Protein Expression and Purification 82 (2012) 332–338

Contents lists available at SciVerse ScienceDirect

Protein Expression and Purification

journal homepage: www.elsevier.com/locate/yprep

Functional heterologous expression and purification of a mammalian methyl-CpG binding domain in suitable yield for DNA methylation profiling assays

Mary E. Boyd, Brandon W. Heimer, Hadley D. Sikes *

Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, United States

ARTICLE INFO

Article history:
Received 2 November 2011
and in revised form 21 January 2012
Available online 2 February 2012

Keywords:
Methyl CpG binding proteins
Recombinant protein expression
Improved soluble yield
Epigenetics
DNA methylation profiling

ABSTRACT

DNA methylation is a major epigenetic modification in mammalian cells, and patterns involving methylation of cytosine bases, known as CpG methylation, have been implicated in the development of many types of cancer. Methyl binding domains (MBDs) excised from larger mammalian methyl-CpG-binding proteins specifically recognize methyl-cytosine bases of CpG dinucleotides in duplex DNA. Previous molecular diagnostic studies involving MBDs have employed *Escherichia coli* for protein expression with either low soluble yields or the use of time-consuming denaturation–renaturation purification procedures to improve yields. Efficient MBD-based diagnostics require expression and purification methods that maximize protein yield and minimize time and resource expenditure. This study is a systematic optimization analysis of MBD expression using both SDS–PAGE and microscopy and it provides a comparison of protein yield from published procedures to that from the conditions found to be optimal in these experiments. Protein binding activity and specificity were verified using a DNA electrophoretic mobility shift assay, and final protein yield was improved from the starting conditions by a factor of 65 with a simple, single-step purification.

Introduction

Methylation of cytosine bases in DNA is an epigenetic modification that has far-reaching and, as of yet, incompletely understood implications in developmental biology [1,2] and in pathology [3,4]. In mammals, a large body of work links particular methylation patterns in particular regions of the genome with several subtypes of cancer, including breast, colon, lung, glioma, leukemia, lymphoma, bladder, kidney, prostate, esophageal, stomach, liver, and ovarian cancers [5–9]. One of these patterns, termed CpG methylation, is described in Fig. 1 and occurs in the promoter regions of tissue-specific subsets of genes in malignant cells [8,10], whereas these dinucleotide duplexes in the gene promoters of normal cells are largely unmethylated [11,12]. Hypermethylation of CpG-rich regions of gene promoters, also known as CpG islands, has been demonstrated to effect the silencing of the downstream coding region; among other gene products, tumor suppressor proteins and DNA repair enzymes are expressed at lower levels as a result of this epigenetic modification [8].

Many analytical approaches for detecting methyl-cytosine in DNA have been developed [13–15], and several rely on specific binding interactions with proteins. Antibodies raised against 5-methylcytosine (5-meC) are one class of protein that binds single-stranded methylated DNA. Immunoprecipitation using 5-meC antibodies is a widely used method of enriching a DNA sample for methylated sequences [16–18]; however, this approach does not provide selectivity for CpG nucleotides or a particular CpG motif (hemi- vs. symmetric as defined in Fig. 1). Methyl-CpG-binding proteins [19] are an alternative set of binding molecules that recognize only methyl-cystosine bases in CpG dinucleotides of DNA in its native duplex form [20–22]. The methyl binding domains (MBDs)[1] [23] of larger methyl-CpG-binding proteins [19,24,25] have been demonstrated to discriminate between unmethylated CpGs, hemi-methylated CpGs, and symmetrically methylated CpGs on the basis of the dissociation constants of each of the above molecular complexes [26–28].

