requires an approach tailored to the nature of the targeting methyltransferase and the targeted fusion protein. Expression of the fusion must be followed by at least partial purification. The cloning, expression, and purification of the NLS-M•*Eco*RII fusion is given as a simple example.

Plasmid pRT103 containing the M•*Eco*RII-NLS fusion protein was constructed by inserting an NLS cassette at the *Nco*I site in the multipurpose clon-

ing site of pRT103. The full-length M•*Eco*RII sequence was isolated from a pBR322 clone and ligated into the *Bam*HI site of the multipurpose cloning site in the pRT103 ***(17)***. This resulted in an in-frame NLS sequence at the N-terminus of the M•*Eco*RII, under the control of the plant 35-S CaMV promoter, which is effective in promoting transcription in *Escherichia coli*. The mammalian nuclear localization signal is located at the 5' end of the transcript. The bacterial translation stop signal from CaMV is located 3' to the *Bam*H1 site. The plasmid was transfected into *E. coli* GM2163 (New England Biolabs) for continuous expression under the control of the 35-S promoter. Bacteria were grown aerobically in Luria broth until they reached stationary state.

Although this particular fusion was formed using restriction cleavage and ligation, we have also found it both convenient and effective to use polymerase chain reaction (PCR) to amplify fusion proteins and the targeting methyltransferases. High-fidelity polymerases are effective in this application, and the use of PCR permits the introduction of any desired restriction site in the bounding primers, thereby facilitating assembly of plasmid systems for fusion production.

2.4.1. Purification

Purification of the fusion protein was essentially as described previously ***(18,19)***. The cells were lysed in a buffer containing 50 m*M* Tris-HCl (pH 7.8), 10% (v/v) glycerol, 1 m*M* EDTA, and 750 µg/mL of lysozyme (Sigma, St. Louis, MO). After 30 min at 0°C, the mixture was incubated for 5 min at 20°C. The viscous mixture was sonicated in 15-s bursts, with 1 min of cooling at 0°C between bursts at the microtip maximum setting with a Branson Sonicator (Branson, Danbury, CT). The temperature during sonication was maintained below 15°C. Cellular debris was removed by centrifuging at 5000*g* for 10 min. The supernatant fluid was then mixed with an equal volume of DEAE Sepharose (Pharmacia, Uppsala, Sweden) that had been equilibrated with a buffer containing 50 m*M* Tris-HCl (pH 7.8), 10% (v/v) glycerol, 1 m*M* EDTA, and 150 m*M* NaCl. After stirring for 30 min, the unbound material was collected by filtration.

The filtrate was then mixed with 0.5 vol of precycled phosphocellulose (P-11; Whatman, Maidstone, UK). After stirring for 30 min, the mixture was filtered and the P-11 cake was resuspended in 2 vol of the same buffer containing 150 m*M* NaCl. The cake was again collected by filtration and resuspended in 2 vol of the same buffer containing 400 m*M* NaCl. Elution was carried out by filtration. The filtered fluid (approx 170 mg of protein) was collected and dialyzed to 150 m*M* NaCl in the elution buffer. The dialyzed material was applied to two tandemly linked 5-mL Cibacron-blue Econo-Pak fast protein liquid chromatography (FPLC) columns (Bio-Rad, Hercules, CA). The column was

washed with 50 mL of the same buffer containing 180 m*M* salt. Elution was carried out with a step increase in salt concentration to 1.5 *M* NaCl in the same buffer. Gradients were prepared and protein detection was carried out with a Waters Millennium 32 System (Waters, Milford MA). Active fractions were then concentrated using Centricon-10 concentrators (Amicon, Beverly, MA). Unlike the M·*Hha*I preparation that was essentially homogeneous by PAGE, the M•*Eco*RII-NLS fusion preparation contained approx 15 visible bands on polyacrylamide gels stained with Coomassie blue R250 (Sigma), suggesting that it made up about 20% of the visible protein. M·*Hha*I was purified as previously described ***(11)***.

2.4.2. Protein Coupling Conditions

Annealed Y-Junction at concentrations ranging from 0.37 to 1.5 μ*M* was exposed to concentrations of M·*Hha*I ranging from 0.81 to 2.43 μ*M* in a binding buffer containing 50 m*M* Tris-HCl (pH 7.8), 10 m*M* EDTA, 5 m*M* β-mercaptoethanol, and 80 μ*M* AdoMet. The final volume of the reaction was 10 μL. The reaction was then incubated for 2.5 h at 37°C.

2.5. Monitoring Final Assembly With Microfluidics Chip-Based Protein Mobility Shift

Electrophoretic mobility shift analysis (EMSA) is a well-characterized and widely employed technique for the analysis of protein–DNA interaction and of transcription factor combinatorics. As currently implemented, EMSA generally involves the use of radiolabeled DNA and PAGE. We have noted ***(20)*** that this technique could be effectively implemented with microfluidics chips designed for the separation of DNA fragments. To accomplish this, samples were run on a 2100 Bioanalyzer (Agilent) using a DNA 500 LabChip or a DNA 7500 LabChip (Caliper Technologies) according to the manufacturer's instructions. **Figure 2** (right virtual gel panel) depicts the results of monitoring the final assembly obtained by mixing the duplex Y-Junction with a mixture of the NLS-M•*Eco*RII fusion protein and M•*Hha*I. When M•*Hha*I addresses were present on each arm of the Y-Junction, the expected intermediates were observed (Y-Junction + M•*Hha*I l; *see* left virtual gel and inset scanned panel in **Fig. 2**).

