DNA-Templated Assembly of Protein Complexes at Nanoscale

G. E. Sroga
Rensselaer Polytechnic Institute
Ricketts 101, 110 8th Street, Troy, NY 12180-3590, USA, srogag@rpi.edu

ABSTRACT

Directed, biologically-driven self-assembly has the potential to yield hybrid multicomponent architectures with applications ranging from sensors and diagnostics to nanoelectronic devices. Critical to these applications is to gain control over the precise orientation and geometry of biomolecules interacting with one-another and with surfaces. Such control has thus far been difficult to achieve in even the simplest biomolecular designs. Using a novel strategy for generation of multicomponent biological nanoarchitectures, the DNA-templated assembly of multiprotein complexes recognizing methylated DNA was achieved. The reassembly of two fragments of TEM-1 β-lactamase, each one fused with a specific DNA recognition factor, into a catalytically active protein was achieved by using the cognate DNA elements of these factors. This strategy could potentially become a useful tool in studies of genomic DNA methylation in the context of cellular epigenetic processes.

Keywords: fragmented enzyme reporter, reassembly driven by methylated dsDNA, nanoscale biomolecular architectures, catalytic function

1 INTRODUCTION

Methylation of genomic DNA is a common characteristic of living organisms such as bacteria, plants or animals. However, its role varies widely between different organisms. Even within the animal genomes, methylation patterns differ substantially from undetectable in nematodes to global methylation in vertebrate genomes. In vertebrates, DNA methylation occurs predominantly at position 5 of cytosines when followed by guanosine (CpG). CpG islands are GC-rich regions of DNA, stretching for an average of about 1 kb, and for example in humans, they coincide with the promoters of approximately 60% of genes transcribed by RNA polymerase II [1]. DNA methylation is known to play an essential role in gene silencing [2, 3] and mammalian development [4]. Thus, the promoters having CpG islands could be subjected to regulation by methylation. However in the case of genomic DNA methylation, there could be subtle differences between identical genomes that escape detection by strategies available nowadays, for example, by current microarray technologies. Therfore, there is a clear need for experimental strategies with capabilities to address emerging issues of epigenetics.

One mechanism by which DNA methylation can cause transcriptional repression is by direct interference with the binding of sequence-specific transcription factors to DNA. Indeed, some transcription factors have been shown to be unable to bind to their target sequences, which became methylated [5, 6]. More indirect mechanism of repression also exists and is supported by observations that DNA methylation can repress transcription at some distance [7, 8] as well as only after chromatin assembly [9]. Several proteins have been identified, which bind specifically to methylated DNA in any sequence context [10]. These proteins have similar structures and make specific contacts in the major grove of methylated DNA, and therefore, became the focus of this study aiming to develop a simple experimental approach to investigate changes in genomic DNA methylations during some developmental processes.

Specifically, two well characterized proteins were selected, human MBD2 protein that binds to methylated CpG islands with K_d = 2.7 nM [11] and Zinf268 that binds to its cognate dsDNA sequence motif with K_d = 6 nM [12]. As MBD2 binding affinity for nonmethylated-CpG sites is at least 70-fold lower then for methylated CpGs, it was inferred that the difference in MBD2 the binding affinities for methylated and nonmethylated DNA should allow selective targeting of the methylated CpG ($_m$CpG) sites. Conversely, it was expected that the exquisite specificity of Zif268 transcription factor for 5'-GCGGGTGC-3' sequence motif would facilitate control over targeting of the engineered protein complexes to the desired DNA sites. In order to be able to detect attachment of the respective DNA-binding proteins to DNA, a reporter was needed that would permit monitoring of the assembly process.

The enzyme-based assays are simple and relatively inexpensive. *E. coli* TEM-1 β-lactamase was selected as a reporter as this monomeric enzyme appears to meet all the essential criteria of a desired protein reporter. It can be easily expressed in *E. coli* and is not toxic to both prokaryotic and eukaryotic cells. As eukaryotic cells do not contain endogenous β-lactamase activity, potential problems caused by an unspecific background would not be encountered. Structure and function of *E. coli* TEM-1 β-lactamase are well characterized [13] and its activity can be simply assayed by the hydrolysis of nitrocefin using a colorimetric assay. It was also shown that the enzyme reassembles from its fragments into a functional protein [14] and this was particularly relevant to the discussed work as one of its main goals has been to explore the structural complementation of the fragmented β-lactamase in conjunction with the presence of investigated DNA.

