Anal. Chem. **2007,** *79,* 6702−6708

Split β-Lactamase Sensor for the Sequence-Specific Detection of DNA Methylation

Jason R. Porter,† Cliff I. Stains,† David J. Segal,‡ and Indraneel Ghosh*,†

Department of Chemistry, University of Arizona, Tucson, Arizona 85721, and Department of Medical Pharmacology and Toxicology, University of California, Davis, California 95616

The methylation pattern of genes at CpG dinucleotide sites is an emerging area in epigenetics. Furthermore, the hypermethylation profiles of tumor suppressor genes are linked to specific tumor types. Thus, new molecular approaches for the rapid determination of the methylation status of these genes could provide a powerful method for early cancer diagnosis as well as insight into mechanisms of epigenetic regulation of genetic information. Toward this end, we have recently reported the first design of a split-protein sensor for the site-specific detection of DNA methylation. In this approach a split green fluorescent protein reporter provided a sequence-specific readout of CpG methylation. In the present work, we describe a sensitive second-generation methylation detection system that utilizes the split enzymatic reporter, TEM-1 β-lactamase, attached to specific DNA binding elements. This system, termed mCpG-SEER-β-Lac, shows a greater than 40-fold specificity for methylated versus nonmethylated CpG target sites. Importantly, the resulting signal enhancement afforded by the catalytic activity of split-β-lactamase allowed for the sensitive detection of 2.5 fmol of methylated target dsDNA in 5 min. Thus, this new sensor geometry represents a 250-fold enhancement in assay time and a 2000-fold enhancement in sensitivity over our first-generation system for the detection of specific sites of DNA methylation.

In addition to the myriad of regulatory genes encoded by the human genome, there exists a second and perhaps equally complex level of regulation based on the selective methylation of DNA. This epigenetic control is dependent on the transfer of a methyl group from *S*-adenosylmethionine to the C5-position of cytosines present in CpG dinucleotides (Figure 1A) and is regulated by DNA methyltransferases, Me-CpG binding proteins, such as the mCpG binding protein MBD2, and demethylation via nucleotide excision or other DNA repair mechanisms.[1−3] This cytosine-specific methylation is present in ∼70% of all CpG dinucleotides found in mammalian somatic cells and contributes to genome stability, repression of transposable elements, and transcriptional silencing through the recruitment of chromatin modification complexes.[4−6] Though CpG methylation is distributed throughout the genome, this chemical modification is normally excluded from promoter-associated CpG-rich regions of a sequence known as CpG islands as methylation is associated with translational repression. Thus, the observed aberrant methylation of tumor suppressor gene associated promoter regions likely results in gene silencing and can be linked to a number of human cancers. Several recent studies have shown that methylation of these CpG islands occurs in a sequence and tumor-type specific manner, leading to the identification of gene hypermethylation profiles for a number of human cancer types.[7,8]

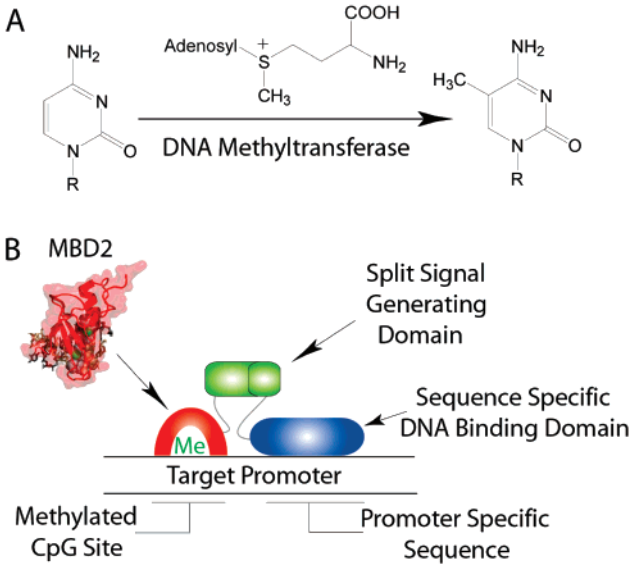

Figure 1. Site-specific determination of promoter methylation (A) C5-Cytosine methylation by DNA methyltransferases (R = deoxyribose). (B) Promoter-specific methylation detection strategy. A methyl-CpG binding protein (MBD2, red) will be used to determine the methylation status of the adjacent CpG site. A sequence-specific DNA binding protein (blue) will direct the sensor to the desired promoter by targeting a unique DNA sequence adjacent to a CpG site being probed for methylation. Simultaneous binding of the two proteins will result in the reassembly of an appended split signal generating domain and produce a detectable signal.

* To whom correspondence should be addressed. E-mail: ghosh@email.arizona.edu. Phone: (520) 621-6331. Fax: (520) 621-8407.
† Department of Chemistry.
‡ Department of Medical Pharmacology and Toxicology.

(1) Bird, A. *Genes Dev.* **2002**, *16*, 6−21.
(2) Berger, J.; Bird, A. *Biochem. Soc. Trans.* **2005**, *33*, 1537−1540.
(3) Barreto, G.; Schafer, A.; Marhold, J.; Stach, D.; Swaminathan, S. K.; Handa, V.; Doderlein, G.; Maltry, N.; Wu, W.; Lyko, F.; Niehrs, C. *Nature* **2007**, *445*, 671−675.
(4) Ehrlich, M.; Gama-Sosa, M. A.; Huang, L. H.; Midgett, R. M.; Kuo, K. C.; McCune, R. A.; Gehrke, C. *Nucleic Acids Res.* **1982**, *10*, 2709−2721.
(5) Meehan, R. R.; Stancheva, I. *Essays Biochem.* **2001**, *37*, 59−70.
(6) Jaenisch, R.; Bird, A. *Nat. Genet.* **2003**, *33 Suppl*, 245−254.

10.1021/ac071163+ CCC: $37.00
Published on Web 08/08/2007

The ability to directly determine CpG methylation in a promoter-specific manner would undoubtedly provide a powerful tool for the detection and possible determination of specific cancer types. Current methylated CpG detection methods are PCR based and employ the bisulfate modification of DNA followed by DNA sequencing or methylation-specific PCR. This treatment provides a basis for the distinction between methylated and nonmethylated cytosines through the conversion of nonmethylated cytosines to uracil while leaving methylated cytosines unchanged.[9,10] Though effective, a number of problems are associated with these techniques, namely, extensive bisulfate reaction times (4−18 h) as well as false positives due to incomplete bisulfate reactions.[11,12] Currently there are no molecular methods available for the direct readout of sequence-specific CpG methylation utilizing antisense-based chemical approaches.

