ONCOLOGY REPORTS 17: 111-116, 2007

The identification of a novel alternatively spliced form of the MBD4 DNA glycosylase

RHIANNON M. OWEN[1], RACHAEL D. BAKER[1], SCOTT BADER[2], MALCOLM G. DUNLOP[3] and IAIN D. NICHOLL[1]

[1]Research Institute in Healthcare Science, School of Applied Sciences, University of Wolverhampton, Wolverhampton; [2]Division of Pathology, Edinburgh Cancer Research Centre, University of Edinburgh; [3]Colon Cancer Genetics Group, University of Edinburgh, School of Molecular and Clinical Medicine and MRC Human Genetics Unit, Western General Hospital, Edinburgh, UK

Received May 22, 2006; Accepted August 16, 2006

Abstract. Methyl-CpG binding protein 4 (MBD4) is a mismatch-specific G:T and G:U DNA glycosylase. During an analysis of MBD4 expression in HeLa cells we noted the presence of an unexpectedly short reverse transcribed product. This cDNA lacked the region encoding the methyl-binding domain and exon 3 of MBD4 but retained the glycosylase domain. Sequence comparison indicates the existence of a previously unreported cryptic splice site in the MBD4 genomic sequence thus illuminating a mechanism whereby a glycosylase acquired a methyl-binding capacity, thus targeting potential mutagenic CpG sites. *In vitro* assays of this highly purified species, refolded in arginine rich conditions, confirmed that this unique, short version of MBD4 possessed uracil DNA glycosylase but not thymine DNA glycosylase activity. We conclude that the identification of a transcript encoding a short version of MBD4 indicates that MBD4 expression may be more complex than previously reported, and is worthy of further investigation.

Introduction

MBD4 is a 66-kDa mammalian methyl-CpG binding protein with DNA *N*-glycosylase activity for uracil, thymine, 5-fluorouracil (5-FU) and 3,N^4-ethenocytosine when mismatched with guanine (1,2). As hydrolytic deamination of 5-methylcytosine (me^{5}C) can result in the production of potentially mutagenic G:T mispairs (3), it has been proposed that the primary function of the methyl-binding domain (MBD) of MBD4 is to enhance the localisation of the DNA *N*-glycosylase activity to regions of the genome where 5-methylcytosine levels are greatest and therefore, where me^{5}C-deamination is most likely to take place (4).

Correspondence to: Dr Iain D. Nicholl, School of Applied Sciences, University of Wolverhampton, Wolverhampton WV1 1SB, UK
E-mail: i.nicholl@wlv.ac.uk

Key words: MBD4, DNA repair, alternative splicing

MBD4 is an unusual DNA glycosylase in that it can bind the DNA mismatch repair (MMR) protein MLH1 via its glycosylase domain (5). This has led to the suggestion that MBD4 is possibly involved in the coordination of base excision repair and MMR activities, when for example, G:T mismatches occur during DNA replication (6). Recently, MBD4 has been shown to have the capacity to bind the Fas-associated death domain protein (FADD) and may therefore participate in DNA damage-induced apoptotic signaling (7). Indeed, MBD4 deficiency has been shown to reduce the apoptotic response to DNA-damaging agents in the murine small intestine (8). A number of groups have additionally shown that MBD4 is frequently mutated in human colonic and extracolonic cancers (9-11). Thus, considerable evidence has accrued, based on *in vitro* and *in vivo* studies, to support the thesis that MBD4 participates in the maintenance of genomic fidelity. Herein we report the identification of a native, novel form of MBD4 lacking the methyl-binding domain but possessing DNA glycosylase activity.

Materials and methods

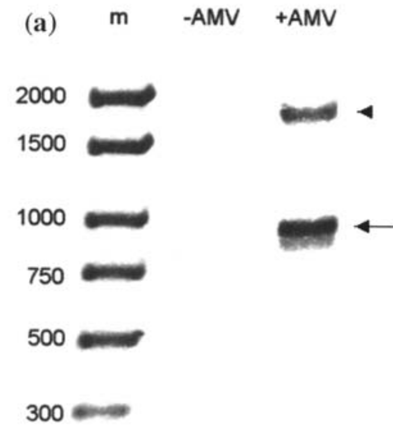
Dulbecco's modified Eagle's medium (DMEM) was from Gibco-BRL. Restriction enzymes were obtained from Roche. APE1 was from New England Biolabs (NEB). Oligonucleotides were purchased from Sigma-Genosys. Ammonium sulphate, imidazole and urea were from BDH. All other reagents were obtained from Sigma unless otherwise noted. HeLa cells were cultured in DMEM supplemented with 10% (v/v) foetal calf serum and glutamine.

