BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS **232,** 10–13 (1997)
ARTICLE NO. RC976211

Mapping the Minimal Domain of hMSH-2 Sufficient for Binding Mismatched Oligonucleotides

Adrian Whitehouse,[1] Jayne Deeble,* Graham R. Taylor,* Pierre J. Guillou,† Simon E. V. Phillips,‡ David M. Meredith, and Alexander F. Markham

*Molecular Medicine Unit, Clinical Sciences Building, *DNA Laboratory, Regional Genetics Service, †Division of Surgery, St. James's University Hospital, and ‡Department of Biochemistry and Molecular Biology, University of Leeds, Leeds, United Kingdom*

Received January 15, 1997

The human MSH-2 gene product is a member of a highly conserved family of proteins involved in post-replication mismatch repair. Germline mutations in this gene have been implicated in hereditary non-polyposis colorectal cancer (HNPCC). Alterations in the coding region of the hMSH-2 gene result in a mutator phenotype with marked instability of microsatellite sequences, indicative of a deficiency in DNA repair. We have previously shown that a region of high homology between MutS proteins of different species containing a nucleotide binding domain, is sufficient to bind DNA containing specific mismatched residues. In order to determine the minimal domain of hMSH-2 necessary for binding mismatch-containing oligonucleotides, deletion analysis of the C-terminal region was performed. We have constructed a 5′ and 3′ deletion series, expressed each deletion as a bacterial fusion protein and assessed it for ATPase activity and its ability to identify mismatch containing DNA. Here we demonstrate that a 585 bp fragment encoding 195 amino acids within the C-terminal domain of hMSH-2 is sufficient to bind to DNA containing mismatches.

The two major inherited forms of colorectal cancer predisposition are hereditary non-polyposis colorectal cancer (HNPCC) and familial adenomatous polyposis (FAP). HNPCC is an autosomal dominant condition. It is believed that 1 in 500 of the western population may be heterozygotes for mutant alleles and predisposed to this form of cancer (2-3). It leads to the development of tumours of the colon, endometrium, stomach, small intestine, hepatobiliary system, kidney, ureter and ovary (4). HNPCC kindreds are characterised by a high incidence of instability in microsatellite sequences which are prone to somatic mutations with the appearance of larger and/or smaller alleles (4). Also, cell lines derived from HNPCC tumours are found to be genetically unstable (5). These observations were indicative of a failure in DNA repair. Subsequently, germlines mutations in HNPCC families were identified in four human genes, hMSH-2, hMLH1, hPMS-1 and hPMS-2, all of which are homologous to bacterial genes involved in mismatch repair (6-12).

In *Escherichia coli (E.coli),* mismatch recognition is mediated by a single protein MutS, which binds to base mispairs or loops of up to 4 base pairs (13). However, in *Saccharomyces cerevisiae,* MSH2, (The MutS homologue) functions with either MSH3 or MSH6 in mismatch repair (14). The recognition of mismatched DNA in humans is thought to be mediated by at least two MutS homologs, hMSH-2 and G/T binding protein (GTBP) forming the heterodimer, hMutSα (15-16). This complex is believed to bind to any mismatches initiating the repair of the heteroduplex. A third protein, hMSH-3 also forms a heterodimer with hMSH-2, hMutSβ, which has been shown to bind loops of up to 4 base pairs (17). The MutL homologs, hPMS2 and hMHL1, are then recruited to the complex and the mismatch is repaired. In bacteria, this repair process involves excision of the tract of single-stranded DNA which contains the mismatched residue, resynthesis of the excised DNA and finally religation (13).

hMSH-2 is of particular importance as mutations in this gene are thought to underlie approximately 40% of HNPCC cases (18). hMSH-2 has also been shown to bind to DNA containing mismatches *in vitro* (19-20). We have previously shown that the C-terminal domain of hMSH-2 displays ATPase activity and is sufficient to bind specific mismatched oligonucleotides (1). In this paper, we have constructed a 5′ and 3′ deletion series within the C-terminal domain of hMSH-2. These mutants were expressed as bacterial fusion proteins, as-

[1] Corresponding author. Fax: 044 113 2444475; Email: msjaw@stjames.leeds.ac.uk.