In order to develop a direct, rapid, and PCR-independent method for the site-specific detection of CpG methylation, we have recently designed a sequence-enabled reassembly (SEER)-based system that utilizes a split-reporter protein in conjunction with sequence-specific and methylation-specific DNA binding proteins. To date, a variety of proteins have been adapted for use as split reporters including the bioluminescent firefly,[13] *Renilla*[14] and *Guassia*[15] luciferases, the green fluorescent protein and its variants,[16−18] and β-lactamase.[19] The use of this technology has been primarily applied to the study of a number of molecular interactions including interacting protein pairs,[20] ligand induced[21] and phosphorylation[22] dependent protein−protein interactions, ligand−receptor interactions,[23] and protein−DNA interactions,[24,25] as well as determining the cellular localization of these interactions.[26,27] Additionally, split protein reporters have been used to image these interactions in vitro and in vivo in both cells and living animals.

Generally, our split-reporter SEER system employs two DNA binding domains, which are capable of reassembling two halves of a tethered split-reporter protein only in the presence of a specific DNA sequence of interest.[24] To create a system capable of site-specific detection of CpG methylation, termed mCpG-SEER, we envisioned that a sequence-specific DNA binding domain tethered to half of split-GFP and a methyl-CpG recognition domain to the other half of the split-GFP would result in a sensor architecture capable of reading out sequence-specific methylation events. Key to this design are the methyl-CpG binding domain (MBD) family of proteins whose members include the MBD1, MBD2, MBD3, MBD4, and MeCP2 proteins.[28,29] Previous studies have demonstrated a high degree of structural similarity between these proteins as well as the specific contacts made between them and the major groove of DNA.[30−32] Thus, our previously reported system based on split-GFP, incorporated the methyl-CpG binding domain, MBD2, in conjunction with the zinc-finger, Zif268, and was able to reassemble the split-GFP in a DNA sequence- and methylation-specific manner. Additionally, this GFP-based mCpG-SEER system was found to be highly specific for methylated CpG sites, showing a greater than 40-fold increase in signal generation in the presence of methylated DNA over nonmethylated DNA.[33] Thus, mCpG-SEER, presents a general strategy for the development of protein-based sensors capable of interrogating CpG methylation in a promoter-specific manner. In these systems, a designed sequence-specific DNA binding protein[34−36] can be used to direct analysis to a DNA sequence unique to the promoter of interest, while a methylation-specific protein can be used to recognize an adjacent methylated CpG site. Simultaneous binding and formation of the ternary complex will result in the reassembly of the appended split reporter resulting in a detectable signal (Figure 1B).

Though capable of site-specific detection of DNA methylation, this initial GFP-based system has drawbacks, namely, insoluble component proteins that couple signal generation to a lengthy denaturation and refolding process. This need for refolding, in conjunction with a slow rate of GFP chromophore formation, results in a minimum assay time of 24 h for maximal fluorescence in the previous system. The current study aims to significantly improve the sensitivity of mCpG-SEER through the incorporation of the enzymatic split reporter TEM-1 β-lactamase. β-Lactamase

(7) Costello, J. F.; Fruhwald, M. C.; Smiraglia, D. J.; Rush, L. J.; Robertson, G. P.; Gao, X.; Wright, F. A.; Feramisco, J. D.; Peltomaki, P.; Lang, J. C.; Schuller, D. E.; Yu, L.; Bloomfield, C. D.; Caligiuri, M. A.; Yates, A.; Nishikawa, R.; Su Huang, H.; Petrelli, N. J.; Zhang, X.; O'Dorisio, M. S.; Held, W. A.; Cavenee, W. K.; Plass, C. *Nat. Genet.* **2000**, *24*, 132−138.
(8) Esteller, M.; Corn, P. G.; Baylin, S. B.; Herman, J. G. *Cancer Res.* **2001**, *61*, 3225−3229.
(9) Furuichi, Y.; Wataya, Y.; Hayatsu, H.; Ukita, T. *Biochem. Biophys. Res. Commun.* **1970**, *41*, 1185−1191.
(10) Clark, S. J.; Harrison, J.; Paul, C. L.; Frommer, M. *Nucleic Acids Res.* **1994**, *22*, 2990−2997.
(11) Herman, J. G.; Graff, J. R.; Myohanen, S.; Nelkin, B. D.; Baylin, S. B. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 9821−9826.
(12) Fraga, M. F.; Esteller, M. *Biotechniques* **2002**, *33*, 632, 634, 636−649.
(13) Ozawa, T.; Kaihara, A.; Sato, M.; Tachihara, K.; Umezawa, Y. *Anal. Chem.* **2001**, *73*, 2516−2521.
(14) Paulmurugan, R.; Gambhir, S. S. *Anal. Chem.* **2003**, *75*, 1584−1589.
(15) Remy, I.; Michnick, S. W. *Nat. Methods* **2006**, *3*, 977−979.
(16) Ghosh, I.; Hamilton, A. D.; Regan, L. *J. Am. Chem. Soc.* **2000**, *122*, 5658−5659.
(17) Ozawa, T.; Nogami, S.; Sato, M.; Ohya, Y.; Umezawa, Y. *Anal. Chem.* **2000**, *72*, 5151−5157.
(18) Hu, C. D.; Kerppola, T. K. *Nat. Biotechnol.* **2003**, *21*, 539−545.
(19) Galarneau, A.; Primeau, M.; Trudeau, L. E.; Michnick, S. W. *Nat. Biotechnol.* **2002**, *20*, 619−622.
(20) Paulmurugan, R.; Umezawa, Y.; Gambhir, S. S. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 15608−15613.
(21) Paulmurugan, R.; Gambhir, S. S. *Anal. Chem.* **2007**, *79*, 2346−2353.
(22) Luker, K. E.; Smith, M. C.; Luker, G. D.; Gammon, S. T.; Piwnica-Worms, H.; Piwnica-Worms, D. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 12288−12293.
(23) Kim, S. B.; Awais, M.; Sato, M.; Umezawa, Y.; Tao, H. *Anal. Chem.* **2007**, *79*, 1874−1880.
(24) Stains, C. I.; Porter, J. R.; Ooi, A. T.; Segal, D. J.; Ghosh, I. *J. Am. Chem. Soc.* **2005**, *127*, 10782−10783.
(25) Ooi, A. T.; Stains, C. I.; Ghosh, I.; Segal, D. J. *Biochemistry* **2006**, *45*, 3620−3625.
(26) Kaihara, A.; Kawai, Y.; Sato, M.; Ozawa, T.; Umezawa, Y. *Anal. Chem.* **2003**, *75*, 4176−4181.
(27) Paulmurugan, R.; Gambhir, S. S. *Anal. Chem.* **2005**, *77*, 1295−1302.
(28) Hendrich, B.; Bird, A. *Mol. Cell. Biol.* **1998**, *18*, 6538−6547.
(29) Hendrich, B.; Abbott, C.; McQueen, H.; Chambers, D.; Cross, S.; Bird, A. *Mamm. Genome* **1999**, *10*, 906−912.
(30) Wakefield, R. I.; Smith, B. O.; Nan, X.; Free, A.; Soteriou, A.; Uhrin, D.; Bird, A. P.; Barlow, P. N. *J. Mol. Biol.* **1999**, *291*, 1055−1065.
(31) Ohki, I.; Shimotake, N.; Fujita, N.; Nakao, M.; Shirakawa, M. *EMBO J.* **1999**, *18*, 6653−6661.
(32) Ohki, I.; Shimotake, N.; Fujita, N.; Jee, J.; Ikegami, T.; Nakao, M.; Shirakawa, M. *Cell* **2001**, *105*, 487−497.
(33) Stains, C. I.; Furman, J. L.; Segal, D. J.; Ghosh, I. *J. Am. Chem. Soc.* **2006**, *128*, 9761−9765.
(34) Pabo, C. O.; Peisach, E.; Grant, R. A. *Annu. Rev. Biochem.* **2001**, *70*, 313−340.
(35) Beerli, R. R.; Barbas, C. F., 3rd. *Nat. Biotechnol.* **2002**, *20*, 135−141.
(36) Ghosh, I.; Stains, C. I.; Ooi, A. T.; Segal, D. J. *Mol. Biosyst.* **2006**, *2*, 551−560.