For sMBD4 cloning, HeLa cells were cultured to 75% confluency. An approximately 950 bp DNA fragment was amplified by RT-PCR using p(dN)6 primers and a 1st strand synthesis cDNA AMV kit (Roche), with total RNA isolated from HeLa cells with a RNAqueous-4 PCR kit (Ambion, Austin, TX). The following oligonucleotides were used as primers: human MBD4, N-terminal, 5'-TTCAGGATCCGCA GCCGGACCCTGCTCG-3'; and C-terminal, 5'-TATACCCG GGTGAAAGCTGCAGAGTTTAAG-3' (12). RT-PCR was carried out with the following conditions: 1st strand synthesis: 1, 25˚C/10 min; 2, 42˚C/60 min; 3, 99˚C/5 min; 4, 4˚C/2 h;

PCR synthesis: 1, 95˚C/2 min; 2, 95˚C/1 min; 3, 55˚C/1 min; 4, 72˚C/2 min; 5, Goto 2 34 X; 6, 72˚C/7 min; 7, 32˚C/15 min; using an Expand HiFidelity kit (Roche). The short cDNA fragment was eluted using an UltraClean™ 15 kit (MO BIO Laboratories Inc.) from a 1.5% (w/v) agarose gel and digested with *Bam*HI and *Sma*I according to the manufacturer's instructions and ligated into the expression vector pQE30 (Qiagen) using T4 DNA ligase. *E. coli* clones containing His-expressing constructs were confirmed by immunoblot analysis using a polyclonal anti-His-tag antibody (Sigma). Plasmids were purified using a QiaPrep Miniprep kit (Qiagen) and sequenced using ABI BigDye terminator chemistry, employing pQE30 sequencing primers. Sequencing reactions were resolved on an ABI PRISM 3700 from PE Biosystems.

For the production and purification of recombinant 'short' MBD4 (sMBD4), *E. coli* strain SG13009[pREP] (Qiagen) transformed with the sMBD4-pQE30 vector was cultured in 1 l of Luria Broth medium at 37˚C to an OD_{600} value of ~0.2 and induced with 1 mM IPTG at 37˚C for approximately 2 h. Cells were lysed with 20 ml BugBuster (Novagen) supplemented with the protease inhibitor cocktail set 1 (Calbiochem) and EDTA to 2 mM, whereupon the sample was homogenised (Dounce, 5X) and incubated for 15 min at room temperature (r.t.). To the lysate 8 ml of denaturing buffer A (25 mM Tris. Cl pH 8.0, 8 M urea, 10 mM NaCl, 1 mM EDTA, 1 mM DTT) was added and the lysate homogenised (Dounce 10x), incubated at r.t. for 60 min and centrifuged at 13,000 rpm for 30 min at r.t.. The supernatant was dialysed against buffer A for 3 h at r.t. to complete protein denaturation. The sample was fractionated over a 1 ml HiTrap Q column (Amersham Biosciences), with significant levels of sMBD4 eluting in the flow through. This sMBD4-containing fraction was dialysed against buffer B (4X PBS, 8 M urea, 20 mM imidazole) and fractionated over a 1 ml HisTrap FF column (Amersham Biosciences). Bound His-tagged sMBD4 was eluted with buffer C (4X PBS, 8 M urea, 0.5 M imidazole). Fractions containing sMBD4 were dialysed against an arginine-rich refolding buffer, buffer D [0.2 M Tris.Cl pH 8.0, 10 mM EDTA, 0.6 M L-arginine HCl, 20% glycerol (v/v), 50 mM NaCl, 1 mM DTT] formulated to minimise insoluble protein aggregation [(13), Roche-Applied Science: http://www.roche-applied-science.com/sis/proteinexpression/literature/manual/chapter5/RTS_176_181.pdf] at 4˚C overnight. sMBD4 purity was confirmed by SDS-PAGE. Fractions containing highly purified sMBD4 were dialysed into 20 mM Tris.Cl pH 8.0, 20% (v/v) glycerol, 50 mM NaCl, 1 mM EDTA, 1 mM DTT (14) and stored at -20˚C prior to further analysis.