3. Applications of Ordered Arrays in Smart Drug Design

The design progression for DNA methyltransferase inhibitors is depicted in **Fig. 3**. Here we see small molecules that must be incorporated into DNA as the first generation of such inhibitors. Single-strand conformers (SSCs) that display the methyltransferase recognition motif and target the enzyme to an inhibitory target are the second-generation inhibitors. Third- and fourth-generation inhibitors can be approached with the bionanotechnology of ordered protein arrays as described in **Subheadings 3.4.** and **3.5.**

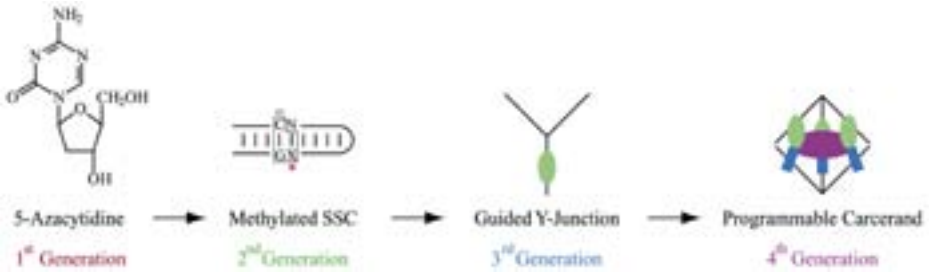

Fig. 3. Schematic of design progression. First-generation inhibitors are small molecules with the capacity to become incorporated in DNA and trap methyltransferases. Second-generation inhibitors are single-strand conformers (SSCs) that carry a methyltransferase trap. In this figure, the three-nucleotide recognition motif characteristic of hDnmt1 is marked with an L-shaped box. The third-generation inhibitors are recombination intermediates with binding sites for methylation activators linked to methyltransferase traps. They are guided to the nucleus of the cell by linked proteins with nuclear localization signals. Fourth-generation inhibitors are self-assembling and programmable hemi carcerands that are capable of sensing their environment and disgorging or exposing an encapsulated or protected entity at a desired cellular or intracellular location.

3.1. Tumor Biology of DNA Methyltransferase Inhibition

Although we are only beginning to understand fully the complexities of *selective* gene activation and repression in normal cells, these processes appear to be mediated by targeted recruitment of chromatin remodeling factors coupled with developmentally regulated protein–protein and protein–DNA interaction combinatorics resulting in promoter activation or repression. Once established, tissue-specific patterns of gene silencing appear to be stably maintained by a self-reinforcing network of protein and DNA modifications involving histone methylation, histone hypoacetylation, and cytosine methylation ***(21)***. These modifications do not appear to *control* gene silencing in the sense that they do not initiate gene silencing during normal differentiation. However, they are clearly important hallmarks *of* silenced genes in mammals that are established by cofactors and changes in nucleic acid structure that alter the activity and specificity of methyltransferases and deacetylases ***(22–26)***.

What is interesting is the general randomization of gene expression patterns during tumorigenesis resulting from the complex interplay of irreversible genetic damage and natural selection during tumor progression. In addition to the attendant alterations in genome structure associated with this process, alterations in patterns of gene expression are clonally established that result in ectopic gene expression and ectopic gene silencing.

Given that clones of tumor cells that are capable of continued expansion have achieved a degree of adaptation to the local environment during tumorigenesis, it is reasonable to suggest that wholesale disruption of the clonally established patterns of gene expression that they have achieved would at least slow their proliferation. This concept has led to a search for methods by which stable patterns of heterochromatinization can be disrupted. As expected, histone deacetylase inhibitors and DNA (cytosine-5) methyltransferase inhibitors act synergistically ***(27,28)*** to destabilize the self-reinforcing modification network in heterochromatin permitting the nonspecific reactivation of silent genes. However, the inhibition of either histone deacetylase or DNA (cytosine-5) methyltransferase is generally sufficient for reactivation. Of these, the DNA methyltransferase inhibitors have received the most attention as potential chemotherapeutic agents (*see* **refs. *29–31*** for reviews).

3.2. First-Generation Inhibitors

The current armamentarium of inhibitors for which at least preliminary human trials have been conducted consists of a series of soluble small molecules that act by becoming incorporated into DNA ***(29–32)***.

As with most general attacks on the machinery of the cell, key parameters in the efficacy of the approach have been the intrinsic sensitivity of the target tumor, the biodistribution of the inhibitory drugs, and the avoidance of potential side effects ***(29–32)***. Among the potential side effects unique to this approach is the real potential for the release of undesirable activities such as multidrug resistance from gene silencing ***(33)***.

Although the patient samples are currently very small, significant responses to the DNA methyltransferase inhibitor decitabine (2'deoxy-5-azacytidine) have been seen with acute leukemias. Complete responses in as many as 33% of patients have been reported for relapsed chronic myelogenous leukemia, with 20 to 30% for relapsed acute myelogenous leukemia ***(34)***. Responses with solid tumors have been less frequent, with melanoma the highest at 5% for the DNA methyltransferase inhibitor 5,6-dihydro-5-azacytidine ***(35)***.

With prostate cancer, although methylation abnormalities have been widely reported in cell lines and clinical specimens, clinical studies with methyltransferase inhibitors are very limited. Nevertheless, decitabine induces cellular differentiation in PC3 and DU145 cell lines ***(36)***. Moreover, in a limited patient study, it has been reported that decitabine is well tolerated and gave stable responses in 2 of 12 cases of metastatic D2 prostate cancer ***(36)***.

3.3. Second-Generation Inhibitors

It is clearly too early to tell whether or not the aforementioned first-generation methyltransferase inhibitors will have chemotherapeutic value alone or in

combination chemotherapy ***(29,31)***. Even so, two new directions in the inhibition of methyltransferases are being explored. Both first-generation inhibitors of DNA methyltransferase ***(32)*** and second-generation inhibitors of methyltransferase action are under study ***(37–41)***. The use of second-generation inhibitors should avoid the side effects produced by the necessity of random incorporation into DNA required for decitabine action because second-generation inhibitors are DNA analogs that act as direct inhibitors of the methyltransferase ***(38,40,42,43)***.

Although the second-generation inhibitors do not use ordered protein arrays or bionanotechnology *per se*, they are potential components of the third-generation inhibitors, as described in **Subheading 3.4.** Thus, it is valuable to review their design properties.

3.3.1. Electronic Structure Activity Relations-Identified Targets and DNA Methyltransferase Inhibitor Design

Electronic structure activity relations (ESAR) have successfully predicted the effects of the mechanism-based inhibitors FdC ***(11)***, dU ***(44)***, and 2-pyrimidinone ***(45)*** on bacterial enzyme. Thus, it appears to be of general utility in determining the effects that a candidate inhibitor will have on methyltransferase action. An algorithm for analyzing candidate targets is depicted in **Fig. 4**. The results of calculations using the algorithm for a series of known and likely inhibitors of methyltransferases are depicted graphically in **Fig. 5**. In light of the known parameters of the human DNA (cytosine-5)methyltransferase, generally termed Dnmt1, the data indicate that the candidate second-generation inhibitors of the enzyme should have the following characteristics:

1. They should generally present a duplex region (that may contain mispairs) comprising about 28 bp of duplex in a hairpin formed of a single strand ***(46,47)*** or duplex formed from two fully or partially complementary oligodeoxynucleotides ***(48)***.
2. The partial duplex should contain nonproductive ***(44)*** or weakly productive three-nucleotide motifs ***(38,48)***. As defined in the cited references, a three-nucleotide recognition motif consists of a cytosine or preferably a 5-methylcytosine on one strand that is paired with a guanine or an inosine residue that is 3' to the targeted base. In short, this motif defines the targeted base. It is depicted in **Fig. 3** (second-generation panel).
3. These nonproductive motifs should target the enzyme to a nonproductive or base-modified nucleotide in which the reference absolute value of the frontier orbital energy difference (*see* **Fig. 4**) for nucleophilic attack is >0.73797 eV, or >7.39287 eV for methyltransfer based on the data given in ***(43)***.