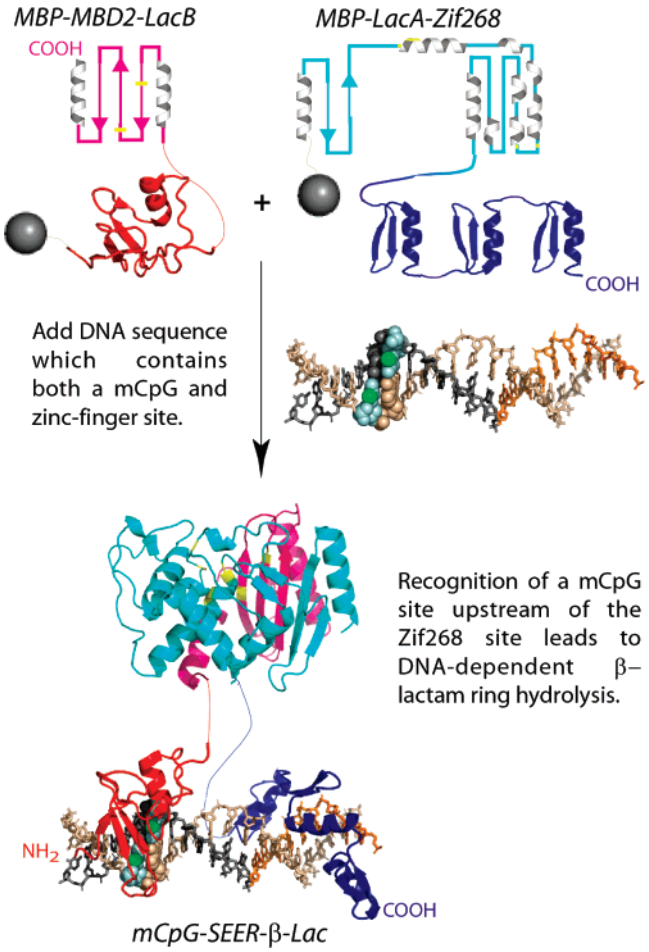

Figure 2. mCpG-SEER-β-Lac system. MBD2-LacB (red and pink) and LacA-Zif268 (cyan and blue) are shown in ribbon form with an N-terminally expressed maltose binding protein (MBP) affinity tag (gray sphere) and highlighted TEM-1 β-lactamase active residues (yellow). The mCpG site is shown as spheres with cytosines (cyan) and methyl groups (green) shown as spheres. The Zif268 binding site (5′-GCGTGGGCG-3′) is shown in orange.

possesses a number of positive attributes for its use in mCpG-SEER, including its small size (29 kDa), ease of expression, and most importantly its prior demonstration as a split-protein reporter for detecting protein−protein interactions.[19] Additionally, a number of substrates are available that allow both colorimetric and fluorogenic detection of β-lactamase activity. Finally, we have previously demonstrated that the split β-lactamase system tethered to zinc-fingers can be utilized to read out normal dsDNA sequence information.[25] Thus, to construct this new β-lactamase-based mCpG-SEER system, referred to herein as mCpG-SEER-β-Lac (Figure 2), the DNA binding proteins MBD2 and Zif268 were chosen for use in our proof-of-principle experiment. The use of MBD2, which binds mCpG sites with a reported 2.7 nM affinity while binding analogous nonmethylated CpG sites with a reported 70-fold decrease in affinity,[37] should provide adequate discrimination between methylated and nonmethylated CpG sites. As previously discussed, in order to uniquely recognize a specific site of DNA methylation, we need to impart sequence specificity to our split protein sensor. The zinc-finger Zif268 (K_d = 10−14 pM)[38] was chosen to impart sequence specificity as demonstrated in the design of our two previous SEER systems. Each mCpG-SEER-β-Lac component was designed to contain a 15-amino acid linker between the DNA binding and split β-lactamase domains to ensure conformational flexibility to allow for ternary complex formation.

