

FOKK2 transcription factor is a novel G/T-mismatch DNA binding protein

Received October 30, 2009; accepted January 5, 2010; published online January 22, 2010

Yoshito Fujii* and Michio Nakamura†

Department of Host-Defense Biochemistry, Institute of Tropical Medicine, Nagasaki University, Nagasaki 852-8523, Japan

*Yoshito Fujii, Department of Host-Defense Biochemistry, Institute of Tropical Medicine, Nagasaki University, Nagasaki 852-8523, Japan, Tel.: +81 95 819 7850, Fax: +81 95 819 7851, email: yfujii@nagasaki-u.ac.jp

†Present address: Section of Internal Medicine, Shounan Hospital, Soh, Kagoshima, 899-8106, Japan

DNA mismatch repair is an important mechanism in the prevention of mutations. We reported the existence in the HL60 cell line of a novel G/T-mismatch DNA-binding protein (nGTBP) requiring strict DNA sequences. In this report, we identify the FOKK2 transcription factor as the nGTBP. FOKK2 fragments were obtained as the only clones with specifically binding activity to the G/T-mismatch DNA by screening of a human brain expression library. The recombinant forkhead domain of FOKK2 specifically recognized the G/T-mismatch DNA. The forkhead domain also recognized hypoxanthine/T and G/uracil, derived from the deamination of the exocyclic amino groups of A/T and G/C, respectively. An electrophoretic mobility shift assay (EMSA) analysis using HL60 cell nuclear extract and antibody raised against FOKK2 resulted in the exclusive binding of FOKK2 to G/T-mismatch DNA. Furthermore, FOKK2 bound to G/T-mismatch DNA with higher affinity than 'match' FOKK2 consensus DNA. We therefore propose that FOKK2 is a G/T-mismatch DNA-binding protein and a deaminated DNA-binding protein.

Keywords: FOKK2/forkhead domain/G/T-mismatch DNA/deaminated base/expression screening.

Abbreviations: EMSA, electrophoretic mobility shift assay; FH, forkhead; Hx, hypoxanthine; FOX, forkhead box; ILF, interleukin enhancer binding factor; IPTG, isopropyl β -D-1-thiogalactopyranoside; mC, 5-methylated cytosine; nGTBP, novel G/T-mismatch DNA binding protein; oG, 8-oxo-guanine; Tg, thymine glycol.

DNA base–base mismatches are caused by DNA polymerase errors or base damages by oxidation, alkylation and deamination (1, 2). If not repaired, these errors or damages can cause gene dysfunctions resulting in altered cellular phenotypes and/or diseases. Numerous mechanisms have evolved to repair these types of base modification causing mutations, including base excision repair (BER) and mismatch repair

(MMR) (1, 3). To initiate the mechanism, mismatches have to be specifically recognized by a sensor protein. For example, in humans, MutS α , a heterodimer of MSH2 and MSH6, recognizes base–base mismatches and a short mismatched insertion/deletion loop (IDL). In addition, thymine DNA glycosylase (TDG) recognizes G/T-mismatch DNA (1, 3).

We have reported the existence of a novel G/T-mismatch specific binding protein (nGTBP) in the human cell line (4). This nGTBP was found to be different from the reported G/T-mismatch DNA-binding proteins and to bind to a minimal 14-mer DNA heteroduplex containing a G/T-mismatch with a strict 5'-TRTGNB-3' sequence (where R stands for purine, N stands for any base, B stands for 'not A' and the underlined G stands for the G/T position). In this study, we identify FOKK2 (alias ILF-1) as the nGTBP. FOKK2 is classified as a forkhead box (FOX) transcription factor, which is defined by a common DNA-binding domain that is referred to as the forkhead (FH) domain.