The hairpin molecules described in **refs.** ***38***, ***41***, ***46***, and ***49*** satisfy these criteria with frontier orbital energy differences for methyl transfer exceeding the requirement in point no. 3 for a normally paired nucleotide analog.

3.3.2. Multiply Mistargeted Bases

Saturation curves for a productive substrate ($^{dC}S_1$) and a multiply mistargeted nonproductive substrate ($^{NdC}S_2$) are depicted in **Fig. 5**. The sequences are as follows:

1. $^{dC}S_1$: CG-methylated oligo targeting normally paired dC

```
                                        m
Oligo CG-methyl   5' GTC-CAC-CAG-ATC-CGG-GCT-ACC-TGG-CCT-CGA 3'
Oligo no-methyl   3' CAG-GTG-GTC-TAG-GCC-CGA-TGG-ACC-GGA-GCT 5'
                                       *
```

2. $^{NdC}S_2$: non-CG-methylated oligo targeting normally paired dC, dT, dA, and dG

```
                          m m m m     m m    m   mm     mm  m
Oligo non-CG-methyl   5' GTC-CAC-CAG-ATC-CGG-GCT-ACC-TGG-CCT-CGA 3'
Oligo no-methyl       3' CAG-GTG-GTC-TAG-GCC-CGA-TGG-ACC-GGA-GCT 5'
                             XX  XX      X*    X   X X    XX  *
```

in which x indicates a mistargeted base, * indicates an appropriately targeted base, and m indicates a targeting methyl group.

Since previous work with substrate $^{dC}S_1$ ***(50,51)*** has shown that tritium incorporation is confined almost exclusively to the methyl-targeted cytosine (*), tritium incorporation into substrate $^{NdC}S_2$ is likely to be directed to the two productive sites in the substrate or other cytosine residues in the oligodeoxynucleotide. Based on the ESAR, 10 of the 12 targeted sites will produce nonproductive binding of the enzyme to sites targeting dT, dG, or dA where LUMO energies exceed that required for nucleophilic attack. The human enzyme has a footprint in the range of 28 bp ***(48)***, suggesting that it would generally bind only once to an oligodeoxynucleotide of this type with a probability of 10/12 (0.833) of binding to a nonproductive site. Thus, the low *Vm* and low *Km* observed for this substrate strongly suggest that the bulk of the interaction with the substrate is at nonproductively targeted dA, dG, and dT sites. Uniform methylation also raises the *Tm* of the oligodeoxynucleotide from 74 to 80°C (data not shown), suggesting that the enzyme may also be able to bind at methylated targeting sites but be slowed by the additional energy required to disrupt the base-stacking interactions as the bases are flipped out of the helix.

Given that the kinetics monitor tritium incorporation into either substrate, the kinetics are described by the reaction scheme given in **Fig. 5A**, left. The general equations describing this reaction ***(52)*** are as follows:

$$v = (Vm_1[S_1]/Km_1 + Vm_2[S_2]/Km_2)/(1 + [S_1]/Km_1 + [S_2]/Km_2) \quad (1)$$