NSTI-Nanotech 2008, www.nsti.org, ISBN 978-1-4200-8503-7 Vol. 1

To this end, a potentially general approach for the design and synthesis of structured and functional protein assemblies reporting the degree of double-strand DNA (dsDNA) methylation was developed. Exchanging MBD2 for other natural or synthetic proteins binding methylated dsDNA and Zif268 for other transcription factors, respectively, may yield more complete set of tools to study global changes of DNA methylation during different biological processes.

2 EXPERIMENTAL PROCEDURES

2.1 Cloning, Expression, and Purification of the Proteins

Unless otherwise stated, all molecular biology methods used in the presented study were performed according to the standard protocols described by Sambrook et al. [15] and Sambrook and Russell [16]. The *E. coli* gene of TEM-1 β-lactamase originated from pCRII-TOPO vector (Invitrogen). Full-length *E. coli* TEM-1 β-lactamase, two fragments of the enzyme, Bla1 and Bla2 [17], and the investigated fusion proteins were generated using the PCR method. The M182T mutation was introduced into the Bla1 sequence to enhance the stability of the protein [18]. The Bla1 fragment of TEM-1 β-lactamase (H26 to G196) was fused with Zif268 and the Bla2 fragment (L198 to W290) with MBD2, respectively. In each case, a 15 amino acid linker separated the Bla domain from the fused protein.

Periplasmic expression of the respective proteins was achieved after recloning of the gene constructs into the pBAD/gIII-D (Kana^R) vector (Invitrogen). Protein expression was conducted according to the manufacturer's instruction. Cell fractionation and extracts' analysis was performed as described by Sroga and Dordick [19].

2.2 Methylation of DNA *in vitro*

Biotinylated DNA fragments (B-DNA) contained Zif268 sequence motif (5'-GCGGGTGC-3') separated from the methyl-CpG ($_m$CG) site (i.e., the binding target for MBD2 domain) by 10 bp DNA spacer. Cognate nonmethylated DNA and random DNA sequence were used in the control experiments. All DNA fragments contained a 16 bp linker sequence at the 3'-end. The DNA fragments were generated using a method published previously [20] and attached to the streptavidin-coated paramagnetic particles (SA-PMPs) [20] after their methylation.

Metylation of dsDNA was conducted using *Sss*I methylase (New England Biolabs) according to the manufacturer's protocol.

2.3 Assembly of Engineered Proteins onto DNA Attached to SA-PMPs

E. coli periplasmic fractions served as the direct source of Bla1-Zif268 and Bla2-MBD2 fusion proteins. To perform the DNA-driven reassembly of TEM-1 β-lactamase, the appropriate periplasmic fractions were mixed in various combinations with PMPs carrying different DNAs. After washings, the PMPs with the attached complexes were directly used in the nitrocefin assay.

2.4 Nitrocefin Assay

The assay was performed using standard 96-well micro-titer plates (MT-plates). To the wells of MT-plate containing 20 or 40 μL of the appropriate DNA-PMPs in the nitrocefin assay buffer (100 mM phosphate, pH 7.0), 20μL of a given protein sample(s) were added to the final volume of 200 μL, mixed and placed immediately in the HTS 7000 Plus Bio Assay Reader (Perkin Elmer, Norwalk, CT) with the HTS 2.0 software. Absorbance readings were recorded at 486 nm at one min intervals for 20 min. The SA-PMPs in 200 μL of the nitrocefin assay buffer served as a blank. Initial reaction rate data were collected and calculated using the HTS 2.0 software taking the molar extinction coefficient for hydrolyzed nitrocefin ($\varepsilon_{486\ \text{hydrolyzed nitrocefin}} = 20.5 \times 10^3\ M^{-1}\ cm^{-1}$) into account.

3 RESULTS AND DISCUSSION

To meet the requirements of a simple assay for *in vitro* detection and quantitation of genomic DNA methylations, two engineered fragments of *E. coli* TEM β-lactamase were tested for their ability to reassemble onto a methylated-dsDNA template. The assembly occurred through the appropriate DNA recognition factor that was fused in frame with a given fragment of β-lactamase (Figure 1).