Materials and Methods

Electrophoretic mobility shift assay

HL60 cell nuclear extraction and electrophoretic mobility shift assay (EMSA) were performed as described previously with minor modifications (5). Briefly, the binding reactions were performed on ice in 20- μ l aliquots. Recombinant protein was allowed to bind to a probe in the EMSA buffer (20 mM HEPES-KOH, pH 7.9, 1 mM DTT, 0.01% Triton X-100, 5% glycerol). In the case of competition assay using nuclear extract, 10 μ g of nuclear extract was pre-incubated on ice for 3 min in EMSA buffer containing the competitor DNA, 50 ng poly(dI)poly(dC) and Complete protease inhibitors (Roche). For the supershift assay, various amounts of goat polyclonal FOKK2 antibody (Novus Biologicals, NB100-1285) or goat polyclonal FOXO3a antibody (Santa Cruz Biotechnology, sc-34897X) was incubated on ice for 3 h with 10 μ g of nuclear extract in EMSA buffer containing 150 ng poly(dI)poly(dC), 200 ng salmon sperm DNA and Complete protease inhibitors. To make each probe, the upper strand oligonucleotides were labelled with T4 DNA ligase (New England Biolabs) and annealed with a 100-fold molar excess of cold counterpart oligonucleotide. The probe was added to the mixture and incubated on ice for 15 min except in Fig. 3. Electrophoresis was performed on a 5% polyacrylamide gel in a cold room (4°C). Images were analysed with a Molecular Imager FX (Bio-Rad). The apparent K_d was determined from the protein concentration at which half of the duplex DNA was protein-bound (6).

Expression screening

All of the DNA sequences used for screening are shown in Fig. 1A. Southwestern screening was performed by using lambda Uni-ZAP XR premade library established from human brain cells (Cerebellum) (Stratagene). Agar plates were plated with 9×10^5 pfu of phage. The isopropyl β -D-1-thiogalactopyranoside (IPTG)-saturated nitrocellulose filter overlays were incubated overnight at 37°C. The filters were treated for 15 min with 6 M guanidine/HCl in EMSA buffer and then re-natured at 4°C with six sequential 1:1 dilutions of EMSA buffer, each lasting 5 min. Then the filters were

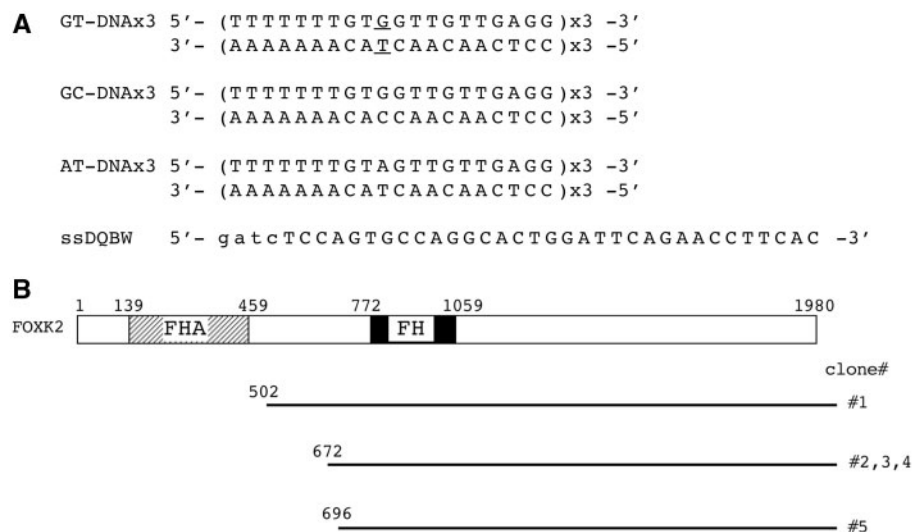


Fig. 1 Expression screening result. (A) DNA sequences used in the screening. The mismatched nucleotide G is underlined. The ssDQBW sequence was used to suppress non-specific YB-1 binding. (B) FOXX2 structure and the obtained clones. The numbers on the diagram are nucleotide positions from A of 1st methionine codon. FHA, forkhead-associated domain; FH, forkhead domain.

blocked with 5% of non-fat dry milk in EMSA buffer for 1 h at 4°C. After this the filters were pre-treated for 2 h at 4°C with EMSA buffer containing 100 mM KCl, 20 µg/ml of salmon sperm DNA, 20 pmol/ml of AT-DNAx3 and 30 pmol/ml of ssDQBW DNA. Protein–DNA binding was performed by adding ³²P-labelled G/T tandem triplicate probe (GT-DNAx3), with a final concentration of 2 × 10⁵ cpm/ml, to the pre-treatment buffer described above for 1 h at 4°C. The second and third selections were performed in EMSA buffer containing 100 mM KCl and 20 µg/ml of salmon sperm DNA with either a GT-DNAx3 probe or a GC-DNAx3 probe for use as a clone. The pBluescript phagemid of GT-DNAx3 specific clones were excised from the Uni-ZAP XR vector using SOLR cells and EXAssist helper phage. The resultant clones were sequenced with a 3730 DNA Analyzer (Applied Biosystems).