Samples of chromatographic fractions were precipitated (15) and analysed by SDS-PAGE using a 12% (w/v) resolving gel (acrylamide:bis ratio 37.5:1, Bio-Rad), with gels routinely stained with Coomassie Blue R250. Molecular masses of species were estimated in comparison to broad range protein standard (Bio-Rad) and fluorescently tagged (Sigma) markers. Immobilization onto nitrocellulose was performed by wet-transfer (Bio-Rad). Immunoblotting was carried out according to standard procedures (16). His-tagged sMBD4 migration was determined using an HRP conjugated monoclonal anti-polyhistidine (clone His-1) antibody (Sigma) and visualised either with DAB peroxidase substrate kit (Vector Labs) or by using ECL and a Storm 840 with ImageQuant software (Amersham Biosciences). Anti-MBD4 polyclonal antibodies, clones V-18 ('internal' epitope) and E-19 (against the C-terminus) were purchased from Santa Cruz Biotechnology.



(b)

PQE-30 RBS *His tag* *start*

atgagaggat cg[**catcacca tcaccatcac**] ggatccgcag ccggaccctg ctcg**ATG**ggc
acgactgggc tggagagtct gagtctgggg gaccgcggag ctgcccccac cgtcacctct
agtgagcgcc tagtcccaga cccgccgaat gacctccgca aagaagatgt tgctatggaa
ttggaaagag tgggagaaga tgaggaacaa atgatgataa aaagaagcag tgaatgtaat
cccttgctac aagaacccat cgcttctgct cagtttggtg ctactgcagg aacagaatgc
c*aagatacca tcccacgaac acagatagaa agaaggaaaa caagcctgta tttttccagc
aaatataaca aagaagctct tagcccccca cgacgtaaag cctttaagaa atggacacct
cctcggtcac cttttaatct cgttcaagaa acactttttc atgatccatg gaagcttctc
atcgctacta tatttctcaa tcggacctca ggcaaaatgg caatacctgt gctttggaag
tttctggaga agtatccttc agctgaggta gcaagaaccg cagactggag agatgtgtca
gaacttctta aacctcttgg tctctacgat cttcgggcaa aaaccattgt caagttctca
gatgaatacc tgacaaagca gtggaagtat ccaattgagc ttcatgggat tggtaaatat
ggcaacgact cttaccgaat tttttgtgtc aatgagtgga agcaggtgca ccctgaagac
cacaaattaa ataaatatca tgactggctt tgggaaaatc atgaaaaatt aagtctatct
taaactctgc agcttt cac ***STOP***

Figure 1. Identification of a short version of MBD4 in HeLa cells. (a), RT-PCR of MBD4 from HeLa mRNA. The unexpectedly short amplicon is arrowed. We presume the larger amplicon (arrowhead) represents full-length MBD4. Amplicons were only noted when AMV was used in the first strand synthesis. Markers (m) in base pairs are indicated. (b), His-tagged sMBD4 nucleotide sequence. The region of sequence change in comparison to full-length MBD4 is indicated with an asterisk. Deposited as EMBL accession no. AM180876.

For the DNA glycosylase assay enzymatic activity of purified sMBD4 was monitored by apurinic site nicking produced by DNA glycosylase activity on double stranded oligonucleotide substrates as previously described (17).

The oligonucleotides employed were: 5'-TAGACATTGC CCTCGAGGTACCATGGATCCGATGTCGACCTCAAAC CTAGACGAATTCCG-3'; 3'-ATCTGTAACGGGAGCTCC ATGGTACCTAGGCTACAGYTGGAGTTTGGATCTGCT TAAGGC-F-5', where F was fluoroscein and Y=C, T or U.