0006-291X/97 $25.00

sessed for ATPase activity and examined for the ability to bind DNA containing mismatched residues.

METHODS

Construction of a hMSH-2 C-terminal 5′ and 3′ deletion series. The full length cDNA of hMSH-2 is contained in the vector, pBShMSH-2 (provided by Prof. B. Vogelstein). Deletion mutants were constructed by digesting pBShMSH-2 with the restriction enzymes shown (Fig.1), blunt ending with T4 polymerase or Klenow DNA polymerase and ligated with pFlag.CTC (IBI), previously digested with *Hind*III and blunt-ended to derive, pAWΔ1-6, respectively.

Expression of hMSH-2 C-terminal deletion mutants as bacterial fusion proteins. Each deletion construct cloned into the bacterial expression vector pFlag.CTC, was used to transform *E. coli* strain DH5α. A fresh overnight culture of transformed *E. coli* was diluted 1 in 20 with LB medium containing ampicillin (100 μg/ml). After growth at 37°C for 2 hours, the culture was induced with IPTG (1 mM) and grown at 37°C for a further 5 hours. The cells were harvested by centrifugation at 3200g for 10 minutes, resuspended in 0.1 volume lysis buffer (100mM Tris, pH 8.0, 1mM EDTA) and incubated on ice with 3 mg/ml of lysozyme for 30 minutes. The cells were then sonicated and lysed by the addition of Tween 20 lysis buffer (100 mM Tris-HCl, pH 8.0, 200 mM NaCl, 1 mM EDTA, 0.3 mg/ml phenylmethylsulphonyl fluoride, 0.8 μg/ml pepstatin, 1 mM DTT, 1% Tween 20). Cellular debris was pelleted by centrifugation at 4,000g.

Detection of fusion proteins by Western blot analysis. Protein extracts of each deletion mutant were mixed with 2× reducing sample buffer (50mM Tris-HCl, pH 6.8, 4% sodium dodecyl sulphate (SDS), 5 mM EDTA, 10% β-mercapthoethanol, 1 mM DTT and 0.01% bromophenol blue). After boiling for 3 minutes, samples were fractionated on a 12% SDS polyacrylamide gel. After electrophoresis the gel was soaked for 10 min in transfer buffer (25 mM Tris, 192 mM glycine, 20% methanol (v:v), and 0.1% SDS), and the proteins were transferred to nitrocellulose membranes by electroblotting for 3 hours at 250 mA. After transfer, the membranes were soaked in PBS and incubated for 2 hr in blocking buffer (PBS containing 5% nonfat dry milk). Membranes were incubated with a 1/3000 dilution of the M2 monoclonal antibody (IgG_1, IBI), washed with PBS and incubated for 1 hr at 37°C with a 1/1000 dilution of rabbit anti-mouse immunoglobulin conjugated with horseradish peroxidase in blocking buffer. After five washes with PBS the nitrocellulose membranes were developed in PBS containing 0.02% 1-chloro-4-naphthol and 0.006% hydrogen peroxide.

ATPase assay. The assay was performed at 37°C in 20 mM Tris-HCl, pH 7.6, 0.5 mM $CaCl_2$, 5 mM $MgCl_2$, 1 mM DTT, 100 μg/ml BSA, 0.1 mM EDTA and 150 ng of hMSH-2 domain protein. Assays were performed using 2, 2.5, 3.3, 5 and 10 μM ATP. Hydrolysis of [α-^{32}P]ATP by the carboxy terminal domains of hMSH-2 was assayed by thin layer chromatography. The radioactive counts for ATP and its hydrolysis products were quantified using a scintillation counter (Packard).

Functional binding assay. Mismatch binding was detected by a nitrocellulose binding assay of labelled oligonucleotides followed by autoradiography as previously described (1). Briefly, oligonucleotides (dCGG ATC CGG ATG T**X**A TGG AAT TCC and dGGA ATT CCA T**X**A CAT CCG GAT CCG) were synthesised and annealed to produce either a perfect matched double-stranded molecule, or a single mismatch (position shown in bold type). Oligonucleotides were annealed and end-labelled using polynucleotide kinase and [γ-^{32}P]ATP. The binding assay used 1 pmole of DNA with 150 ng hMSH-2 domain in a total volume of 10 μl. After 1 hour on ice the mixture was slowly filtered over pure prewetted nitrocellulose (Millipore, 0.45μm) and washed in STM buffer. The filter was then allowed to air dry and bound DNA was detected by autoradiography.