Plasmid construction

The FH domain of FOXX2 was cloned with the polymerase chain reaction (PCR) from obtained phagemid clone using PrimeSTAR HS DNA polymerase with GC buffer (TAKARA). The following primers were used: forward, 5'-CGGGATCCGAAGCCGCTTAC TCCTACGC-3'; and reverse, 5'-CCAAGCTTACCTAGGCCGTC GTTTCCTAAA-3'. The PCR product was digested with BamHI and HindIII and directionally subcloned into a pET29b vector (Novagen). The sequence was subsequently confirmed. As the purity of the resulting purified protein from the pET52 system was better than from the pET29 system, the product of the pET52b vector (constructed by subcloning with BamHI and NotI digestion) was used in the experimental trials.

Recombinant protein preparation

The FOXX2 FH domain was expressed in Rosetta-gamiB(DE3) plysS cells (Novagen) in the presence of 50 µg/ml carbenicillin, 34 µg/ml chloramphenicol, 15 µg/ml kanamycin and 12.5 µg/ml tetracycline. Cultures were grown to a density of 0.6 A₆₆₀. Then, IPTG was added to a final concentration of 1.0 mM. This solution was induced for 3 h at 37°C and the cells were then separated from suspension using centrifugation. The resulting cell pellets were resuspended in BugBuster protein extraction reagent (Novagen) containing 25 unit/ml of Benzonase nuclease and Complete protease inhibitors. Protein solubilization and nucleotide degradation was performed in a rotating tube at room temperature for 20 min. The insoluble material was removed by centrifugation at 15,000g at 4°C for 10 min. Purification of StrepTagII fusion protein was carried out using StrepTactin purification kits (Novagen) according to the manufacturer's instructions. Affinity-purified protein was treated with 50 unit/ml benzonase nuclease in dialysis buffer (20 mM HEPES-NaOH pH 7.1, 150 mM NaCl, 20 mM MgCl₂, 1 mM DTT) containing 4 M urea at room temperature for 3 h to remove

bound bacterial nucleic acid. Protein was renatured by subjected to stepwise dialysis in dialysis buffer and affinity-purified again to remove the nuclease. No nucleic acid was detected in the purified protein, as revealed by ethidium bromide staining. Purified protein was analysed using Tricine-SDS-PAGE on a Novex precast gel (Invitrogen). The protein concentration was determined by the Bradford method (Bio-Rad) using γ-globulin as a standard.

Results and Discussion

Identification of FOXX2 as nGTBP

In an attempt to identify the G/T-mismatch DNA-binding protein, we performed expression screening on the HL60 lambda cDNA library using GT-DNAx3 and GC-DNAx3 (Fig. 1A) as probes. We could not isolate any G/T-mismatch DNA specific clone and instead isolated YB-1, which was determined to be a binding protein to both GT-DNAx3 and GC-DNAx3. The human brain library was used for further screening with a buffer containing ssDQBW (Fig. 1A), which has been reported to be a single-stranded DNA that strongly binds to YB-1 (7), to block the binding. Five GT–DNAx3 specific clones, #1 to #5, were isolated after three rounds of selection using GT-DNAx3 and GC-DNAx3 probes (Fig. 1B). All clones contained FOXX2 (GenBank accession no.NM_004514) cDNA. These clones lacked the N-terminal coding region with the FH-associated (FHA) domain (139–459 bp, from A of the 1st methionine codon) but contain the FH domain (772–1059 bp) (Fig. 1B).