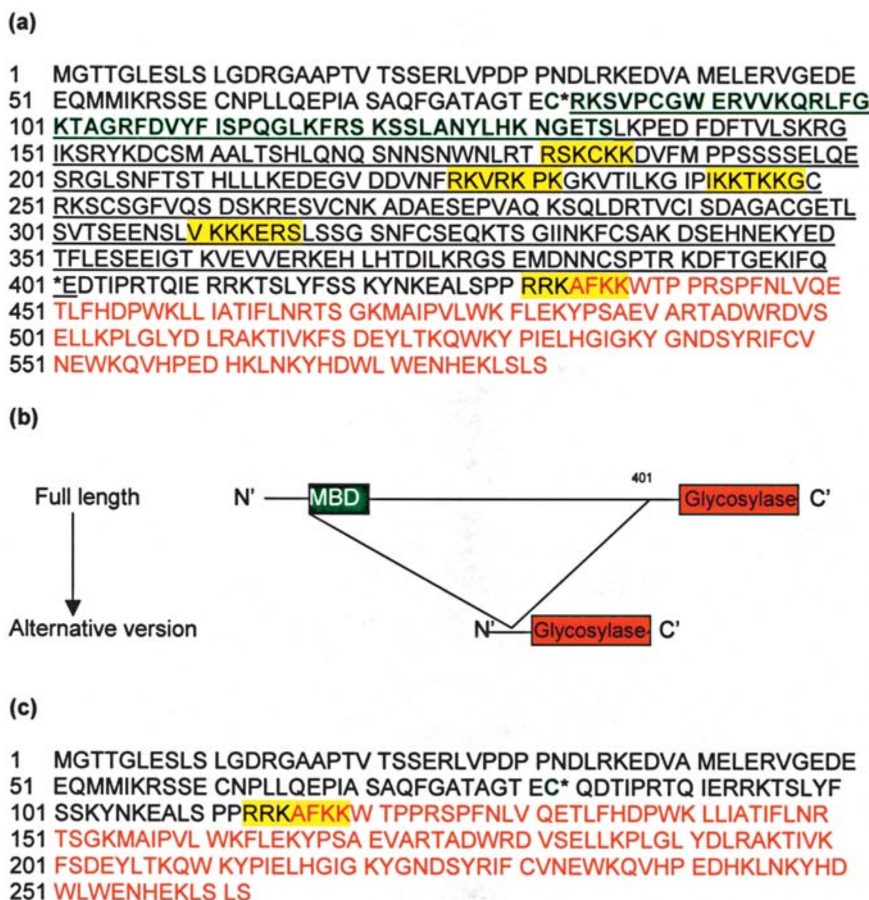


Figure 2. Shortened MBD4 retains the glycosylase domain but does not have an MBD. (a), Previously reported full-length MBD4 sequence. (b), Schematic diagram of MBD4 in comparison to predicted sMBD4 sequence. Splicing was seen to occur between the nucleotides coding for cysteine and arginine (codons 82/83) and glutamate (codon 401) residues, and results in the conversion of glutamate 401 to a glutamine in sMBD4. 319 amino acids are deleted (underlined) to form sMBD4. (c), Predicted sMBD4 sequence showing the potential nuclear localisation sequence and the glycosylase domain. Putative nuclear localisation signals are indicated (yellow) (5). The MBD (green) is denoted here to span codons 82-135 (see text). The glycosylase domain is highlighted in red (12,14).

The substrates were prepared by annealing oligonucleotides: a 0.5 M solution of the labelled lower strand and 1.0 M solution of the upper strand was heated at 95˚C in 10 mM Tris.Cl pH 8.0, 50 mM NaCl for 5 min. The solution was cooled to r.t. overnight. Assays were as follows: 50 nM labelled duplex DNA was incubated with approximately 140 nM protein in APE buffer (NEB), in the absence or presence of commercially available APE1 (NEB) for 60 min at 37˚C as indicated in Results.

Products were precipitated by adding 1/10 vol of 3 M sodium acetate pH 5.5, and 3 vol cold ethanol (-20˚C) and left overnight at -20˚C whereupon samples were centrifuged at 14,000 rpm for 60 min. The precipitates were rinsed with 80% ethanol (-20˚C) and air-dried. Samples were resuspended in loading buffer (15% TE pH 8.0, 85% deionized formamide), denatured at 95˚C for 5 min, cooled on ice and resolved on a 15% (w/v) denaturing (7 M urea) polyacrylamide gel (acrylamide-bis ratio of 29:1, National Diagnostics). Reaction products were visualised using a Storm 840 using ImageQuant software.