MATERIALS AND METHODS

General Materials. 2xYT media was purchased from Becton Dickinson, DTT was purchased from Research Products International, and $ZnCl_2$ was obtained from EM Sciences. All other materials were obtained from Sigma-Aldrich unless otherwise noted.

Cloning, Expression, and Purification of mCpG-SEER-βLac Proteins. Human MBD2 DNA was obtained by PCR from a pUC19 vector containing an *Escherichia coli* codon optimized form of the human MBD2 gene (residues 147−215) (GenScript). The following primers were used to clone MBD2 into an existing pMAL-c2X vector containing a 15-amino acid linker followed by the C-terminal portion of TEM-1 β-lactamase (LacB), residues 198−290: Fwd, 5′-GCGCGCGAATTCGAAAGCGGCAAACGC-3′; and Rev, 5′-CGGTTAACCGGTCAT TTTGCCGGTACG-3′. The resulting pMAL-c2X plasmid containing the MBD2-LacB gene was transformed by electroporation into BL21 (DE3) cells (Novagen). An overnight culture of these cells was used to inoculate a 100-mL culture of 2xYT media supplemented with 100 μg/mL ampicillin, 100 μM $ZnCl_2$, and 2% glucose at an OD_{600} of 0.05. Protein expression was induced at an OD_{600} of 0.8 with 1 mM isopropyl β-D-thiogalactopyranoside for 4 h at 37 °C. Cells were pelleted by centrifugation and frozen overnight at −20 °C. Pelleted cells were thawed on ice and resuspended in zinc buffer A (100 mM Tris at pH 7.5, 90 mM KCl, 1 mM $MgCl_2$, 100 μM $ZnCl_2$, and 5 mM DTT). Resuspended cells were lysed by sonication. The lysate was cleared by centrifugation at 18 000 cfm for 30 min. MBP-tagged MBD2-LacB was purified over amylose resin (New England Biolabs) using the following protocol. 3 mL of resin was loaded into a PD-10 column and equilibrated with 5 column volumes of zinc buffer A. Following the addition of cleared lysates, the column was washed with 10 column volumes of zinc buffer A supplemented with 2 M NaCl followed by an additional wash with 8 column volumes of zinc buffer A. MBP-MBD2-LacB was eluted with 5 mL of zinc buffer A supplemented with 10 mM maltose and was characterized by SDS-PAGE. Concentrations were determined in 6 M guanidine using absorbance measurements at 280 nm (calculated ϵ = 94 310 M^{-1} cm^{-1}).

The MBP-LacA-Zif268 fusion protein (containing residues 26−196 of TEM-1 β-lactamase) was expressed and purified as described above. MBP-LacA-Zif268 was characterized by SDS-PAGE; concentrations were determined in 6 M guanidine using absorbance measurements at 280 nm (calculated ϵ = 79 300 M^{-1} cm^{-1}).

Colorimetric Activity Assays. DNA oligonucleotides containing target DNA sequences (HPLC-purified) were obtained from IDT, and sheared herring sperm DNA was obtained from Invitrogen. Target DNA oligos in 1× BamHI buffer (NEB) were annealed using a Techne Genius thermocycler with the following protocol: 95 °C for 7 min, decrease to 56 °C at 1 °C/min, with

(37) Fraga, M. F.; Ballestar, E.; Montoya, G.; Taysavang, P.; Wade, P. A.; Esteller, M. *Nucleic Acids Res.* **2003**, *31*, 1765−1774.

(38) Nalefski, E. A.; Nebelitsky, E.; Lloyd, J. A.; Gullans, S. R. *Biochemistry* **2006**, *45*, 13794−13806.

equilibration for 5 min at 56 °C, and finally decrease to 25 °C at 1 °C/min.

All experiments were conducted in 96-well plates (CoStar, Corning). All target DNA sequences used for determining mCpG-SEER-β-Lac specificity and target-site spacing dependence are given in Figure 5 and Supporting Information Table S1, respectively. All experiments were performed in triplicate.

The background activity of each mCpG-SEER-β-Lac component was interrogated by incubating MBP-MBD2-LacB or MBP-LacA-Zif268 alone (125 nM each) in zinc buffer A (volume of 180 μL) for 30 min at room temperature. Following incubation, 20 μL of 1 mM nitrocefin in phosphate buffer (100 mM, pH 7.0) was added to a final concentration of 100 μM and a final volume of 200 μL.

To determine mCpG-SEER-β-Lac dependence on target site spacing and target site specificity, MBP-MBD2-LacB and MBP-LacA-Zif268 (125 nM each) were mixed with target dsDNA (5 nM) in zinc buffer A to a volume of 180 μL and allowed to incubate for 30 min at room temperature. Following the 30-min incubation, 20 μL of 1 mM nitrocefin in phosphate buffer (100 mM, pH 7.0) was added to a final concentration of 100 μM and a final volume of 200 μL.

To determine the minimum amount of dsDNA needed to produce a detectable signal, MBP-MBD2-LacB and MBP-LacA-Zif268 (125 nM each) were mixed in zinc buffer A, in a volume of 180 μL, with decreasing concentrations of a dsDNA target containing a 2-bp spacing between the mCpG and Zif268 binding sites. Following a 30-min incubation at room temperature, 20 μL of 1 mM nitrocefin in phosphate buffer (100 mM, pH 7.0) was added to a final concentration of 100 μM and a final volume of 200 μL. For experiments using the fluorescent substrate CCF2FA, 10 μL of 100 μM CCF2FA in PBS was added to a 90 μL of solution containing mCpG-SEER-β-Lac proteins (125 nM each) and target dsDNA to a final concentration of 10 μM in a final volume of 100 μL. Generated signal is defined as the rate of CCF2FA hydrolysis between 30 and 50 min after CCF2FA addition.

Absorbance Measurements. All absorbance spectra were acquired on a μQuant plate reader (Bio-Tek Instruments) at 486 nm over a period of 15 min. Signal generation is defined as the rate of nitrocefin hydrolysis between 0.5 and 2 min after nitrocefin addition. Activity was determined by subtracting background rate of hydrolysis for mCpG-SEER-β-Lac proteins at the same concentration in the absence of DNA; these samples were then normalized to the sample with the highest rate of hydrolysis.