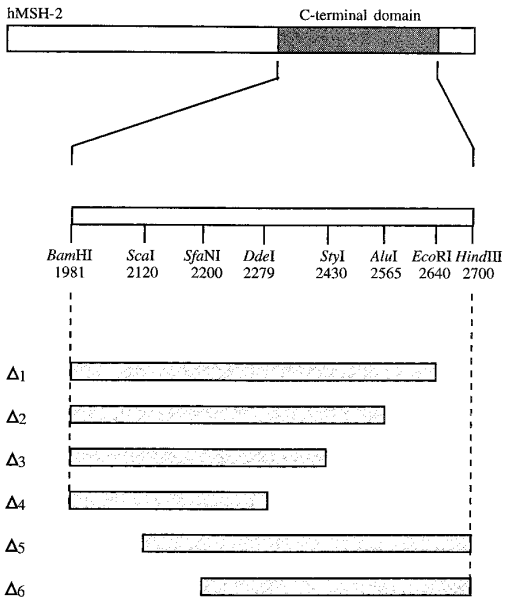

FIG. 1. Diagrammatic representation of the deletion series from the C-terminal domain of hMSH-2. A series of 5′ and 3′ mutants was constructed by digesting pBShMSH-2 with the restriction enzymes shown, blunt ending with T4 polymerase or Klenow enzyme and ligating with pFlag.CTC.

RESULTS AND DISCUSSION

Cloning of 5′ and 3′ deletion mutants. A DNA fragment encoding the C-terminal domain of hMSH-2 has been previously shown to be sufficient to bind mismatched oligonucleotides (1). To further analyse the sequences within this domain which are required for mismatch recognition a 5′ and 3′ deletion series within the C-terminal domain of hMSH-2 was produced (Fig. 1). Deletions within the C-terminal domain of hMSH-2 were constructed by digesting pBShMSH-2 with restriction enzymes which progressively delete larger fragments from the 5′ and 3′ regions of the gene. These fragments were ligated into the bacterial expression vector, pFlag.CTC, in phase with respect to the ATG translational start codon immediately upstream of the multiple cloning site (MCS) and also in frame with the C-terminal coding sequence immediately downstream of the MCS to ensure proper fusion to the C-terminal Flag peptide (Asp Tyr Lys Asp Asp Asp Asp Lys), to derive Δ1-6.

Expression of the hMSH-2 C-terminal deletion mutants. Expression of the deletion mutants resulted in a range of protein species detected by Western blot analysis on SDS-PAGE (Fig. 2). The anti-Flag M2 monoclonal (IgG1) mouse antibody (IBI) was used to specifically bind to the eight amino acid Flag peptide, which identified the deletion mutant fusion proteins pAWΔ1-6. This detected proteins of 27, 24, 19, 13, 24 and 21 kDa, respectively, comprising the hMSH-2 C-

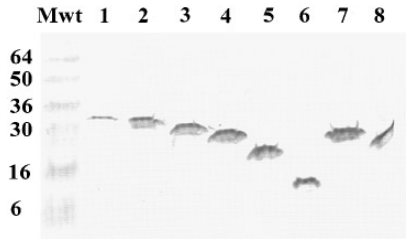

FIG. 2. Western blot analysis of mutant fusion proteins. Extracts were resolved by SDS PAGE, transferred to nitrocellulose, and incubated with M2 monoclonal antibody (IgG_1) and immune complexes were detected by using rabbit anti mouse immunoglobulin conjugated with horseradish peroxidase. The nitrocellulose membranes were developed in PBS containing 0.02% 1-chloro-4-naphthol and 0.006% hydrogen peroxide. Lane 1-2, hMSH-2 C-terminal domain-uninduced and induced; Lane 3, Δ1-induced; Lane 4, Δ2-induced; Lane 5, Δ3-induced; Lane 6, Δ4-induced; Lane 7, Δ5-induced; Lane 8, Δ6-induced.