Binding specificity of FOXX2 FH domain

As the FH domain was reported as a DNA-binding domain, we examined the DNA-binding specificity of the bacterially expressed FOXX2 FH domain. The affinity-purified domain, containing no bacterial nucleic acid, was used for the experiment (Fig. 2A). In the EMSA assay, when we systematically introduced all 16 possible pairs into the X position of the probe, the FOXX2 FH domain displayed strong

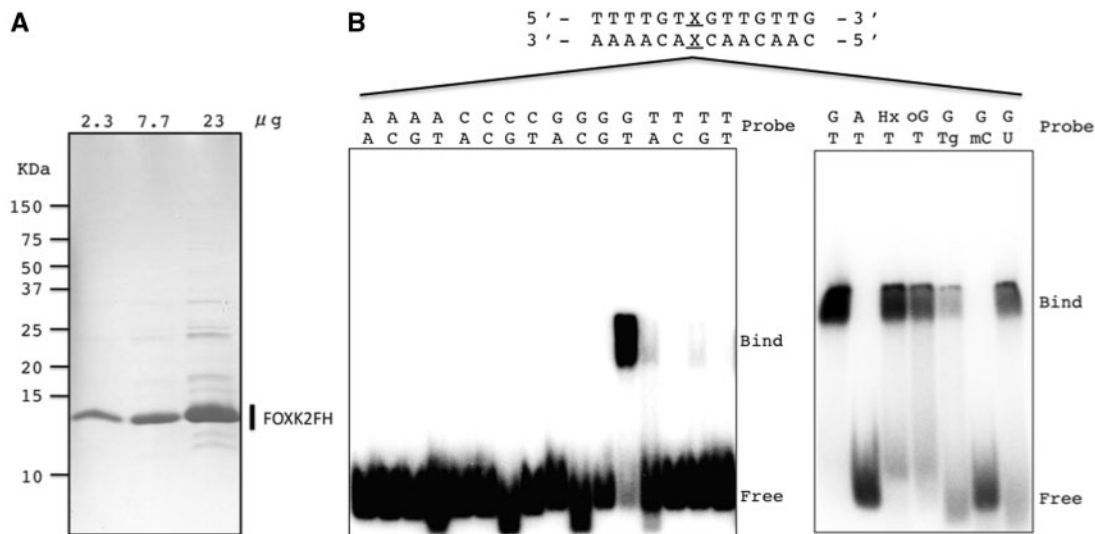


Fig. 2 Binding specificity of FOXK2 FH domain to ^{32}P -labelled DNA probes. (A) Tricine-SDS-PAGE analysis of purified recombinant StrepTagII-tagged FOXK2 FH domain by staining with Coomassie blue. The position of the molecular weight markers is indicated on the left. kDa, kilo-dalton. (B) EMSA analysis using the 14-mer probe with various base pairs at position X. The binding reaction was performed on ice for 15 min with 1×10^4 cpm probe and 1 pmol FOXK2 FH domain. Left: all possible 16 pair probes. Right: modified base probes. oG, 8-oxo-guanine; mC, 5-methyl cytosine; U, uracil; Hx, hypoxanthine; Tg, thymine glycol.

binding to G/T-mismatch DNA and slight binding to T/A-match DNA and T/G-mismatch DNA (Fig. 2B, left). We also examined the influence of the neighbouring sequences around G/T-mismatch DNA as previously reported (4). The results indicated that 5'-TRTGNB-3' sequence was required for FOXK2 FH domain binding to G/T-mismatch DNA (data not shown). These highly selective binding of the FOXK2 FH domain to the G/T-mismatch DNA probe was in good accordance with our previous nuclear extract data (4).

FOXK2 FH domain binds to deaminated derivatives from match DNA

In order to obtain some clues, which would imply a biological role for mismatch DNA recognition by FOXK2, we examined FOXK2 FH domain binding to probes that had modified bases introduced into position X (Fig. 2B, right), as previously reported with HL60 cell nuclear extract (4). The modified bases had deaminated C and A [uracil (U) and hypoxanthine (Hx)], oxidized G and T [8-oxo-guanine (oG) and thymine glycol (Tg)], and 5-methylated cytosine (mC). Hx/T and G/U probes displayed comparable binding with the G/T probe, suggesting that FOXK2 could be recruited for use to deaminated A/T and G/C. G/mC and G/C (Fig. 2B, left) displayed no binding, in contrast to G/U and G/T, and this suggested that the C-4 oxo group of pyrimidine contributed to probe recognition and the C-5 methyl group did not. A/T displaying no binding, in contrast to G/T and Hx/T, and this suggested that the C-6 oxo group of purine contributed to probe recognition and the C-2 amino group did not. oG/T displayed significant binding and this suggested that the C-8 oxo group of purine did not have any effect on G/T probe recognition. G/Tg displayed weak binding and this suggested that the

oxidation of pyrimidine had a partial inhibitory effect on G/T probe recognition.