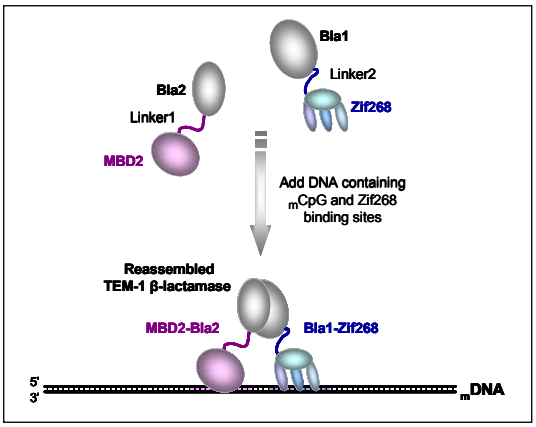

Figure 1: Schematic showing the principle of the methylated-dsDNA detection through the assembly of the split TEM-1 β-lactamase.

Each protein fragment of β-lactamase was designed with a 15 amino acid linker that separated it from the DNA-binding factor in order to prevent potential steric clashes. Conversely, selection of the 10 bp spacing between the binding sites was based on the typical distance observed

between natural DNA binding motifs (also named boxes) within promoters for various transcription factors as well as the results from the previous work [20].

Catalytic activity of TEM-1 β-lactamase was determined using the colorimetric substrate nitrocefin that changes color from yellow to red (peak absorbance at 486 nm) upon hydrolysis. The activity of the enzyme was restored only after addition of methylated-DNA-PMPs to the protein samples containing a mixture of the Zif268-Bla1 and MBD2-Bla2 fusion proteins (Figure 2). The observed initial reaction rate, $v_{0\ [Hydrolyzed\ nitrocefin]}$, for hydrolytic activity of reassembled TEM-1 β-lactamase was 0.56 ± 0.04 μM/40 μL periplasmic fraction min^{-1} (Table 1). The specificity of the binding was further confirmed by showing the lack of enzymatic activity in the samples containing only one type of the assembly component and methylated cognate DNA or both protein constructs and either nonmethylated-cognate or non-cognate DNA (Figure 2). In summary, this part of the investigation demonstrated that the Bla1 and Bla2 fragments can be brought together to form an active enzyme only in the presence of methylated cognate dsDNA.

% of DNA on SA-PMPs **A**: $_{m}$CpG-$(N)_{10}$-$TFBS_{Zif268}$ **B**: CpG-$(N)_{10}$-$TFBS_{Zif26}$	v_0 [Hydrolyzed nitrocefin] [μM/40 μL periplasmic fraction min^{-1}]
100% **A**	0.56 ± 0.04
50% **A** + 50% of **B**	0.29 ± 0.03
100% **B**	ND[1)]

[1)] ND = Not Detected. Values are the average of four parallel assays. Average amount of DNA attached to 75 μl particles was 600 ng (21 pmoles). To ensure similar binding-space availability, the length of all tested DNA fragments was ca. 75 bp. Thus, molar concentrations and masses of the DNAs used were comparable.

Table 1: Initial reaction rates ($v_{0\ [Hydrolyzed\ nitrocefin]}$) for hydrolytic activity of reassembled TEM-1 β-lactamase investigated using different percentage of methylated and nonmethylated cognate dsDNA fragments attached to SA-PMPs.

Catalytic activity of the reassembled TEM-1 β-lactamase can be observed as an increasing intensity of the red color product (Figure 2). To test that proteins of *E. coli* periplasm do not cause background hydrolysis of nitrocefin, periplasmic fraction of *E. coli* host was incubated with nitrocefin under the standard assay conditions. No observable hydrolysis was detected over the assay period as well as within 1 hr of its completion. The rationale for testing periplasmic fractions directly in the assay instead of the purified proteins was to develop more robust and faster to perform experimental procedure. Notably, similar results were obtained when the assay was performed using purified fusion proteins.

In a separate set of experiments, two- and three-fold higher concentration of engineered proteins was used for attachment and this resulted in approximately 1.8-fold and 2.8-fold increase in v_0, respectively. These data imply that in addition to the identification of DNA methylation, a concentration of the reassembled TEM-1 β-lactamase from its fragments can be simply calculated, as under the usual *in vitro* enzyme assay conditions reaction rate is directly proportional to the enzyme concentration, and hence, to the level of recombinant DNA-binding factors. Thus, it was calculated that on average periplasmic fractions contained about 18 – 19 % of a given fusion protein.