$$Km_2 = Km_1[S_2](v - Vm_2)/(Vm_1[S_1] - v([S_1] + Km_1)) \quad (2)$$

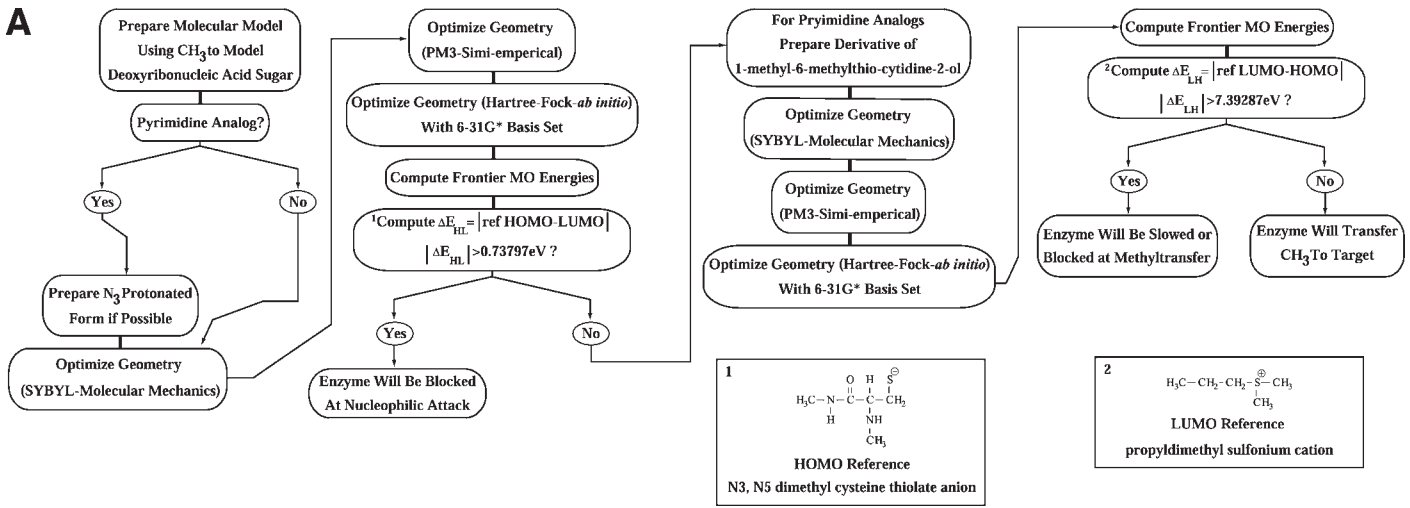

Fig. 4. ESAR for second-generation methyltransferase inhibitors. **(A)** ESAR algorithm for deciding whether nucleophilic attack or methyltransfer will be affected in inhibition. **(B)** Frontier orbital picture comparing models of targets that permit nucleophilic attack at a cytosine model compound (acceptor attacked) with models of targets that do not permit nucleophilic attack (acceptor not attacked): (I) 1-methyl-5-pyrimidinone cation (N3 protonated); (II) 1-methyl-5-fluorocytosine cation (N3 protonated); (III) 1-methyl-5-azacytosine cation (N3 protonated); (IV) 1-methylcytosine cation (N3 protonated); (V) 1-methyl-4-thiouracil; (VI) 1-methyl-3-bromouracil; (VII) 1-methyl-2-pyrimidinone; (VIII) 1-methyl-5-fluorouracil; (IX) 1-methyluracil; (X) 1-methylpseudouracil (ψ); (XI) 1-methylthymine; (XII) 1-methyl-5-azacytosine; (XIII) 1-methylcytosine; (XIV) 9-methyl-8-oxoguanine; (XV) 9-methyladenine; (XVI) 9-methylguanine; (XVII) 9-methyl-7-deazaguanine; (XXII) N3, N5 dimethyl cysteine thiolate anion.

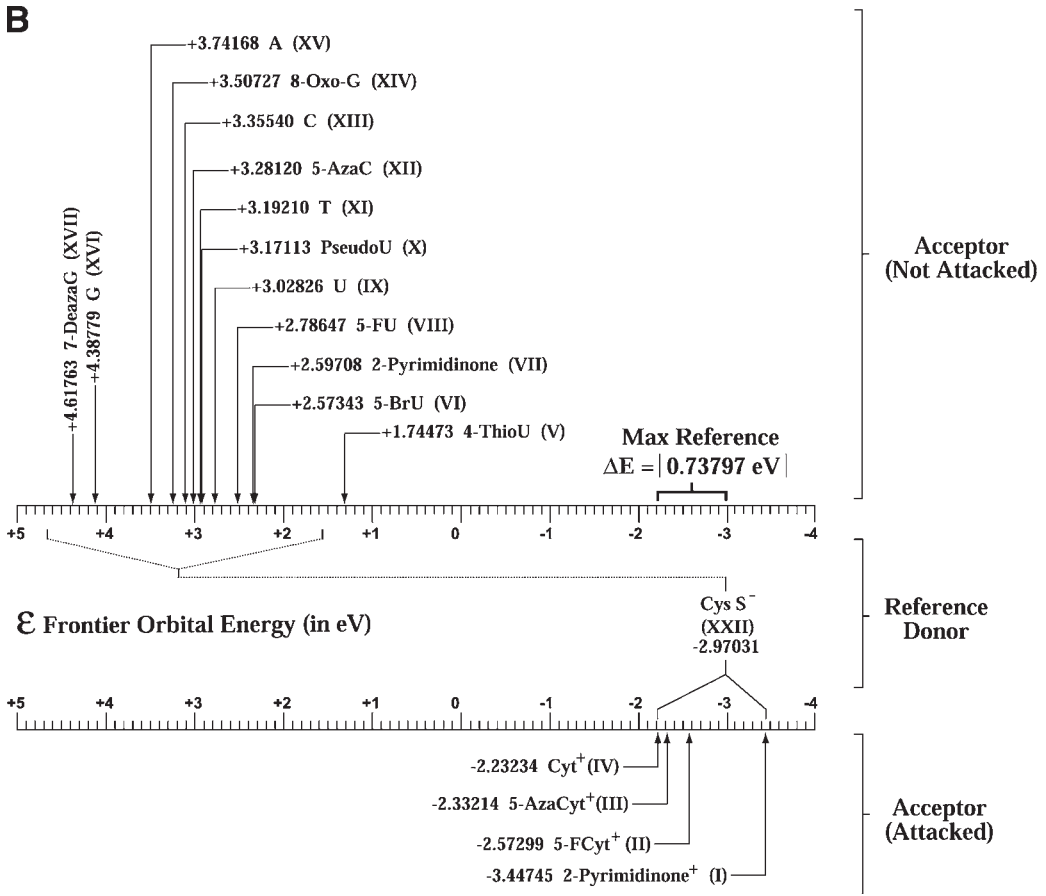

Fig. 4. *(continued)*

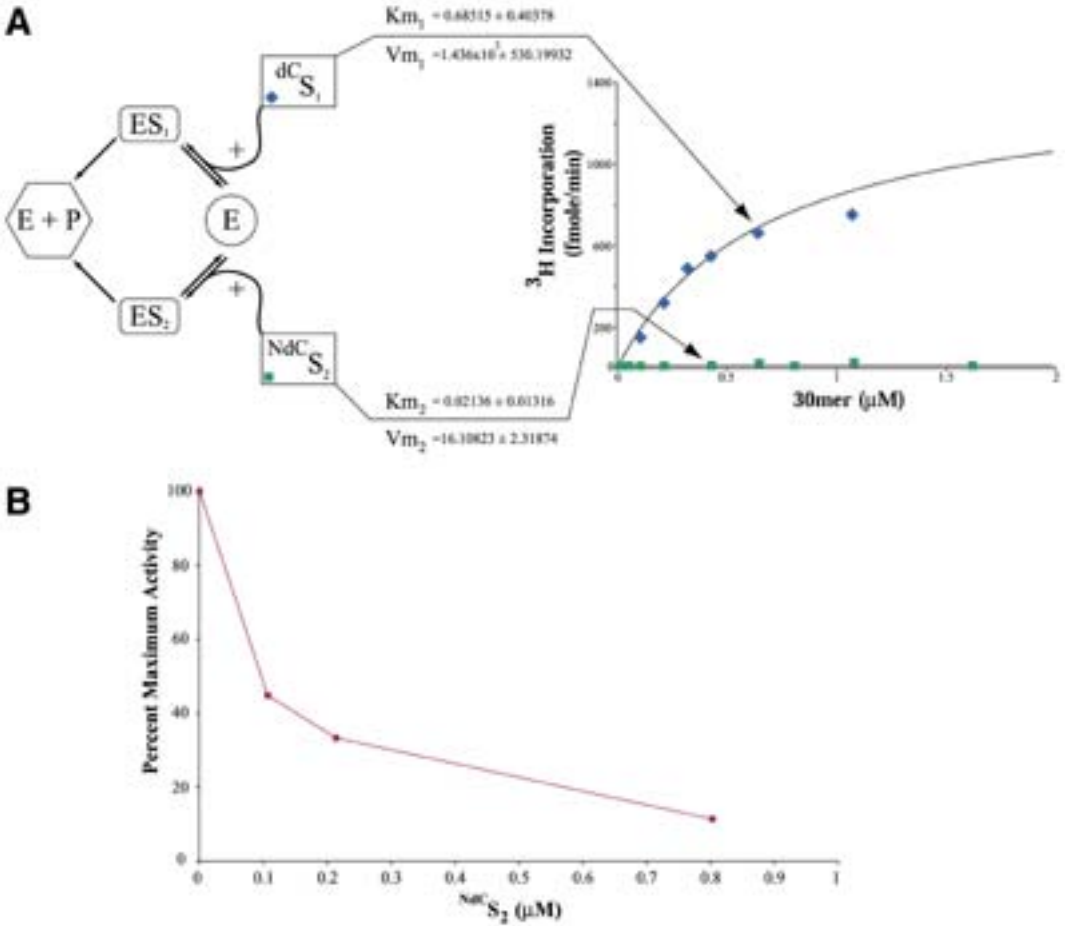
A
Km1 = 0.68515 ± 0.40378
Vm1 = 1.436x10^3 ± 536.19932
dCS1
ES1
+
E + P
E
ES2
+
NdCS2
Km2 = 0.02136 ± 0.01316
Vm2 = 16.10823 ± 2.31874
3H Incorporation (fmole/min)
30mer (μM)
B
Percent Maximum Activity
NdCS2 (μM)