Recombinant FOXK2 FH domain binds to GT-mismatch DNA with higher affinity than consensus DNA

FOXK2 has been reported as a binding protein to 'match' consensus DNA and to be involved in transcriptional regulation (8–10). To evaluate the affinity of FOXK2, we compared the binding affinity of various amounts of recombinant FOXK2 FH domain towards the G/T probe and the consensus ILF probe (8) (Fig. 3). Our results indicated that the FOXK2 FH domain bound to the G/T probe with a higher affinity than the ILF probe. The apparent K_d values of the G/T probe and ILF probe were ~ 30 and 230 nM, respectively. The K_d value towards the G/T probe was two orders of magnitude higher than those we reported previously (0.11 nM) for nGTBP in HL60 cell nuclear extract (4). The differences between the two experimental procedures, such as recombinant protein fragment vs nuclear extract may have affected the K_d values.

nGTBP is FOXK2 and binds to G/T-mismatch DNA with higher affinity than FOXK2 consensus DNA

Antibodies were subjected to EMSA assay using HL60 cell nuclear extract and a G/T probe (Fig. 4A). A major shift in the nGTBP band was observed (4). The nGTBP band was almost completely disrupted by $3 \mu\text{g}$ of FOXK2 antibody in a $20\text{-}\mu\text{l}$ reaction. In contrast, the FOXO3A antibody and control goat IgG (data not shown) retained the binding capacity. Therefore, we concluded that nGTBP is FOXK2. To evaluate the binding affinity of nGTBP, a competition assay was performed with various DNA (Fig. 4B). G/T and ILF competitors effectively inhibited the binding, in contrast to A/T and ILFmt competitors. This result

indicates that nGTBP binds specifically to both G/T-mismatch DNA and FOXK2 consensus DNA. Furthermore, the G/T competitor inhibited the binding more effectively than the ILF competitor ($P < 0.05$, in all doses). This indicates that nGTBP binds to G/T-mismatch DNA with higher affinity than to consensus DNA. These data were consistent with the FOXK2 FH domain data (Fig. 3).

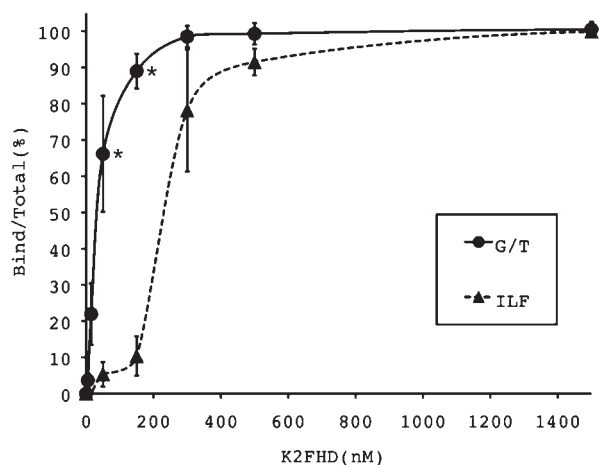


Fig. 3 Comparison of the binding affinity of the recombinant FOXK2 FH domain toward G/T probe with that toward ILF probe. Various amounts of FOXK2 FH domain were incubated on ice for 30 min in the presence of G/T or ILF probe (200 pM) and analysed using EMSA and binding fractions were quantified by means of phosphor imaging. G/T probe: 5'-TTTGTGGTGTG-3' (with the G/T position underlined). ILF probe: 5'-TGTTGTAAACAATACA-3'. The fraction of DNA bound was plotted vs the protein concentration, and the K_d was determined to be the concentration of protein at which 50% of the DNA was bound. Asterisks indicate where the differences between the G/T probe affinity and the ILF probe affinity are statistically significant ($P < 0.05$).