Fluorescence Measurements. All fluorescence measurements were acquired on a Spectra Max Gemini (Molecular Devices) plate reader with excitation at 409 nm and monitored emission at 450 and 520 nm. Cutoffs were set at 435 and 495 nm for the 450- and 520-nm emissions, respectively. Given the decreased rate of catalysis of CCF2FA (k_{cat} = 29 s^{-1})[39] compared to nitrocefin (k_{cat} = 930 s^{-1})[40] a 30-min incubation period after CCF2FA addition was required. After incubation, fluorescence measurements were acquired at 450 and 520 nm every 2 min for 20 min. Activity was determined by subtracting background signal for mCpG-SEER-β-Lac proteins at the same concentration without DNA.

(39) Zlokarnik, G.; Negulescu, P. A.; Knapp, T. E.; Mere, L.; Burres, N.; Feng, L.; Whitney, M.; Roemer, K.; Tsien, R. Y. *Science* **1998**, *279*, 84−88.
(40) Raquet, X.; Lamotte-Brasseur, J.; Fonze, E.; Goussard, S.; Courvalin, P.; Frere, J. M. *J. Mol. Biol.* **1994**, *244*, 625−639.

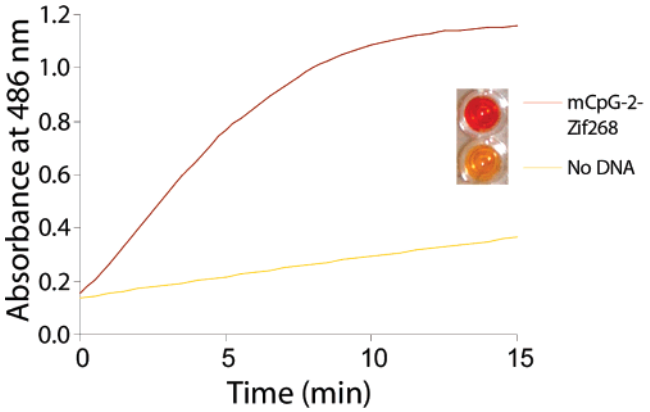

Figure 3. Absorbance spectra of MBP-LacA-Zif268 (125 nM) + MBP-MBD2-LacB (125 nM) in the presence (red) and absence (yellow) of 5 nM dsDNA target (mCpG-2-Zif268) in zinc buffer A.

RESULTS

Initial mCpG-SEER-β-Lac Activity. β-Lactamase activity assays were conducted using nitrocefin, a chromogenic substrate that undergoes a distinct color change from yellow (λ_{max} = 390 nm) to red (λ_{max} = 486 nm) upon β-lactam hydrolysis. Previous mCpG-SEER-GFP studies showed that a dsDNA target containing a 2-bp separation between the methylated CpG and zinc-finger binding sites provides an adequate signal upon protein binding.[33] To evaluate the use of split β-lactamase for the detection of methylated CpG sites, the hydrolysis activity of 125 nM MBP-LacA-Zif268 plus 125 nM MBP-MBD2-LacB was evaluated in the presence and absence of 5 nM mCpG-2-Zif268, a methylated dsDNA target containing a 2-bp spacing between the mCpG and the zinc-finger binding sites. Proteins were incubated for 30 min followed by the addition of nitrocefin to a final concentration of 100 μM. Absorbance at 486 nm was monitored for 15 min after the addition of nitrocefin. A marked increase in nitrocefin hydrolysis was observed in the presence of the dsDNA target mCpG-2-Zif268, as opposed to minimal hydrolysis in the absence of dsDNA (Figure 3). This initial observation of a marked increase in activity only in the presence of a target methylated dsDNA served to initially validate our design of mCpG-SEER-β-Lac and its ability to address specific sites of DNA methylation.

Background Activity of mCpG-SEER-β-Lac Proteins. To interrogate the possibility of alternate sources of background activity of mCpG-SEER-β-Lac, nitrocefin hydrolysis was monitored in the presence of each mCpG-SEER-β-Lac protein (125 nM) alone, along with both proteins in the presence and absence of the dsDNA target mCpG-2-Zif268 (5 nM). No nitrocefin hydrolysis was detected for either mCpG-SEER-β-Lac protein alone, and minimal hydrolysis was observed with both proteins in the absence of target dsDNA (Figure 4). This minimal background, a result of self-association of the fusion proteins also observed by Michnick and us,[19,25] contributes to less than 3% of the total signal observed in the presence of target dsDNA (5 nM). Overall, this demonstrated that the signal produced is DNA sequence dependent and facilitated by the simultaneous DNA binding of each mCpG-SEER-β-Lac protein to afford an enzymatically active ternary complex.

mCpG-SEER-β-Lac Binding Site Specificity. In order for mCpG-SEER-β-Lac to serve as a tool for interrogating promoter-specific methylation, the system must be able to distinguish

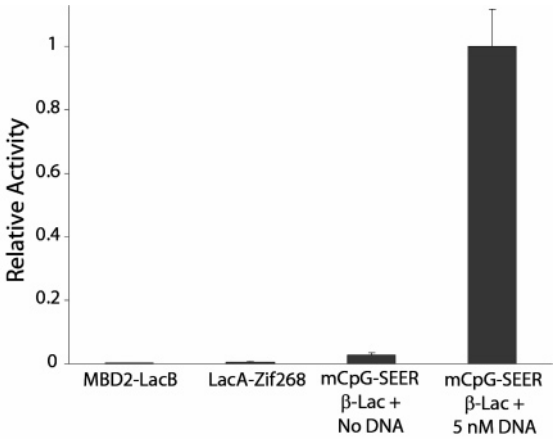

Figure 4. Nitrocefin hydrolysis in the presence of either MBP-MBD2-LacB or MBP-LacA-Zif268 (125 nM) alone, and mCpG-SEER-β-Lac (MBP-MBD2-LacB (125 nM) + MBP-LacA-Zif268 (125 nM) in the presence and absence of dsDNA target mCpG-2-Zif268.