Results

During an analysis of MBD4 expression in the HeLa cell line by RT-PCR using previously reported primers (12), we noted the presence of two amplicons, the larger corresponding to a size commensurate with MBD4 (Fig. 1a). The short amplicon was cloned into a His-tag expression system and sequenced (Fig. 1b). Unexpectedly, the sequence (deposited as accession no. EMBL AM180876) was identical to full-length MBD4 (accession no. EMBL AF072250) but did not possess nucleotides 144-231 of exon 2 and exhibited loss of exon 3. This corresponds to the loss of the parent codons 83-401, resulting in a 262-amino acid species instead of the 580 amino acids for full-length MBD4 (Fig. 2). When compared to previously published reports of human MBD4 domain structure (4), this shortened version of MBD4 (denoted herein *sMBD4*) does not possess the poorly conserved central region of MBD4 and most significantly, lacks most or all of the methyl-binding domain (MBD). *Pfam* analysis (http://www.ensembl.org/Homo_sapiens/protview?peptide=ENSP00000249910&db=core) indicates that the MBD in MBD4 spans from codons 76-151 whilst others have reported the MBD spans codons 82-135 (18). Sequence alignment indicates that the MBD in MBD4 is most similar to that of the MBD in the gene implicated in Rett syndrome, MECP2 (18-20). However, the highly conserved C-terminal region with 95% homology between mouse and man from codons 401-580, which contains a glycosylase domain from codons 434-580 (4,12,14), was retained in sMBD4. A putative nuclear localisation sequence, cor-

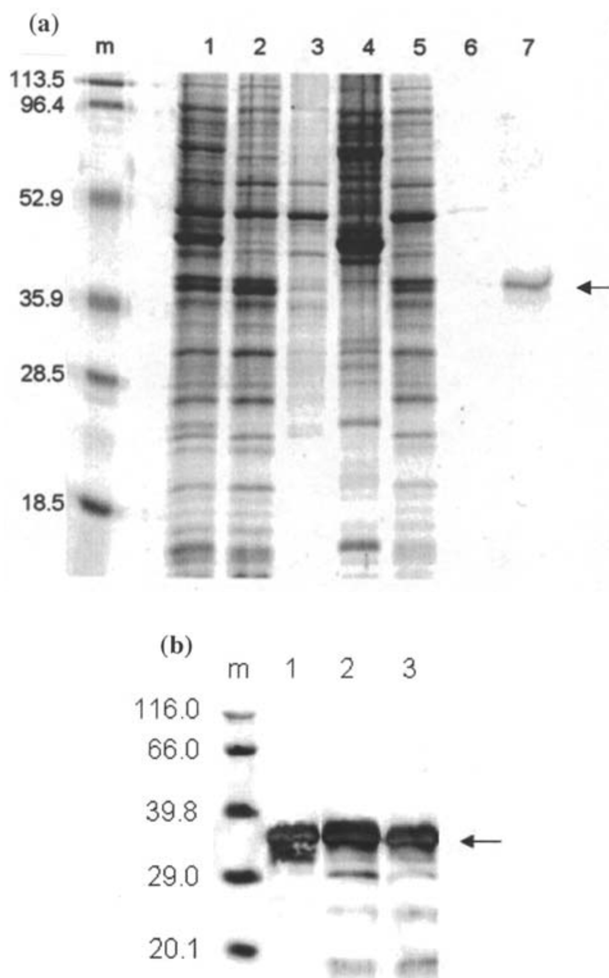


Figure 3. sMBD4 purification by denaturing anion-exchange and nickel-affinity chromatography. (a), SDS-PAGE of fractions. Coomassie blue stained. Lane 1, bacterial extract; lane 2, flow through from HiTrap Q column; lane 3, column wash; lane 4, eluted protein; lane 5, flow through from nickel-affinity column; lane 6, column wash; lane 7, eluted sMBD4, indicated by arrow. (b), Immunoblot analysis of sMBD4: lane 1, anti-His Tag antibody; lane 2, V-18 antibody; lane 3, E-19 antibody. Markers (m) in kDa in (a) and (b) are indicated.

responding to codons 431-437 of the wild-type human protein (5) is also present. The molecular mass of His-tagged sMBD4 was estimated as 32.2 kDa based on analysis of amino acid sequence.