Fig. 5.

3.3.3. Singly Mistargeted Mispaired Bases

Although the methyltransferase reaction involves both AdoMet and DNA as substrates, at constant AdoMet concentrations, the DNA saturation kinetics are hyperbolic for either the highly productive or weakly productive substrate. When a constant concentration of the productive substrate is present and varying amounts of the weakly productive substrate are added, inhibition by the weakly productive substrate follows the kinetics predicted for a two-DNA substrate system. In many cases, incorporation into weakly productive competing substrates is very low. This makes it difficult to estimate Km_2 from direct saturation experiments. In these cases, we have found it useful to arrive at an upper limit for Vm_2 and determine an average value for Km_2 from **Eq. 2**.

In general, we have noted that multiply mistargeted second-generation inhibitors are more potent than singly mistargeted inhibitors based on side-by-side comparisons. An inhibition curve for the multiply mistargeted inhibitor $^{NdC}S_2$ is provided in **Fig. 5B**.

3.4. Third-Generation Inhibitors

One form of third-generation inhibitors under study comprises recombination intermediates with binding sites for methylation activators that are linked to methyltransferase traps like those found in second-generation inhibitors. These inhibitors are to be guided to the nucleus of the cell by linked proteins with nuclear localization signals, as shown schematically in **Fig. 6**. The design parameters of such an inhibitor are given in **Fig. 7**. To construct this machine, the NLS-M•*Eco*RII fusion was cloned and the fusion protein was purified as described in **Subheading 2.4.** Final assembly was monitored by mobility shift analysis using a Bioanalyzer 2100 and DNA analysis microfluidics chips. As can be seen in **Fig. 2**, each step in the final assembly process could be monitored by the appearance of the expected intermediate in the microfluidics analysis. The data clearly show that methyltransferase-directed nucleoprotein

---

Fig. 5. *(previous page)* Nucleic acid–based inhibitor action. **(A)** Saturation plots for hemimethylated DNA and non-CG methylated DNA. Although each substrate carries methyl groups at cytosines on only one strand, the hemimethylated substrate ($^{dC}S_1$) targets hDnmt1 to a single cytosine residue. The second substrate ($^{NdC}S_2$) mistargets the enzyme to dG, dA, and dT residues as well. Based on the two-substrate reaction scheme on the right, increasing the concentration of ($^{NdC}S_2$) at a constant concentration of ($^{dC}S_1$) is expected to inhibit the enzyme. **(B)** Inhibition of human DNA methyltransferase by singly and multiply mistargeted substrate analogs. The active substrate ($^{dC}S_1$) gives rapid incorporation of ^{3}H-methyl groups from AdoMet into DNA (normalized to 100% activity). This ongoing reaction was inhibited by increasing the amounts of a second substrate, $^{NdC}S_2$.

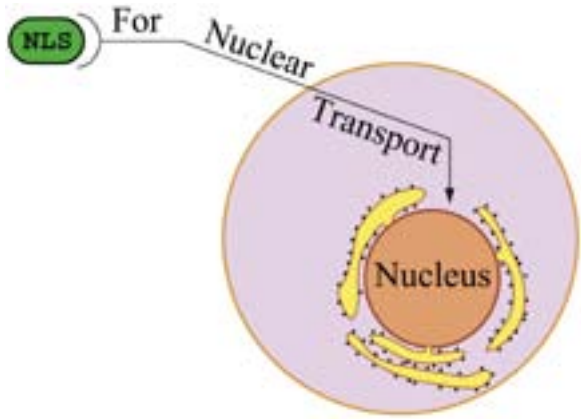

Fig. 6. Schematic of effect of NLS in directing protein transport to the nucleus.

assembly can be used to produce a machine for inhibition of human DNA methyltransferase based on a bionanotechnological design (*see* **Fig. 2**, Y-Junction + M•*Eco*RII lane). The design is now being tested for its efficacy in the inhibition of hDnmt1 in vitro and in vivo.

More sophisticated designs based on arrayed protein signals or binding sites are possible. Designs of this type, however, would require four-arm or X-Junctions in order to accommodate the arrayed functions. A detailed discussion of these designs is beyond the scope of this chapter.

3.5. Fourth-Generation Inhibitors

Fourth-generation inhibitors are analogous to the small-molecule carcerands first studied by Warmuth and Yoon ***(53)***. These self-assembling and programmable macromolecular hemi carcerands ***(54)*** would be capable of sensing their environment and disgorging or uncoating an encapsulated or hidden entity at a desired cellular or intracellular location. A representative design is given in **Fig. 8**. The value of using an ordered protein array in a design of this type is the availability of DNA-binding proteins that bind to DNA or release only in the presence of a small-molecule ligand. Many proteins of this type are known, and it will be possible to engineer responses to desired ligands that will cause the hemi carcerand to disassemble in the presence of high local concentrations of the ligand. Disassembly could be used to release a macromolecule or expose the internal surface of the structure so as to attack or modify certain cellular systems localized in the region of high ligand concentration.