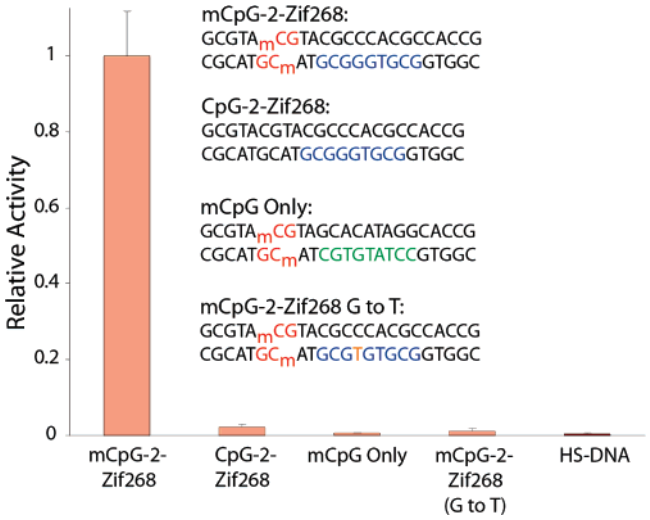

Figure 5. Nitrocefin hydrolysis in the presence of MBP-LacA-Zif268 (125 nM) + MBP-MBD2-LacB (125 nM) in the presence of indicated dsDNA targets.

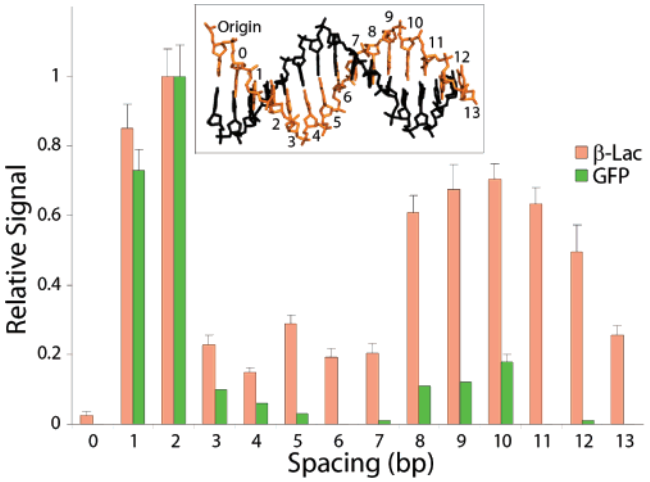

Figure 6. Relative signal of mCpG-SEER-β-Lac (125 nM) and mCpG-SEER-GFP in the presence of dsDNA target as a function of increased number of base pairs between the mCpG and Zif268 sites (inset).

methylated from nonmethylated CpG sites and to differentiate promoter-specific target sequences from nontarget sequences. To interrogate the binding site specificity of the mCpG-SEER-β-Lac system, a series of target dsDNA sequences were designed with mutations intended to abolish binding of either MBD2 or Zif268 to their respective binding sites. The following dsDNA target sequences containing a 2-bp spacing between the mCpG island and the Zif268 binding site were used: (1) nonmethylated CpG (CpG-2-Zif268), (2) Zif268 site absent (mCpG only), (3) a single G to T base mutation in the second finger recognition site of the Zif268 binding site (mCpG-2-Zif268, G to T), and (4) herring sperm DNA (15 ng). Mixtures of the two proteins, MBP-MBD2-LacB (125 nM) and MBP-LacA-Zif268 (125 nM), were allowed to incubate for 30 min in the presence of 5 nM concentrations of each dsDNA target described above followed by hydrolysis rate determination.

The observed rates of hydrolysis demonstrate the ability of mCpG-SEER-β-Lac to distinguish cognate binding sites from nonspecific, nonmethylated, or mutated binding sites (Figure 5). In the presence of a nonmethylated CpG site, the observed activity is 42-fold less than that observed in the presence of a methylated CpG site, showing the ability of mCpG-SEER-β-Lac to clearly distinguish methylated from nonmethylated CpG sites. In addition to methylation, the need for a cognate Zif268 binding site is apparent in the 170-fold activity decrease in the absence of a Zif268 binding site. As a testament to the superb sequence specificity of Zif268, the single G to T base mutation resulted in a 93-fold decrease in activity compared to the nonmutated Zif268 site. Last, the 220-fold decrease in activity in the presence of nonspecific herring sperm DNA reaffirms the requirement of both a methylated CpG site and a specific zinc-finger binding site for signal generation.

Distance Dependence of mCpG-SEER-β-Lac Activity. To evaluate the effect of distance and proximity on signal generation, a series of dsDNA targets were designed that have an increasing distance between the mCpG and Zif268 binding sites. In all, 14 dsDNA targets were constructed placing 0−13 bps between the mCpG and Zif268 binding sites. Mixtures of the two mCpG-SEER-β-Lac proteins, MBP-MBD2-LacB (125 nM) and MBP-LacA-Zif268 (125 nM), were allowed to incubate for 30 min in the presence of 5 nM concentrations of each dsDNA target.

The results from these experiments provide some interesting insight into the signal generation profile of the mCpG-SEER-β-Lac system (Figure 6, tan). Virtually no activity is observed when the MBD2 and zinc-finger binding domains are placed next to one another (0-bp spacing). This is most likely due to the fact that the MBD2 binding domain occupies at least 1 bp on the surface of DNA beyond the canonical mCpG recognition site. As the distance between the two binding sites increases to 1 and 2 bps, a sharp increase in activity is observed, with maximal signal being generated when the mCpG and zinc-finger binding sites are separated by 2 bps. As the number of bps increases, a stepwise decrease in activity is observed for spacings of 3−7 bps, with a marked increase occurring again with a second signal maximum (70% of the signal generated at a distance of 2 bps) when the two binding sites are separated by 10 bps. As the distance between the two sites increases from 11 to 13 bps, a decrease in activity is again observed. This pattern serves to graphically recapitulate the

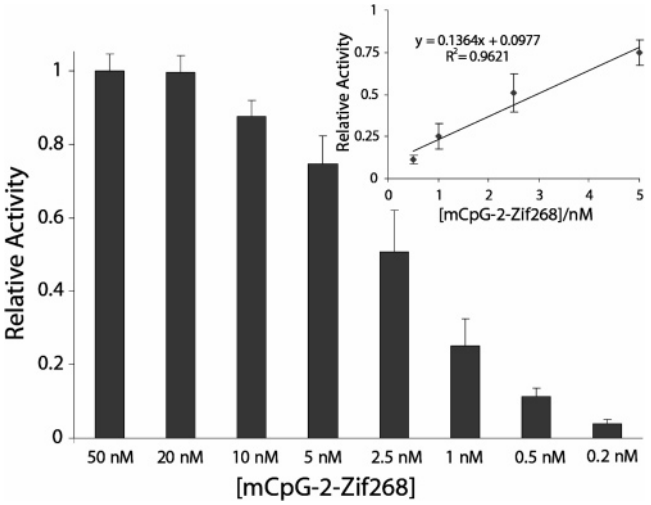

Figure 7. Relative hydrolysis of nitrocefin in the presence of MBP-LacA-Zif268 (125 nM) + MBP-MBD2-LacB (125 nM) in the presence of decreasing concentrations of dsDNA target mCpG-2-Zif268. (Inset) Linear fit from 5 to 0.5 nM mCpG-2-Zif268.