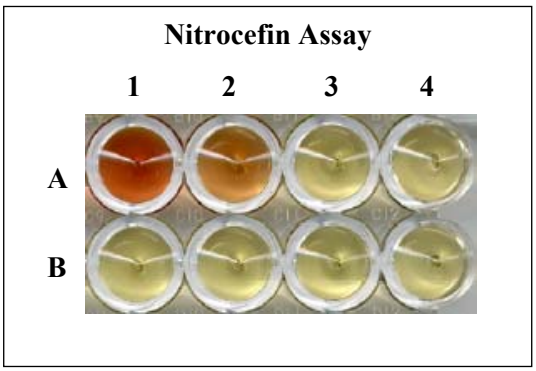

Figure 2: The nitrocefin assay in MT-plate format. Sample **A1**: 20 μL $_{m}$CpG-$TFBS_{Zif268}$ dsDNA-PMPs with 160 ng (6 pmoles) of bound DNA plus Bla1-Zif268 and MBD2-Bla2; **A2**: 10 μL $_{m}$CpG-$TFBS_{Zif268}$ dsDNA-PMPs with 80 ng (3 pmoles) of bound DNA plus Bla1-Zif268 and MBD2-Bla2; **A3**: 20 μL $_{m}$CpG-$TFBS_{Zif268}$ dsDNA-PMPs and Bla1-Zif268; **A4**: 20 μL $_{m}$CpG-$TFBS_{Zif268}$ dsDNA-PMPs and Bla2-MBD2; **B1**: 20 μL CpG-$TFBS_{Zif268}$ dsDNA-PMPs (i.e., non-methylated) plus Bla1-Zif268 and MBD2-Bla2; **B2**: 20 μL random-dsDNA-PMPs plus Bla1-Zif268 and MBD2-Bla2; **B3**: 20 μL $_{m}$CpG-$TFBS_{Zif268}$ dsDNA-PMPs and the periplasmic fraction of the *E.coli* host; **B4**: 20 μL SA-PMPs, Bla1-Zif268 and MBD2-Bla2. The amount of DNA attached to the paramagnetic particles used in the samples A3 to B4 was 160 ng (6 pmoles).

The major challenge in developing methods to study diverse biological processes *in vitro* is to be able to extrapolate the results to the *in vivo* cellular events. This issue is especially relevant to the study of methylation-dependent modulation of gene expression. The use of SA-PMPs permits generation of the particles with various percentage of cognate methylated-dsDNA on the "background" of cognate nonmethylated-dsDNA and this could mimic the variation in the levels of biological genome methylation. One may expect that more engineered proteins would assemble onto the larger number of

methylated-dsDNA sites as compared to the particles carrying less of these sites and this would correspond to quantitative differences in the observed enzyme activity. Indeed, the measured v_0 values for the reassembled TEM-1 β-lactamase confirmed the aforementioned hypothesis (Table 1).

The availability of rapid and robust methods for detection of DNA methylation is important for elucidating roles DNA methylation-based gene regulation and related epigenetic processes. The strategy presented here should be useful for selective, qualitative and quantitative analysis of methylated-DNA - protein interactions through the enzymatic activity of the reassembled tag. Because under the usual *in vitro* assay conditions reaction rate is directly proportional to the enzyme concentration – and in this study also to the concentration of the fused DNA-binding proteins – the degree DNA methylation can be simply evaluated in the terms of the enzyme arbitrary units. The work presented here relies on the concept of the enzyme/protein-fragment complementation (PFC) that was developed for the quantitative detection of dynamic protein-protein interactions *in vitro* and *in vivo* [21]. The crucial feature of PFC is that the fragments are designed not to fold spontaneously without being brought into close proximity by the interaction of the proteins to which they are fused. Without a spontaneous folding, there is no a false positive signal and this is in contrast to the assay systems that rely on naturally occurring and spontaneously associating subunits of the enzymes fused to interacting proteins. The central problem of those assays is that subunits, even if weakly associating, are always able to do this to some extend, meaning that there is a constant background of spontaneous assembly.