To explore sMBD4 functionality it would be ideal to express and purify recombinant sMBD4 to homogeneity, and determine whether it retained DNA glycosylase activity. Initial attempts to purify the sMBD4 reproducibly under non-denaturing conditions were unsuccessful (data not shown), most likely due to proteolysis (IDN, unpublished data). We attempted to purify the protein under denaturing conditions, employing anion-exchange and Nickel affinity chromatography. Highly purified recombinant sMBD4 was obtained (Fig. 3a), and observed to migrate with a molecular mass of approximately 37 kDa. The recombinant protein was dialysed against an arginine-rich buffer more typically utilised in pulse-refolding experiments to facilitate protein renaturation (Materials and methods). Reactivity with commercially available polyclonal anti-MBD4 antibodies confirmed the origin of the protein as MBD4 (Fig. 3b).

This purified sMBD4 was tested in a glycosylase 'nicking' assay using 60-mer duplex oligonucleotides containing a G:U or G:T mismatch, as described in Materials and methods. Commercially available uracil DNA glycosylase (UDG) and AP endonuclease (APE1) was used as a positive control to produce the 23-mer product expected to arise upon exposure of the G:U containing substrate to the glycosylase. Robust glycosylase activity against the G:U containing oligonucleotide was reproducibly detected with sMBD4, whilst negligible activity was noted against the G:T containing oligonucleotide (Fig. 4).

Discussion

Herein we report the identification of a novel transcript in HeLa cells, predicted to encode a shortened version of the DNA glycosylase MBD4 that we have termed *sMBD4*. By sequence comparison with full-length MBD4, splicing to create the sMBD4 transcript occurs at regions of critical importance to MBD4 function [Fig. 2 (4)]. However, the cloned sMBD4 construct retained uracil DNA glycosylase activity (Fig. 4). We suggest it is unlikely that this splicing is artefactual. It is formally possible that a cryptic splice donor site in exon 2 is present in the coding sequence for MBD4 in the HeLa cell line at codon 83, although we are not aware of reports of a germline mutation at this position. However, it is notable that within exon 2 at nucleotide position (exonic) 248-253 the sequence GTAAGT is present: this sequence, or an A→T germline mutation therein resulting in GTATGT, can act as a

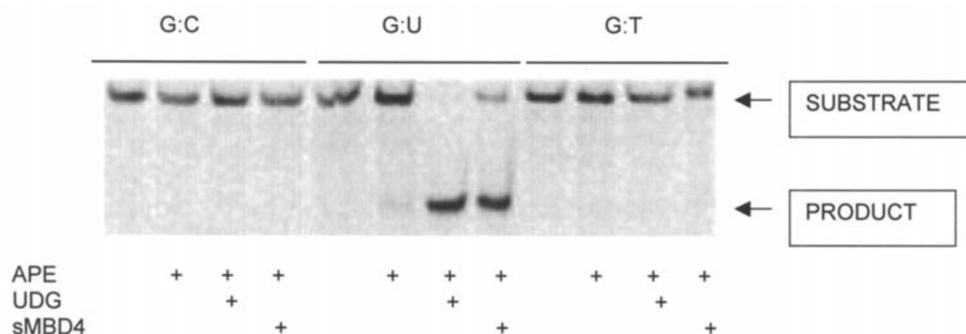


Figure 4. A G:U but not G:T mispair can act as a substrate for sMBD4. The oligonucleotide substrate containing the indicated mismatch was incubated with sMBD4, UDG (0.5 units) and/or APE1 (4 units) in a 'nicking assay'. The result shown is a fluorescence scan of a denaturing 15% acrylamide gel.

splice donor site (21). Whilst the existence of alternatively spliced forms of human and murine MBD4 transcripts have been reported before (http://us.expasy.org/cgi-bin/get-all-varsplic.pl?O95243), none of the forms had alterations which naturally affected coding of the MBD region (19).