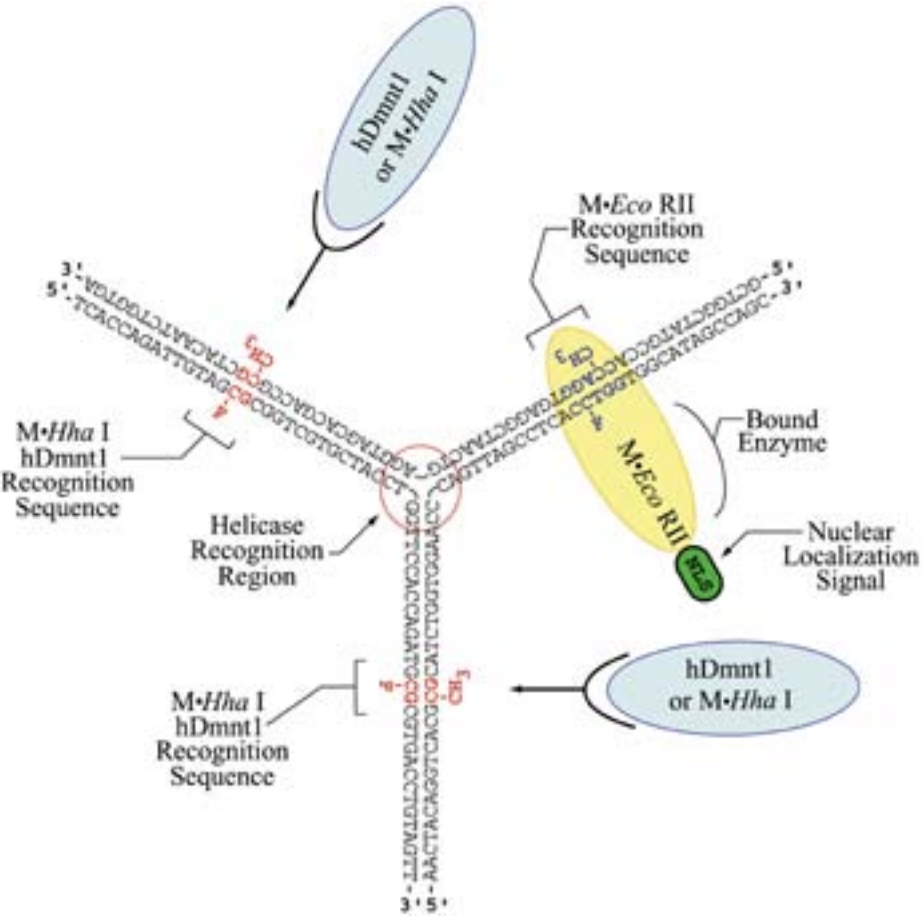

Fig. 7. Bionanotechnological machine for inhibition of human DNA methyltransferase: a guided Y-Junction. This machine comprises a Y-Junction protein array scaffold that can act as a recombination intermediate so as to activate further the already activated methyltransferase targeting sites (hDnmt1 recognition sequence). The device also has the potential to attract helicases (helicase recognition region). The arms of the Y-Junction carrying the human recognition sequences are each 30 bp long. Thus, they each present an adequate kinetic footprint for the hDnmt1 at each of the dual trapping regions of the inhibitor. The entire system is stabilized by the presence of a trapped bacterial methyltransferase (M•*Eco*RII) carrying an NLS that is designed to link the machine to the cellular nuclear transport system so as to place it effectively in the cell nucleus where the bulk of hDnmt1 acts.

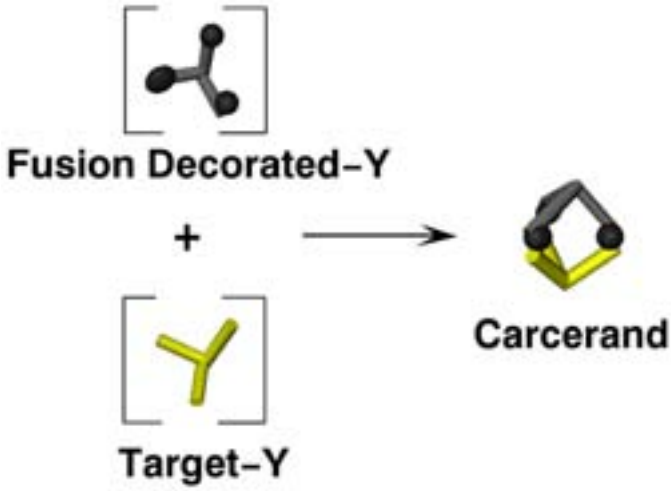

Fig. 8. Self-assembling trigonal dipyramid as a macromolecular carcerand. A methyltransferase-targeted DNA-binding protein is used to decorate a Y-Junction so as to place the fusion proteins on the same face of the helix in each arm of the Y-Junction. The targeted Y-Junction contains binding sites for the DNA-binding protein on each of its arms. Mixing the two components should result in a closed structure with the edge topology of a trigonal dipyramid (Johnson Solid). The resulting lumen could be used to house a macromolecule, or to hide an inhibitory nucleic acid surface. Disassembly of the structure (not shown) would occur if a ligand-sensitive DNA-binding protein were used as the fusion protein of choice to produce a hemi carcerand. Depending on the choice of fusion, the hemi carcerand would disassemble in a region of low or high ligand concentration because DNA binding or release by the fusion would be induced by changes in ligand concentration.

4. Conclusion

Ordered protein arrays provide a facile approach to a variety of problems in biology and nanoscience. Current demonstration systems use either nucleic acid tethers ***(55)*** or methyltransferase fusions ***(54)*** in protein targeting to nucleic acid scaffolds. These systems are now capable of imitating the earliest forms of prebiotic organisms and can be expected to reach the complexity of a small virus in the near future. The use of microfluidics chips in assessing final assembly of these designs ***(20)*** should facilitate their construction and testing. The number of assemblies that can be contemplated is quite large, ranging from tethered enzyme couples ***(9)*** to the smart chemotherapeutic agents described here. Thus, one can anticipate that a large number of interesting and complex array-based assemblies and devices will appear in the near future.

References

1 Yun, C. S., Khitrov, G. A., Vergona, D. E., Reich, N. O., and Strouse, G. F. (2002) Enzymatic manipulation of DNA-nanomaterial constructs. *J. Am. Chem. Soc.* **124,** 7644, 7645.

2 Seeman, N. C. (1991) Construction of three-dimensional stick figures from branched DNA. *DNA Cell. Biol.* **10,** 475–486.