terminal domain coupled to the Flag peptide at its carboxy terminus. As reported earlier (1), these domains only remain soluble in the presence of greater than 1 mM DTT, and as yet this protein has proved refractory to complete purification (1). Purification studies of the deletion mutant fusion proteins produced here were again complicated by insolubility problems. However, binding to mismatched oligonucleotides and ATPase activities, analogous to those previously observed of extracts from induced cultures of the carboxy terminal domains demonstrated the specificity of the enzymes as no activity was detected in uninduced cells. This suggests that these activities are generated from the constructs alone.

ATPase analysis of mutant fusion proteins. It has been shown that the carboxy terminal domain of hMSH-2 contains an ATPase activity (1). In order to determine whether these mutants hydrolyse ATP, to ADP and Pi, [α-^{32}P]ATP was incubated with each mutant fusion protein and the products separated using TLC. To determine K_m and k_{cat} values for the mutants, ATPase activity was measured in the presence of various concentrations of ATP (Fig. 3). The results show that the Δ1 and Δ2 deletions have limited effects on the ATPase activity of the domain. Wild type K_m and k_{cat} values were 7.96 μM and 0.5 s^{-1}, respectively. Δ3 and Δ4 display low ATPase activities. However, Δ5 and Δ6 had no detectable activity, probably due to the removal of the nucleotide binding domain. In control experiments, nonenzymatic hydrolysis of ATP in the absence of the wild type expressed domain was less than 5%.

Functional analysis of the mutant fusion proteins. A mismatch binding assay was developed to measure the hMSH-2 C-terminal domains' activity (1). Mismatch recognition was detected by nitrocellulose binding of labelled oligonucleotides containing a mismatch at position 11 within the context of a double-stranded 24-mer oligonucleotide pair. We found that the wild type C-terminal domain of hMSH-2 selectively bound all specific mismatches apart from A/A and C/C, in agreement with results described previously (1). The control vector did not bind to any labelled oligonucleotide pair. Applying this assay to the mutant proteins, we found that Δ1 and Δ2 bound the same specific mismatches as the wild type domain albeit to a slightly lesser extent. This may be due to the deletion of sequences leading to reduced recognition of the mismatches or a reduced affinity of these proteins once bound to a mismatch resulting in separation from the mismatch in the washing procedures of the assay. Δ3 and Δ4 were essentially unable to bind mismatches indicating that sequences downstream of 2430 bp (of the published cDNA sequence) are essential for mismatch binding. Δ5 and Δ6, which have the nucleotide binding domain deleted, are also unable to bind mismatches demonstrating that ATPase activity is essential for mismatch binding (Fig. 4).

We have previously reported that a large C-terminal domain of hMSH-2 exhibits ATPase activity and is sufficient to bind mismatched oligonucleotides (1). In this report we have produced and studied a range of deletion mutants to map function in the carboxy terminal domain. We have shown that the minimal region for mismatch binding is contained within a 585 bp fragment, which encodes a protein fragment of 195 amino acids. The minimal 585 base pair fragment is highly homologous between hMSH-2 and its homologues in other species and contains the type A nucleotide binding consensus sequence. The inability of Δ5 and Δ6, which have the nucleotide binding sequence removed, to bind mismatches is consistent with our hypothesis

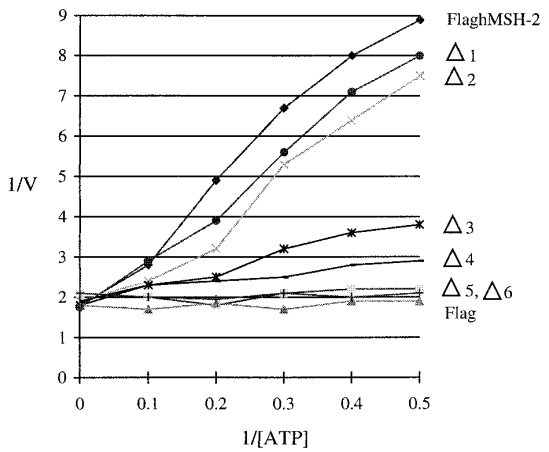

FIG. 3. ATPase analysis of the mutant bacterial fusion proteins. Hydrolysis of various substrate concentrations of [α-^{32}P]ATP by the carboxy terminal domain of hMSH-2, pFlag and Δ1-6 were assayed by thin layer chromatography, and quantified using a scintillation counter.