A short version of MBD4 is not restricted to cells of human origin. The chick homolog of MBD4 does not possess an MBD, but retains G:T mismatch and 5-methylcytosine DNA glycosylase activity on hemimethylated DNA (12). We find merit in the suggestion that the full-length MBD4 could have arisen from the fusion of the glycosylase domain to an MBD, that results in a glycosylase that targets methylated mismatches (4).

It is worth noting that in previously published work where HeLa cell MBD4 has been analysed, a shorter version in addition to the major full-length species is clearly observable by immunoblot analysis (22). This could arise via proteolytic action during extraction or represent the existence of a truncated, or alternatively spliced version of MBD4.

With respect to the lack of thymine DNA glycosylase (TDG) activity present in purified sMBD4, whilst it is formally possible the assay conditions employed were not ideal for G:T repair activity, we note that one genetically engineered form of MBD4 lacking the MBD was found to lack TDG activity (4). Furthermore, it has been previously stated that TDG actually has a preference for G:U as a substrate over G:T mismatches (23); *in vitro* kinetic analysis also indicates that the MBD and methylation of the mismatched CpG substrate are not required for efficient catalysis by MBD4 (1).

The production of a shortened version of MBD4 may have particular significance to cancer biology. A human MBD4 construct lacking the MBD, but retaining the glycosylase domain and the intervening central region (codons 155-580) ('ΔMBD') was previously intimated to behave in a dominant negative fashion and interfere with DNA MMR (5). Thus paradoxically, whilst sMBD4 expression could result in an enzyme capable of effectively acting as a DNA glycosylase and participate in base excision repair, it may possibly act to negatively interfere with DNA MMR. However, HeLa cells are MMR proficient and it thus remains a formal possibility that the finding of microsatellite instability induced in a reporter gene in the SW480 colorectal cancer cell line by the MBD4 (ΔMBD) deletion mutant (5,6) cannot be generalised to a different cellular context.

Reduced MBD4 expression is correlated with poorer tumour differentiation in hepatocellular carcinoma (24), whilst increased expression of MBD4 is positively correlated with the grade of malignancy in gliomas (25). On the basis of our observation, we suggest that dysregulation of MBD4 function may also be a consequence of splicing dysregulation, producing protein species in tumour cells that could conceivably confer a selective growth advantage. Clearly, the control of the expression of full-length and production of alternatively spliced forms of MBD4 not only in cancerous cell lines but in normal tissue, merits further investigation.

Acknowledgements

We acknowledge the contribution of Paul Hooley, Lesley Stark, Ralf Zwacka, and colleagues of the MRC-Human Genetics Unit and the University of Wolverhampton. The work was partly funded by the University of Wolverhampton, the University of Edinburgh and through grants to MGD/IDN: Cancer Research UK Programme Grant (C348/A3758), Chief Scientist Office (K/OPR/2/2/D333).