The potential utility of MBDs in molecular diagnostic applications has been suggested in recent studies [26,29,30]. In these proof-of-concept studies, MBD coding regions were excised from the larger mammalian (murine or human) gene, cloned into bacterial expression vectors, and expressed in *Escherichia coli*. Yields of purified proteins were not reported, and some groups used lengthy denaturation–renaturation purification procedures to recover the large fraction of expressed protein in inclusion bodies. For diagnostic applications of MBDs, both in the assay development phase and in the end-use phase, expression and purification procedures that

* Corresponding author. Fax: +1 617 253 2272.
E-mail address: sikes@mit.edu (H.D. Sikes).

[1] *Abbreviations used:* MBDs, methyl binding domains; MGMT, O6-methylguanine DNA methyl- transferase; CV, column volumes.

1046-5928/$ - see front matter © 2012 Elsevier Inc. All rights reserved.
doi:10.1016/j.pep.2012.01.016

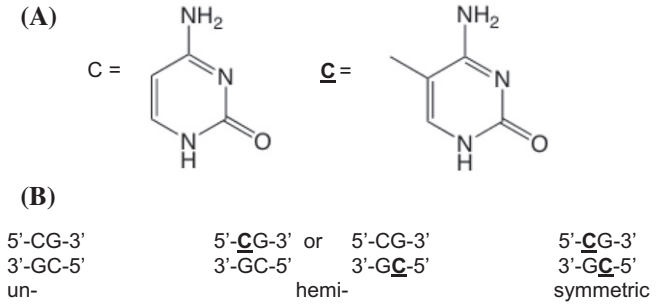

Fig. 1. (A) Cytosine, 5-methylcytosine, and (B) all possible CpG methylation patterns found in dsDNA. Underlining and bold text signify 5-methyl cytosine.

Table 1
DNA used in EMSA for MBD activity. Underlining and bold text indicate 5-methyl cytosine.

Methylation pattern	DNA sequence
o_o_o	5′-NH_2-CGGCAAACGGCATCAAACGGCTTTGCGGTCCGCTGCCCGACCC-3′
o_o_o	3′-AAACGCCAGGCGACGGGCTGGG-5′
o_m_o	5′-NH_2-CGGCAAACGGCATCAAACGGCTTTGCGGTC**C**GCTGCCCGACCC-3′
o_m_o	3′-AAACGCCAGG**C**GACGGGCTGGG-5′

result in maximal protein yields with minimal investment of resources are highly desirable. The present study reports yields for published procedures [26] and a series of systematic optimization experiments coupled with a simple, single-step affinity purification that resulted in final yields improved by a factor of 65 over the starting conditions.

Materials and methods

Bacterial strain and plasmids

A recombinant pET30(b+) plasmid containing the 1×MBD-GFP construct [26] was obtained from Steve Blair's lab (University of Utah, Salt Lake City); sequence data for this gene is provided in Supplementary Fig. 1. This fusion was created by inserting the GFP gene from the pUB-GFP plasmid (Addgene, Plasmid 11155) into the pET30(b+)-1×MBD plasmid [31] from Adrian Bird's lab (Edinburgh, UK). The sequence encoding the N-terminal His_6-tag and nuclear localization signal (NLS) preceding the cDNA of murine MBD1 (Entrez Gene 17190) amino acids 1–75 remained intact with downstream addition of the GFP gene. Amino acids 1–75 encode the methyl binding domain (MBD) of the larger MBD1 protein. *E. coli* BL21 (DE3) cells were transformed with the recombinant 1×MBD-GFP plasmid via electroporation.

Optimization of protein expression conditions using SDS–PAGE and fluorescence microscopy

Protein expression conditions were systematically studied with the goal of maximizing the expression level of soluble protein. For these studies, a 14 mL culture tube (17 × 100 mm, VWR) containing 5 mL of TB medium (12 g tryptone, 24 g yeast extract from BD Biosciences and 4 mL glycerol from BDH dissolved in 900 mL of distilled water and autoclaved, then cooled and brought to 1 L with the addition of 100 ml of a sterile solution of 0.17 M KH_2PO_4 and 0.72 M K_2HPO_4) supplemented with 50 μM kanamycin was inoculated with 500 μL of overnight culture and incubated at 37 °C with orbital shaking. When cultures reached an OD_{600} of 0.6, recombinant protein expression was induced at temperatures ranging from 20 to 37 °C and IPTG concentrations from 0.05 to 1 mM. Due to slower rates of cellular processes at lower temperatures, longer expression times (in addition to the 3 h time commonly reported in the literature for 37 °C) were investigated for lower temperatures.