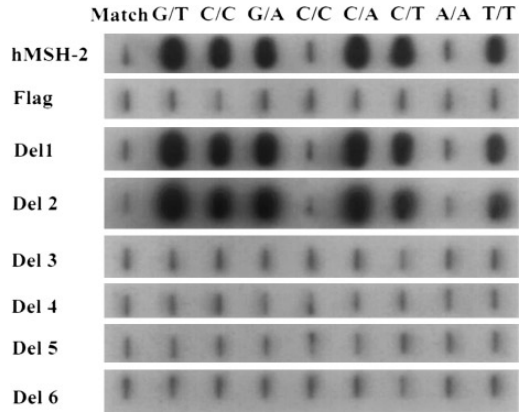

FIG. 4. Functional analysis of the mutant bacterial fusion proteins. Oligonucleotides containing either a perfect match or a range of single mismatches were radiolabelled using polynucleotide kinase. One pmole of labelled DNA was incubated for 1 hour with protein extracts of hMSH-2, pFlag or Δ1-6. After the incubation period the mixtures were slowly filtered over prewet nitrocellulose. Washed and bound DNA was detected using autoradiography.

that this ATPase capability is essential. Mutational analysis of this domain has previously demonstrated that a variant protein containing an alteration within the nucleotide binding consensus sequence of Lys675 to an alanine residue severely impairs mismatch binding (21), suggesting that this sequence plays a critical role in mismatch recognition.

It will be of interest to examine other regions of hMSH-2 in the future to identify other separate functions or specific interactions. Mismatch recognition is thought to be mediated by the heterodimers, hMutSα and hMutSβ (15-16). Deletion analysis of the N-terminal region of hMSH-2 may help in identifying specific domains necessary for the formation of these heterodimers by interaction with GTBP or hMSH-3, respectively.

ACKNOWLEDGMENTS

We thank the Yorkshire Cancer Research Campaign and West Riding Medical Research Trust for financial support for this project. Research in the authors' laboratory is also supported by the Medical Research Council, Northern and Yorkshire Regional Health Authority and the Wellcome Trust. SEVP is an International Research Scholar of the Howard Hughes Medical Institute. We are particularly grateful to Bert Vogelstein for originally providing the clone pBShMSH-2.