In summary, these results indicated that human FOXK2 specifically recognizes G/T-mismatch DNA via FH domain with higher affinity than 'match' consensus DNA. Furthermore, deaminated DNA can also be recognized by FOXK2. While the biological meanings of the results are unclear, the high specificity and high affinity of FOXK2 to mismatched DNA strongly suggest that FOXK2 initializes DNA repair mechanisms as a sensor for such mismatches.

FOXK2 (ILF1) was reported as the potential activator of interleukin-2 (IL-2) gene transcription by binding to the FOX consensus sequence (10). Some proteins have been shown to be involved in both transcriptional regulation and DNA repair. For example, YB-1, identified as a non-specific binding protein in our screening experiment, has been reported to bind to 'match' inverted CCAAT box, mismatch DNA and cisplatin-modified DNA (11, 12). YB-1 has a strand separation activity (11) and a 3' to 5' exonuclease activity (13). It interacts with proliferating cell nuclear antigen (PCNA), a component of several DNA repair systems (14), via the C-terminal region of YB-1 (12). We tested the interaction of FOXK2 and PCNA by immunoprecipitation using an antibody against C-terminal region of FOXK2 and found no significant binding of the two (data not shown). The C-terminal region of FOXK2, the target of the antibody, does not contain any PCNA-interacting protein (PIP) box sequences (14). We also tested the interaction of the two by EMSA using a PCNA antibody and also found no significant binding (data not shown). These results suggest that FOXK2 does not make a complex with PCNA. YB-1 has been reported to interact with other DNA repair proteins such as oxidized base-specific DNA glycosylases (NTH1 and NEIL2), DNA ligase III α , DNA polymerases (β and δ), MSH2, Ku80 and

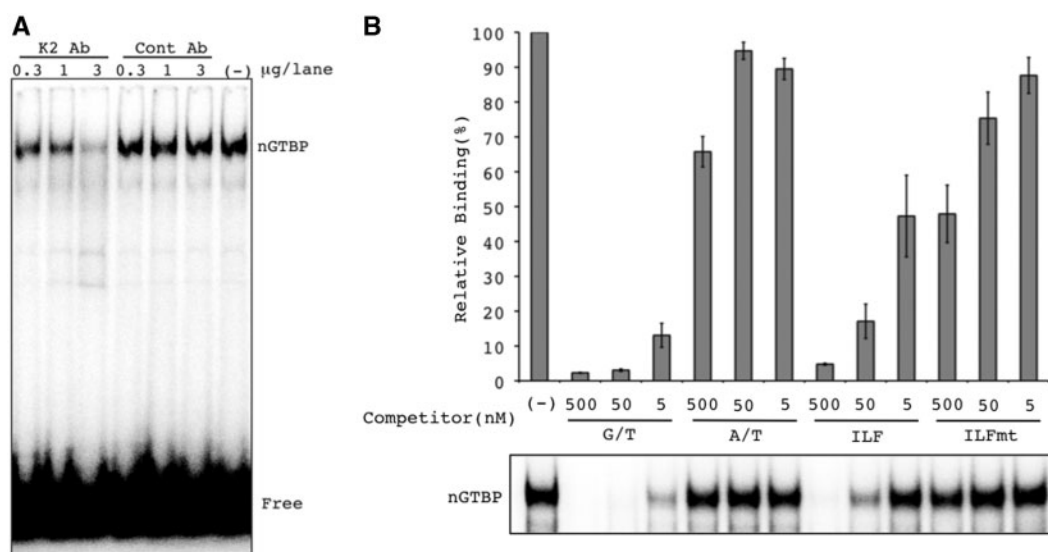


Fig. 4 EMSA analysis of HL60 cell nuclear extracts. (A) Antibody Supershift assays with FOXK2 antibody. HL60 nuclear extract was incubated on ice for 3 h in various amounts of FOXK2 antibody (goat) or FOXO3a antibody (goat) as a control before adding the G/T probe (1×10^4 cpm). (B) The HL60 nuclear extracts were incubated on ice for 30 min in the presence of various amounts of cold competitor DNA before adding the G/T probe (1×10^4 cpm; ~ 500 pM). The 14-mer G/T and A/T DNA sequences are shown in Fig. 2B. The ILF probe sequence is shown in Fig. 3. The ILFmt probe: 5'-TGTTGTAAAAAATACA-3' (with the mutated position underlined). Shifted bands were quantified by means of phosphor imaging. The radioactivity associated with nGTBP binding in the absence of competition was taken as the control value and defined as 100%. A typical gel image is displayed at the bottom.