In addition to induction of protein expression using IPTG, auto-induction [32] conditions were studied. For these experiments, cultures were inoculated according to the above protocol in Overnight Express™ Instant TB Medium (Novagen) supplemented with 50 μM kanamycin. Cultures were incubated at 37 °C with orbital shaking until reaching an OD_{600} of 0.2–0.3. At this point, cultures were either moved to 18 °C for 16–24 h or maintained at 37 °C for 8–24 h. Cells were harvested in 1 mL samples via centrifugation at 12,000*g* and 4 °C for 10 min. Both SDS–PAGE and fluorescence microscopy were used to compare the relative amounts of soluble and insoluble MBD-GFP protein produced with each set of conditions. For SDS–PAGE analysis, 1 mL cell pellets were lysed with 300 μL Bugbuster HT Protein Extraction Reagent (EMD4Biosciences) and treated according to the manufacturer's instructions. The soluble portion was subsequently separated from insoluble cell debris via centrifugation at 12,000*g* and 4 °C for 15 min. For every condition, approximately 0.25 μL of cell pellet (i.e. insoluble fraction) and 10 μL of supernatant (i.e. soluble fraction) were each diluted with 10 μL gel loading buffer. These sample amounts were chosen because this loading provides a clear resolution of bands in PAGE analysis; intensities in pellet lanes can be meaningfully compared with intensities in other pellet lanes and intensities in lysate lanes may be meaningfully compared with intensities in other lysate lanes. After heat denaturation (95 °C for 5 min, ice for 5 min), 20 μL of each sample and 15 μL of a MW standard (BioRad Precision Plus Protein™, Dual Color) were separated on a 4–15% SDS–PAGE gradient gel for 35 min at 150 V and visualized with Coomassie Brilliant Blue R250 stain. Culture volumes, lysis reagent volumes, and sample volumes for gel loading were each carefully controlled and held constant for all growth conditions.

In addition to SDS–PAGE analysis, light microscopy was used to visualize the degree of intracellular localization of MBD-GFP. Prior to microscope analysis, 1 mL cell pellets were resuspended and diluted 100-fold in Tris buffer (25 mM Tris–HCl, 150 mM NaCl, pH 7.4). Intact cells were examined with both bright field and fluorescence microscopy using an Olympus IX81 microscope equipped with a 60× oil-immersion objective. Fluorescence images of MBD-GFP were acquired using a filter cube comprising Semrock 472/30 excitation, 495 dichroic and 520/35 emission filters. During fluorescence imaging, the lamp intensity setting (10%, Prior Lumen 200) and the image acquisition time (5 ms) were each held constant for all samples. Binning was not used.

Protein purification

Larger-scale protein purification studies were carried out for growth conditions previously established in the literature and for the optimal conditions among those described above with the goal of quantitatively comparing yields for each set of conditions. Recombinant protein expression was induced using IPTG at an OD_{600} of 0.6 for two sets of growth conditions: 37 °C for 3 h with 1 mM IPTG in accordance with previous work [26] and 20 °C for 16 h with 0.05 mM IPTG in accordance with the results the optimization experiments described above. MBD-GFP expression in auto-induction medium was accomplished by allowing cultures to reach an OD_{600} of 0.2 at 37 °C and then moving the cultures to 18 °C for 24 h. Cultures expressed for 3 h at 37 °C were prepared in a 1 L flask containing 250 mL of TB medium supplemented with 50 μM kanamycin and inoculated with 5 mL of overnight culture. Because of the high level of expression at 20 °C for 16 h, purification