REFERENCES

1. Whitehouse, A., Deeble, J., Taylor, G. T., Phillips, S. E. V., Meredith, D. M., and Markham, A. F. (1996) *Biochem. Biophys. Res. Commun.* **225,** 289–295.
2. Dunlop, M. G. (1992) *Br. J. Surgery* **79,** 488.
3. Parker, S., Tong, T., Bolden, S., and Wingo, P. (1996). *CA Cancer J. Clin.* **46,** 5–27.
4. Lynch, H. T., Smyrk, T. C., Watson, P., Lanspa, S. J., Lynch, J. F., Cavalieri, R. J., and Bolland, C. R. (1993) *Gastroentrology* **104,** 1535–1549.
5. Loeb, L. A. (1994) *Cancer Res.* **54,** 5059–5063.
6. Parsons, R., Li, G. M., Longleym, M. J., Fang, W., Papadopoulos, M., Jen, J., de la Chapelle, A., Kinzler, K. W., Vogelstein, B., and Modrich, P. (1993) *Cell* **75,** 1227–1235.
7. Leach, F. S., Nicolaides, N. C., Papadopoulos, N., Liu, B., Jen, J., Parsons, R., Peltomaki, P., Sistonen, P., Aaltonen, L. A., Nystrom-Lahti, M., Guan, X.-Y., Zhang, L., Meltzer, P. S., Yu, J.-W., Kao, F.-T., Chen, D. J., Cerosaletti, K. M., Fournier, R. E. K., Todd, S., Lewis, T., Leach, R. J., Naylor, S. L., Weissenbach, J., Mecklin, J.-P., Jarvinen, H., Petersen, G. M., Hamilton, S. R., Green, J., Jass, J., Watson, P., Lynch, H. T., Trent, J. M., de la Chapelle, A., Kinzler, K. W., and Vogelstein, B. (1993) *Cell* **75,** 1215–1225.
8. Bronner, C. E., Baker, S. M., Morrison, P. T., Warren, G., Smith, L. G., Lescoe, M. K., Kane, M., Earabino, C., Lipford, J., Lindbolm, A., Tannergard, P., Bollag, R. J., Godwin, A. R., Ward, D. C., Norden-skjold, M., Fishel, R., Kolodner, R., and Liskay, R. M. (1994) *Nature* **368,** 258–261.
9. Kolodner, R. D., Hall, N. R., Lipford, J., Kane, M. F., Rao, M. R. S., Morrison, P., Wirth, L., Finan, P. J., Burn, J., Chapman, P., Earabino, C., Merchant, E., and Bishop, D. T. (1994) *Genomics* **24,** 516–526.
10. Kolodner, R. D., Hall, N. R., Lipford, J., Kane, M. F., Morrison, P., Wirth, L., Finan, P. J., Burn, J., Chapman, P., Earabino, C., Merchant, E., and Bishop, D. T. (1994) *Cancer Res.* **55,** 242–248.
11. Liu, B., Nicolaides, N. C., Markowitz, S. S., Willson, J. K. V., Parsons, R. E., Jen, J., Papadopoulos, N., Peltomaki, P., de la Chapelle, A., Hamilton, S. R., Kinzler, K. W., and Vogelstein, B. (1995) *Nature Genetics* **9,** 48–55.
12. Papadopoulos, N., Nicolaides, N. C., Wei, Y.-F., Ruben, S. M., Carter, K. C., Rosen, C. A., Haseltine, W. A., Fleischmann, R. D., Fraser, C. M., Adams, M. D., Venter, J. C., Hamilton, S. R., Petersen, G. M., Watson, P., Lynch, H. T., Peltomaki, P., Mecklin, J.-P., de la Chapelle, A., Kinzler, K. W., and Vogelstein, B. (1994) *Science* **263,** 1625–1629.
13. Modrich, P. (1991) *Annu. Rev. Genet.* **25,** 229–253.
14. Johnsen, R. E., Kovvali, G. K., Prakash, L., and Prakash, S. (1996). *J. Biol. Chem.* **271,** 7285–7288.
15. Drummond, J. T., Li, G.-M., Longley, M. J., and Modrich, P. (1995) *Science* **268,** 1909–1912.
16. Palambo, F., Gallinari, P., Iaccarino, I., Lettieri, T., Hughes, M., D'Arrigo, A., Truong, O. J., Hsuan, J., and Jiricny, J. (1995) *Science* **268,** 1912–1914.
17. Palambo, F., Iaccarino, I., Nakajima, E., Ikejima, M., Shimada, T., and Jiricny, J. (1996) *Current Biology* **6,** 1181–1184.
18. Lui, B., Parsons, R., Papadopoulos, N., Nicolaides, N. C., Lynch, H. T., Watson, P., Jass, J. R., Dunlop, M., Wyllie, A., Peltomaki, P., de la Chapelle, A., Hamilton, S. R., Vogelstein, B., and Kinzler, K. W. (1996) *Nature Medicine* **2,** 169–174.
19. Fishel, R., Ewel, A., Lee, S., and Lescoe, M. K. (1994) *Cancer Res.* **54,** 5539–5542.
20. Fishel, R., Ewel, A., Lee, S., Lescoe, M. K., and Griffith, J. (1994) *Science* **266,** 1403–1405.
21. Whitehouse, A., Parmar, R., Deeble, J., Taylor, G. R., Phillips, S. E. V., Meredith, D. M., and Markham, A. F. *Biochem. Biophys. Res. Commun.* **229,** 147–153.