Although a number of cell biological applications use GFP and other fluorescent proteins as tags [22], the split-enzyme reporter was selected instead of GFP for a number of reasons. For example, the somewhat disappointing sensitivity of the GFP tag is probably inherent result of its lack of signal amplification. Unlike enzymes, GFP can not catalytically process an indefinite number of substrate molecules. Each GFP molecule produces at most one fluorophore. It has been estimated that 1 μM well-folded wild-type GFP molecules is required to equal the endogenous autofluorescence of a typical mammalian cell [23], that is, to double the fluorescence over the background. Mutant GFPs with improved extinction coefficients might improve this detection limit, but still 0.1 μM GFP is approx. 10^5 copies per typical cell of 1 – 2 pL volume. Moreover, this estimate already assumes perfect GFP maturation; any maturation/folding problems would raise the threshold copy number even further.

In conclusion, an approach facilitating detection of methylated dsDNA was developed. One of the most attractive aspects of this strategy would be to extend some of its concepts to develop an assay functioning in the *in vivo* cellular setting, and this is currently undergoing investigation.

This work was supported by the Chemical Engineering Instructor Program, School of Engineering, Rensselaer Polytechnic Institute.

REFERENCES

[1] F. Antequera and A. Bird, Proc. Natl. Acad. Sci. USA 90, 11995 – 11999, 1994.
[2] B. Neumann and D. P. Barlow, Curr. Opin. Genet. Dev. 6, 159 – 163, 1996.
[3] A. Razin and H. Cedar, Cell 77, 473 – 476, 1994.
[4] E. Li, T. H. Bestor and R. Jaenisch, Cell 69, 915 – 926, 1992.
[5] M. Comb and H. M. Goodman, Nucleic Acid Res. 18, 3975 – 3982, 1990.
[6] S. M. M. Iguchi-Ariga and W. Schaffner, Genes Dev. 3, 612 – 619, 1989.
[7] H. Cedar, Cell 53, 3 – 4, 1988.
[8] S. U. Kass, J. P. Goddard and R. L. P. Adams, Mol. Cell. Biol. 13, 7372 – 7379, 1993.
[9] G. Buschhausen, B. Wittig, M. Graessmann and A. Graessmann, Proc. Natl. Acad. Sci. USA 84, 1177 – 1181, 1987.
[10] B. Hendrich and A. Bird, Mol. Cell. Biol. 18, 6538 – 6547, 1998.
[11] M. F. Fraga, E. Ballestar, G. Montoya, P. Taysavang and P. A. Wade, Nucleic Acid Res. 31, 1765 – 1774, 2003.
[12] N. P. Pavletich and C. O. Pabo, Science 252, 809 – 817, 1991.
[13] L. Maveyraud, R. F. Pratt and J.-P. Samama, Biochemistry 37, 2622 – 2628, 1998.
[14] A. Galarneau, M. Primeau, L.-E. Trudeau and S. W. Michnick, Nature Biotechnol. 20, 619 – 622, 2002.
[15] J. Sambrook, E. F. Fritsch and T. Maniatis, "Molecular Cloning: A Laboratory Manual," 2nd ed., CSH Laboratory Press, Cold Spring Harbor, NY, 1989.
[16] J. Sambrook and J. D. W. Russell, "Molecular Cloning: A Laboratory Manual," 3rd ed., CSH Laboratory Press, Cold Spring Harbor, NY, 2001.
[17] A. Galarneau, M. Primeau, L.-E. Trudeau and S. W. Michnick, Nature Biotechnol. 20, 619 – 622, 2002.
[18] V. Siedarki, W. Huang, T. Palzkill and H. F. Gilbert, Proc. Natl. Acad. Sci. USA 98, 283 – 288, 2001.
[19] G. E. Sroga and J. S. Dordick, Biotechnol. Bioeng. 78, 761 – 769, 2002.
[20] G. E. Sroga and J. S. Dordick, Biotechnol. Bioeng. 94, 312 – 321, 2006.
[21] I. Remy and S. W. Michnick, BioTechniques 42, 137 – 144, 2007.
[22] I. Ghosh, C. I. Stains, A. T. Ooi and D. J. Segal, Mol. BioSyst. 2, 551 – 560, 2006.
[23] K. D. Niswender, S. M. Blackman, L. Rhode, M. A. Magnuson and D. W. Piston, J. Microsc. 180, 109 – 116, 1995.