helical turn of dsDNA by not only increasing the linear distance between the bound mCpG-SEER-β-Lac proteins but also varying the orientation of each protein relative to each other on the surface of dsDNA. Interestingly a very similar pattern was observed when GFP is used as the signal generating protein (Figure 6, green), but shows significantly less signal generation at separations of 8−10 bps.

dsDNA Concentration Dependence of mCpG-SEER-β-Lac Activity. The above experiments demonstrated that the maximal signal intensity is obtained when the target dsDNA oligo contains a 2-bp spacing between the mCpG and zinc-finger binding sites. With this is mind, the dsDNA target oligo mCpG-2-Zif268 was used to determine the minimum amount of dsDNA that can be detected by mCpG-SEER-β-Lac. Mixtures of the two mCpG-SEER-β-Lac proteins, MBP-MBD2-LacB (125 nM) and MBP-LacA-Zif268 (125 nM), were allowed to incubate for 30 min in the presence of decreasing concentrations of mCpG-2-Zif268 (50−0.2 nM) (Figure 7). These experiments show that 200 pM (40 fmol) mCpG-2-Zif268 is visible above background and that activity scales linearly from 5 to 0.5 nM when the chromogenic substrate nitrocefin is used (Figure 7, inset).

To further increase the sensitivity of mCpG-SEER-β-Lac for dsDNA, the fluorogenic substrate CCF2FA was utilized to optimize the detection capability of mCpG-2-Zif268. CCF2FA is a FRET-based fluorescent β-lactamase substrate and has been previously reported to detect as little as 50 molecules of β-lactamase in a single living cell.[39] Starting with an amount of dsDNA near the lower linear limit of what can be detetcted with nitrocefin, experiments were performed as described above in the presence of decreasing concentrations of mCpG-2-Zif268 (500−25 pM) (Figure 8a). With the use of CCF2FA, mCpG-SEER-β-Lac was capable of detecting target dsDNA concentrations as low as 25 pM (2.5 fmol), resulting in a 16-fold improvement in sensitivity over nitrocefin. Additionally, activity was found to scale linearly from 500 to 25 pM (Figure 8b).

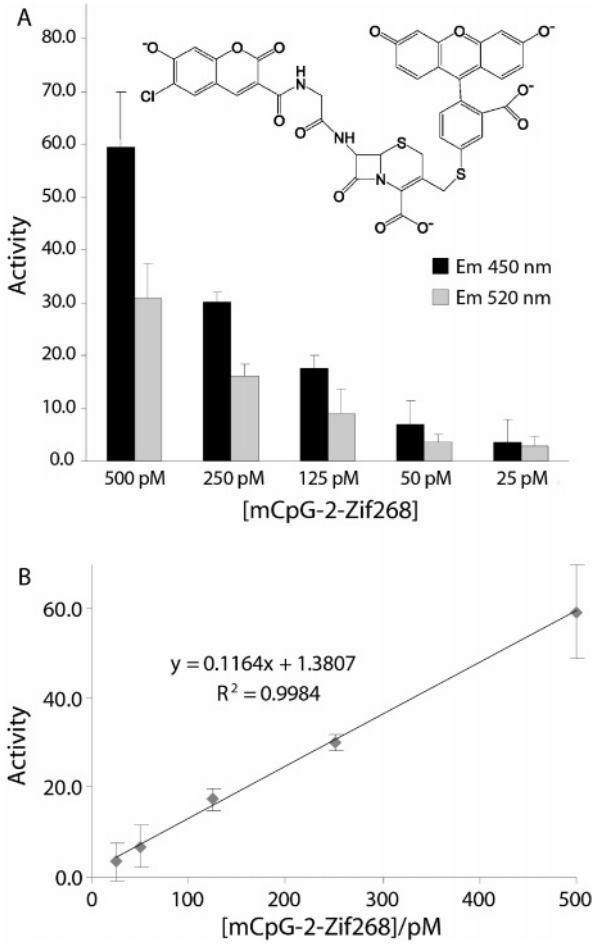

Figure 8. CCF2FA hydrolysis. (A) Hydrolysis of CCF2FA in the presence of MBP-LacA-Zif268 (125 nM) + MBP-MBD2-LacB (125 nM) in the presence of decreasing concentrations of the dsDNA target mCpG-2-Zif268. Hydrolysis was determined by either the increase in emission at 450 nm (black) or the decrease in emission at 520 nm (shown in gray as the negative decrease). (Inset) Structure of CCF2FA. (B) Linear fit of the emission at 450 from 500 to 25 pM mCpG-2-Zif268.

DISCUSSION

A large volume of work has been compiled demonstrating the ability to design and create split-protein biosensors capable of interrogating a wide variety of protein−protein and protein−nucleic acid interactions.[15,41,42] Additionally, a number of selection and design strategies have been utilized to develop proteins that can selectively bind specific DNA sequences.[34−36] Previously, the split GFP variant sg100 was utilized to develop a new methodology, termed mCpG-SEER, for the site-specific detection of DNA methylation. The current study builds upon this work to create mCpG-SEER-β-Lac through the employment of the split reporter protein TEM-1 β-lactamase. In this study, the expression vector pMAL-c2X was used, which appends a maltose binding protein domain to the N-terminus of the expressed recombinant proteins, resulting in fusion proteins with increased stability and solubility compared to the GFP-based system. The use of soluble mCpG-SEER-β-Lac proteins, coupled with the enzymatic signal amplification afforded by TEM-1 β-lactamase, allows for the detection of methylated dsDNA in less than 5 min when the chromogenic substrate nitrocefin was used. When the fluorogenic substrate CCF2FA was utilized, low-femtomole quantities of methylated target dsDNA could be detected, though longer assay times were

(41) MacDonald, M. L.; Lamerdin, J.; Owens, S.; Keon, B. H.; Bilter, G. K.; Shang, Z.; Huang, Z.; Yu, H.; Dias, J.; Minami, T.; Michnick, S. W.; Westwick, J. K. *Nat. Chem. Biol.* **2006**, *2*, 329−337.
(42) Schmidt, U.; Richter, K.; Berger, A. B.; Lichter, P. *J. Cell Biol.* **2006**, *172*, 373−381.

required due to a decreased rate of hydrolysis. This sensor thus provides a greater than 250-fold time improvement over the previously reported GFP-based mCpG-SEER system in terms of assay time. In addition to an increase in assay time, the currently described β-lactamase-based system also provides a 2000-fold improvement in the minimum amount of methylated target dsDNA needed to generate a detectable signal above background.