References

1. Petronzelli F, Riccio A, Markham GD, *et al*: Biphasic kinetics of the human DNA repair protein MED1 (MBD4), a mismatch-specific DNA N-glycosylase. J Biol Chem 275: 32422-32429, 2000.
2. Meyers M, Wagner MW, Mazurek A, Schmutte C, Fishel R and Boothman DA: DNA mismatch repair-dependent response to fluoropyrimidine-generated damage. J Biol Chem 280: 5516-5526, 2005.
3. Scharer OD and Jiricny J: Recent progress in the biology, chemistry and structural biology of DNA glycosylases. Bioessays 23: 270-281, 2001.
4. Hendrich B, Hardeland U, Ng H-H, Jiricny J and Bird A: The thymine glycosylase MBD4 can bind to the product of deamination at methylated CpG sites. Nature 401: 301-304, 1999.
5. Bellacosa A, Cicchillitti L, Schepis F, *et al*: MED1, a novel human methyl-CpG-binding endonuclease, interacts with DNA mismatch repair protein MLH1. Proc Natl Acad Sci USA 96: 3969-3974, 1999.
6. Bellacosa A: Role of MED1 (MBD4) gene in DNA repair and human cancer. J Cell Physiol 187: 137-144, 2001.
7. Screaton RA, Kiessling S, Sansom OJ, *et al*: Fas-associated death domain protein interacts with methyl-CpG binding domain protein 4: a potential link between genome surveillance and apoptosis. Proc Natl Acad Sci USA 100: 5211-5216, 2003.
8. Sansom OJ, Zabkiewicz J, Bishop SM, Guy J, Bird A and Clarke AR: MBD4 deficiency reduces the apoptotic response to DNA-damaging agents in the murine small intestine. Oncogene 22: 7130-7136, 2003.
9. Bader S, Walker M, Hendrich B, Bird A, Bird C, Hooper M and Wyllie A: Somatic frameshift mutations in the MBD4 gene of sporadic colon cancers with mismatch repair deficiency. Oncogene 18: 8044-8047, 1999.
10. Riccio A, Aaltonen LA, Godwin AK, *et al*: The DNA repair gene MBD4 (MED1) is mutated in human carcinomas with microsatellite instability. Nat Genet 23: 266-268, 1999.
11. Yamada T, Koyama T, Ohwada S, *et al*: Frameshift mutations in the MBD4/MED1 gene in primary gastric cancer with high-frequency microsatellite instability. Cancer Lett 181: 115-120, 2002.
12. Zhu B, Zheng Y, Angliker H, *et al*: 5-Methylcytosine DNA glycosylase activity is also present in the human MBD4 (G/T mismatch glycosylase) and in a related avian sequence. Nucleic Acids Res 28: 4157-4165, 2000.
13. Vallejo LF and Rinas U: Strategies for the recovery of active proteins through refolding of bacterial inclusion body proteins. Microb Cell Fact 3: 1-12, 2004.
14. Wu P, Qiu C, Sohail A, Zhang X, Bhagwat AS and Cheng X: Mismatch repair in methylated DNA. Structure and activity of the mismatch-specific thymine glycosylase domain of methyl-CpG-binding protein MBD4. J Biol Chem 278: 5285-5291, 2003.
15. Wessel D and Flugge UI: A method for the quantitative recovery of protein in dilute solution in the presence of detergents and lipids. Anal Biochem 138: 141-143, 1984.
16. Harlow E and Lane D: Antibodies: A Laboratory Manual. Cold Spring Harbor Laboratory Press, NY, 1988.
17. Sartori AA, Schar P, Fitz-Gibbon S, Miller JH and Jiricny J: Biochemical characterization of uracil processing activities in the hyperthermophilic archaeon *Pyrobaculum aerophilum*. J Biol Chem 276: 29979-29986, 2001.
18. Roloff TC, Ropers HH and Nuber UA: Comparative study of methyl-CpG-binding domain proteins. BMC Genomics 4: 1-9, 2003.
19. Hendrich B and Bird A: Identification and characterization of a family of mammalian methyl-CpG binding proteins. Mol Cell Biol 18: 6538-6547, 1998.
20. Ballestar E and Wolffe AP: Methyl-CpG-binding proteins. Targeting specific gene repression. Eur J Biochem 268: 1-6, 2001.

21. O'Driscoll M, Ruiz-Perez VL, Woods CG, Jeggo PA and Goodship JA: A splicing mutation affecting expression of ataxia-telangiectasia and Rad3-related protein (ATR) results in Seckel syndrome. Nat Genet 33: 497-501, 2003.
22. Drummond JT and Bellacosa A: Human DNA mismatch repair *in vitro* operates independently of methylation status at CpG sites. Nucleic Acids Res 29: 2234-2243, 2001.
23. Yoon JH, Iwai S, O'Connor TR and Pfeifer GP: Human thymine DNA glycosylase (TDG) and methyl-CpG-binding protein 4 (MBD4) excise thymine glycol (Tg) from a Tg:G mispair. Nucleic Acids Res 31: 5399-5404, 2003.
24. Saito Y, Kanai Y, Sakamoto M, Saito H, Ishii H and Hirohashi S: Expression of mRNA for DNA methyltransferases and methyl-CpG-binding proteins and DNA methylation status on CpG islands and pericentromeric satellite regions during human hepatocarcinogenesis. Hepatology 33: 561-568, 2001.
25. Schlegel J, Guneysu S and Mennel H-D: Expression of the genes of methyl-binding domain proteins in human gliomas. Oncol Rep 9: 393-395, 2002.