3 Gil, A., de Pablo, P. J., Colchero, J., Gómez-Herrero, J., and Baró, A. M. (2002) Electrostatic scanning force microscopy images of long molecules: single-walled carbon nanotubes and DNA. *Nanotechnology* **13,** 309–313.

4 Yurke, B., Turberfield, A. J., Mills, A. P. Jr., Simmel, F. C., and Neumann, J. L. (2000) A DNA-fuelled molecular machine made of DNA. *Nature* **406,** 605–608.

5 Gibson, T. J. and Lamond, A. I. (1990) Metabolic complexity in the RNA world and implications for the origin of protein synthesis. *J. Mol. Evol.* **30,** 7–15.

6 Woese, C. R. (2002) On the evolution of cells. *Proc. Natl. Acad. Sci. USA* **99,** 8742–8747.

7 Niemeyer, C. M., Adler, M., Gao, S., and Chi, L. (2000) Supramolecular nanocircles consisting of streptavidin and DNA. *Angew. Chem. Int. Ed. Engl.* **39,** 3055–3059.

8 Niemeyer, C. M., Sano, T., Smith, C. L., and Cantor, C. R. (1994) Oligonucleotide-directed self-assembly of proteins: semisynthetic DNA—streptavidin hybrid molecules as connectors for the generation of macroscopic arrays and the construction of supramolecular bioconjugates. *Nucleic Acids Res.* **22,** 5530–5539.

9 Niemeyer, C. M., Koehler, J., and Wuerdemann, C. (2002) DNA-directed assembly of bienzymic complexes from in vivo biotinylated NAD(P)H:FMN oxidoreductase and luciferase. *Chembiochem.* **3,** 242–245.

10 Smith, S. S. (1995) Nucleoprotein-based nanoscale fabrication. *Biol. Biomed. Sci. Technol. Div.* Office of Naval Research (ONR) 34196-3, 161, 162.

11. Smith, S. S., Niu, L., Baker, D. J., Wendel, J. A., Kane, S. E., and Joy, D. S. (1997) Nucleoprotein-based nanoscale assembly. *Proc. Natl. Acad. Sci. USA* **94,** 2162–2167.

12 Klimasauskas, S., Kumar, S., Roberts, R. J., and Cheng, X. (1994) HhaI methyltransferase flips its target base out of the DNA helix. *Cell* **76,** 357–369.

13 Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000) The protein data bank. *Nucleic Acids Res.* **28,** 235–242.

14 Ponder, J. W. and Case, D. A. (2003) Force fields for protein simulations. *Adv. Protein Chem.* **66,** 27–85.

15 Schroeder, S. G. and Samudzi, C. T. (1997) Structural studies of EcoRII methylase: exploring similarities among methylases. *Protein Eng.* **10,** 1385–1393.

16 Smith, S. S. (2001) A self-assembling nanoscale camshaft: implications for nanoscale materials and devices constructed from proteins and nucleic acids. *Nano Lett.* **1,** 51–56.

17 Topfer, R., Matzeit, V., Gronenborn, B., Schell, J., and Steinbiss, H. H. (1987) A set of plant expression vectors for transcriptional and translational fusions. *Nucleic Acids Res.* **15,** 5890.

18 Kosykh, V. G., Solonin, A. S., Buryanov Y. I., and Bayev, A. A. (1981) Overproduction of the EcoRII endonuclease and methylase by Escherichia coli strains carrying recombinant plasmids constructed in vitro. *Biochim. Biophys. Acta* **655,** 102–106.
19 Buryanov, Y. I., Bogdarina, I. G., and Bayev, A. A. (1978) Site specificity and chromatographic properties of E. coli K12 and EcoRII DNA-cytosine methylases. *FEBS Lett.* **88,** 251–254.
20 Clark, J., Shevchuk, T., Swiderski, P. M., Dabur, R., Crocitto, L. E., Buryanov, Y. I., and Smith, S. S. (2003) Mobility-shift analysis with microfluidics chips. *Biotechniques* **35,** 548–554.
21 Richards, E. J. and Elgin, S. C. (2002). Epigenetic codes for heterochromatin formation and silencing: rounding up the usual suspects. *Cell* **108,** 489–500.
22 Smith, S. S. and Crocitto, L. (1999) DNA methylation in eukaryotic chromosome stability revisited: DNA methyltransferase in the management of DNA conformation space. *Mol. Carcinog.* **26,** 1–9.
23 Smith, S. S. (2000) Gilbert's conjecture: the search for DNA (cytosine-5) demethylases and the emergence of new functions for eukaryotic DNA (cytosine-5) methyltransferases. *J. Mol. Biol.* **302,** 1–7.
24 Baylin, S. and Bestor, T. H. (2002) Altered methylation patterns in cancer cell genomes: cause or consequence? *Cancer Cell* **1,** 299–305.
25 Smith, S. S. (1994) Biological implications of the mechanism of action of human DNA (cytosine-5)methyltransferase. *Prog. Nucleic Acid Res. Mol. Biol.* **49,** 65–111.
26 Smith, S. S., Laayoun, A., Lingeman, R. G., Baker, D. J., and Riley, J. (1994) Hypermethylation of telomere-like foldbacks at codon 12 of the human c-Ha-ras gene and the trinucleotide repeat of the FMR-1 gene of fragile X. *J. Mol. Biol.* **243,** 143–151.
27 Chiurazzi, P., Pomponi, M. G., Pietrobono, R., Bakker, C. E., Neri, G., and Oostra, B. A. (1999) Synergistic effect of histone hyperacetylation and DNA demethylation in the reactivation of the FMR1 gene. *Hum. Mol. Genet.* **8,** 2317–2323.
28 Cameron, E. E., Bachman, K. E., Myohanen, S., Herman, J. G., and Baylin, S. B. (1999) Synergy of demethylation and histone deacetylase inhibition in the re-expression of genes silenced in cancer. *Nat. Genet.* **21,** 103–107.
29 Goffin, J. and Eisenhauer, E. (2002) DNA methyltransferase inhibitors: state of the art. *Ann. Oncol.* **13,** 1699–1716.
30 Brown, R. and Strathdee, G. (2002) Epigenetic cancer therapies: DNA methyltransferase inhibitors. *Expert Opin. Investig. Drugs* **11,** 747–754.
31 Christman, J. K. (2002) 5-Azacytidine and 5-aza-2'-deoxycytidine as inhibitors of DNA methylation: mechanistic studies and their implications for cancer therapy. *Oncogene* **21,** 5483–5495.
32 Stewart, D. J., Donehower, R. C., Eisenhauer, E. A., Wainman, N., Shah, A. K., Bonfils, C., MacLeod, A. R., Besterman, J. M., and Reid, G. K. (2003) A phase I pharmacokinetic and pharmacodynamic study of the DNA methyltransferase 1 inhibitor MG98 administered twice weekly. *Ann. Oncol.* **14,** 766–774.