The crux of the mCpG-SEER methodology is the ability of the designed components to both distinguish methylated from nonmethylated CpG sites and accurately recognize promoter-specific DNA sequences. In the absence of cytosine methylation, a drastic 42-fold decrease in activity is seen, showing sensitivity to the methylation status of the CpG site of interest. This discrimination between methylated and nonmethylated DNA was further evaluated in the presence of a large excess of genomic herring sperm DNA (HS-DNA). A signal of >3-fold above background was observed for 15 ng of methylated DNA target in the presence of 1860 ng of HS-DNA (Supporting Information Figure S6). Additionally, a single bp mismatch in the adjacent zinc-finger binding site results in a 93-fold decrease in activity while the complete elimination of the adjacent zinc-finger binding site results in a 170-fold decrease in activity. These series of control experiments clearly demonstrate that the current mCpG-SEER-β-Lac design shows a remarkable ability to selectively detect methylated CpG sites in a sequence-dependent manner, attesting to the specificity of both MBD2 and zinc-finger DNA binding elements.

The distance-dependence profile of the mCpG-SEER-β-Lac system provides a number of insights into split β-lactamase reassembly and the design and application of the current mCpG-SEER-β-Lac system to promoter methylation analysis. Unlike the previously described GFP-based mCpG-SEER system (Figure 6), which shows minimal signal generation at a spacing of 8, 9, and 10 bps, the current mCpG-SEER system shows significant signal generation at a spacing of 8−13 bps. Unlike GFP, which refolds to adopt a single domain β-barrel fold, and is considered to be a single domain protein, β-lactamase comprises two separate domains, between which β-lactamase has been split. As a result, β-lactamase likely possesses a higher degree of conformational flexibility, allowing catalytic activity even under conditions that subtly distort the interface between these domains. This natural conformational flexibility likely accounts for the increased signal generation observed with β-lactamase over GFP at separations greater than 2 bp. The accurate analysis of methylated promoters will depend on the sequence specificity of the promoter-specific DNA binding domain; thus, the observed activity of mCpG-SEER β-Lac at different separation distances could potentially provide inaccurate results in the detection of site-specific methylation events. However, this flexibility may be a boon in disguise as it may allow one to potentially increase the amount of targetable sequence space. This is relevant since the tunability of zinc-fingers though extremely promising cannot at present target all DNA sequences with high sequence selectivity. Furthermore, the observed flexibility in β-lactamase reassembly will likely afford an avenue to tune the enzymatic activity by redesigning the 15-residue linkers that separate the lactamase halves from the DNA recognition domains.

Previous experiments have demonstrated the detection of as few as 50 molecules of β-lactamase using the fluorogenic substrate CCF2.[39] Using the free acid form of CCF2, our in vitro experiments clearly detect as little as 2.5 fmol of target methylated dsDNA. These results show that the use of β-lactamase provides a 2000-fold decrease in the minimal amount of dsDNA needed to generate a visible signal above background when compared to our split-GFP system. In the current study, fluorescence measurements were carried out using a standard top-read fluorescence plate reader, yielding a detection limit of 2.5 fmol of target dsDNA. Given the limitations of top-read plate readers, it is conceivable that lower detection limits will be achievable by future optimization of optical detection methods combined with reductions in sample volume. Moreover, modifications of the fluorescent lactamase substrate as well as re-engineering the lactamase can also potentially increase sensitivity. Future efforts will aim to extend this system for detecting DNA methylation by utilizing the split-luciferase system as well as target promoter sequences of genes known to be hypermethylated in cancerous cells with designed zinc-fingers.[34−36]

This report clearly demonstrates that split-β-lactamase provides a significantly improved split-reporter system domain for the site-specific detection of DNA methylation. With a rapid total assay time and a minimal detection of less than 3 fmol of methylated dsDNA, mCpG-SEER-β-Lac has the potential to provide a new strategy for determining the epigenetic modification of dsDNA. Given the currently available zinc-finger lexicon, new methods for determining sequence preferences of DNA binding proteins,[43,44] and the modularity of the SEER design, mCpG-SEER-β-Lac provides a new platform for directly interrogating the specific methylation status of promoters of interest. Thus, this methodology has the potential for providing new diagnostic reagents for understanding epigenetic control as well as aid in the early diagnosis of cancer.

ACKNOWLEDGMENT

This research was supported by R21CA122630 (NIH). J.R.P. and C.I.S were supported by an NIH training grant. We thank members of the Ghosh lab, Segal lab, Craig Aspinwall, and Kevin Braun for helpful comments. Models were rendered using PyMol; DeLano, W. L. (www.pymol.org).

SUPPORTING INFORMATION AVAILABLE

Details of cloning, sequencing, purification, and characterization in addition to a complete list of oligonucleotide targets used in this study. This material is available free of charge via the Internet at http://pubs.acs.org.

(43) Warren, C. L.; Kratochvil, N. C.; Hauschild, K. E.; Foister, S.; Brezinski, M. L.; Dervan, P. B.; Phillips, G. N., Jr.; Ansari, A. Z. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 867−872.

(44) Berger, M. F.; Philippakis, A. A.; Qureshi, A. M.; He, F. S.; Estep, P. W., 3rd; Bulyk, M. L. *Nat. Biotechnol.* **2006**, *24*, 1429−1435.

Received for review June 1, 2007. Accepted June 29, 2007.

AC071163+