33. El-Osta, A. (2003) On the use of DNA methylation inhibitors and the reversal of transcriptional silencing. *Blood* **101,** 1656; author reply 1657, 1658.
34 Momparler, R. L., Rivard, G. E., and Gyger, M. (1985) Clinical trial on 5-aza-2'-deoxycytidine in patients with acute leukemia. *Pharmacol. Ther.* **30,** 277–286.
35 Creagan, E. T., Schaid, D. J., Hartmann, L. C., and Loprinzi, C. L. (1993) A phase II study of 5,6-dihydro-5-azacytidine hydrochloride in disseminated malignant melanoma. *Am. J. Clin. Oncol.* **16,** 243, 244.
36 Thibault, A., Figg, W. D., Bergan, R. C., Lush, R. M., Myers, C. E., Tompkins, A., Reed, E., and Samid, D. (1998) A phase II study of 5-aza-2'deoxycytidine (decitabine) in hormone independent metastatic (D2) prostate cancer. *Tumori* **84,** 87–89.
37 Smith, S. S., Lingeman, R. G., and Kaplan, B. E. (1992) Recognition of foldback DNA by the human DNA (cytosine-5-)-methyltransferase. *Biochemistry* **31,** 850–854.
38 Smith, S. S., Kaplan, B. E., Sowers, L. C., and Newman, E. M. (1992) Mechanism of human methyl-directed DNA methyltransferase and the fidelity of cytosine methylation. *Proc. Natl. Acad. Sci. USA* **89,** 4744–4748.
39 Knox, J. D., Araujo, F. D., Bigey, P., Slack, A. D., Price, G. B., Zannis-Hadjopoulos, M., and Szyf, M. (2000) Inhibition of DNA methyltransferase inhibits DNA replication. *J. Biol. Chem.* **275,** 17,986–17,990.
40 Bigey, P., Knox, J. D., Croteau, S., Bhattacharya, S. K., Theberge, J., and Szyf, M. (1999) Modified oligonucleotides as bona fide antagonists of proteins interacting with DNA: hairpin antagonists of the human DNA methyltransferase. *J. Biol. Chem.* **274,** 4594–4606.
41. Smith, S. S. and Kaplan, B. E. (1996) Mechanism-based inhibitors of DNA methyltransferase. US patent 5,503,975.
42 Kho, M. R., Baker, D. J., Laayoun, A., and Smith, S. S. (1998) Stalling of human DNA (cytosine-5) methyltransferase at single-strand conformers from a site of dynamic mutation. *J. Mol. Biol.* **275,** 67–79.
43 Clark, J., Shevchuk, T., Kho, M. R., and Smith, S. S. (2003) Methods for the design and analysis of oligodeoxynucleotide-based DNA (cytosine-5)methyltransferase inhibitors. *Anal. Biochem.* **321,** 50–64.
44 Wendel, J. A. and Smith, S. S. (1998) Uracil as an alternative to 5-fluorocytosine in addressable protein targeting. *Nanotechnology* **9,** 297–304.
45 Zhou, L., Cheng, X., Connolly, B. A., Dickman, M. J., Hurd, P. J., and Hornby, D. P. (2002) Zebularine: a novel DNA methylation inhibitor that forms a covalent complex with DNA methyltransferases. *J. Mol. Biol.* **321,** 591–599.
46 Smith, S. S., Lingeman, R. G., and Kaplan, B. E. (1992) Recognition of foldback DNA by the human DNA (cytosine-5)-methyltransferase. *Biochemistry* **31,** 850–854.
47 Christman, J. K., Sheikhnejad, G., Marasco, C. J., and Sufrin, J. R. (1995) 5-Methyl-2'-deoxycytidine in single-stranded DNA can act in cis to signal de novo DNA methylation. *Proc. Natl. Acad. Sci. USA* **92,** 7347–7351.
48 Laayoun, A. and Smith, S. S. (1995) Methylation of slipped duplexes, snapbacks and cruciforms by human DNA(cytosine-5)methyltransferase. *Nucleic Acids Res.* **23,** 1584–1589.

49 Szyf, M. (1998) Targeting DNA methyltransferase in cancer. *Cancer Metastasis Rev.* **17,** 219–231.
50 Smith, S. S., Hardy, T. A., and Baker, D. J. (1987) Human DNA (cytosine-5)methyltransferase selectively methylates duplex DNA containing mispairs. *Nucleic Acids Res.* **15,** 6899–6916.
51 Smith, S. S., Kan, J. L., Baker, D. J., Kaplan, B. E., and Dembek, P. (1991) Recognition of unusual DNA structures by human DNA (cytosine-5)methyltransferase. *J. Mol. Biol.* **217,** 39–51.
52. Chaplin, M. F. and Bucke, C. (1990) *Enzyme Technology*, University Press, Cambridge.
53 Warmuth, R. and Yoon, J. (2001) Recent highlights in hemicarcerand chemistry. *Acc. Chem. Res.* **34,** 95–105.
54 Smith, S. S. (2002) Designs for the self-assembly of open and closed macromolecular structures and a molecular switch using DNA methyltransferase to order proteins on nucleic acid scaffolds. *Nanotechnology* **13,** 413–419.
55. Niemeyer, C. M. (2000) Self-assembled nanostructures based on DNA: towards the development of nanobiotechnology. *Curr. Opin. Cell Biol.* **4,** 609–618.
56 Fuller, R. A., Clark, J., Kretzner, L., Korns, D., Blair, S. L., Crocitto, L. E., and Smith, S. S. (2003) Use of microfluidics chips for the detection of human telomerase RNA. *Anal. Biochem.* **313,** 331–334.
57. Nachamkin, I., Panaro, N. J., Li, M., Ung, H., Yuen, P. K., Kricka, L. J., and Wilding, P. (2001) Agilent 2100 bioanalyzer for restriction fragment length polymorphism analysis of the Campylobacter jejuni flagellin gene. *J. Clin. Microbiol.* **39,** 754–757.