WRN (11, 15, 16). Because FOXK2 has no catalytic domain usually found in DNA repair proteins (e.g. glycosylase, nuclease and strand separation), it may make a complex with a DNA repair protein such as YB-1 having a catalytic DNA repair activity. The FHA domain found in FOXK2 has been reported to be a small protein module that recognizes phosphothreonine epitopes, and involved in a DNA damage checkpoint response (17). Thus, FOXK2 might be engaged in DNA repair functions through activating the signalling pathway of its FHA domain.

Since the FH domain is the conserved DNA binding domain of FOX family proteins (18), a possible role in mismatch DNA repair system might be observed in the FOX family proteins in common even if small variations in mismatched DNA sequences might be required by the individual family proteins. An actual interaction of FOXK2 (and other FOX family proteins) with known and/or unknown members of DNA repair systems and biological roles of its mismatch-binding activity are remaining for future elucidation. Finally, it should be noted that the strong binding of FOXK2 to certain DNA mismatch sequences might freeze further transcription and replication of this mismatch containing gene and genome, respectively, which might result in cell death.

Acknowledgements

We thank Dr Tomoyuki Maekawa for valuable discussions and suggestions. We also thank Ms. Kanae Tanaka for technical assistance.

Funding

This work was supported by Grants-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology, Japan; grants for the Center of Excellence and Collaboration Research from the Institute of Tropical Medicine, Nagasaki University; and president's discretionary fund of Nagasaki University.

Conflict of interest

None declared.

References

- Kunz, C., Saito, Y., and Schär, P. (2009) Mismatched repair: variations on a theme. *Cell. Mol. Life Sci.* **66**, 1021–1038
- Yonekura, S., Nakamura, N., Yonei, S., and Zhang-Akiyama, Q. (2009) Generation, biological consequences and repair mechanisms of cytosine deamination in DNA. *J. Radiat. Res.* **50**, 19–26
- Dalhus, B., Laerdahl, J.K., Backe, P.H., and Bjørås, M. (2009) DNA base repair - recognition and initiation of catalysis. *FEMS Microbiol. Rev.* **33**, 1044–1078
- Takata-Yahiro, M., Fujii, Y., Nodarse, J.F., Islam, M.R., Oda, S., Zhang, Q., Yonei, S., and Nakamura, M. (2003) A novel GT-mismatch binding protein that recognizes strict DNA sequences with high affinity. *Tohoku J. Exp. Med.* **200**, 211–229
- Fujii, Y., Kumatori, A., and Nakamura, M. (2003) SATB1 makes a complex with p300 and represses gp91^{phox} promoter activity. *Microbiol. Immunol.* **47**, 803–811
- Carey, J. (1991) *Gel Retardation in Methods in Enzymology* (Sauer, R.T., ed.), Vol. 208, pp. 103–117, Academic Press, San Diego
- Hasegawa, S.L., Doetsch, P.W., Hamilton, K.K., Martin, A.M., Okenquist, S.A., Lenz, J., and Boss, J.M. (1991) DNA binding properties of YB-1 and dbpA: binding to double-stranded, single-stranded, and abasic site containing DNAs. *Nucleic Acids Res.* **19**, 4915–4920
- Tsai, K.L., Huang, C.Y., Chang, C.H., Sun, Y.J., Chuang, W.J., and Hsiao, C.D. (2006) Crystal structure of the human FOXK1a-DNA complex and its implications on the diverse binding specificity of winged helix/forkhead proteins. *J. Biol. Chem.* **281**, 17400–17409
- Li, C., Lai, C., Sigman, D.S., and Gaynor, R.B. (1991) Cloning of a cellular factor, interleukin binding factor, that binds to NFAT-like motifs in the human immunodeficiency virus long terminal repeat. *Proc. Natl Acad. Sci. USA* **88**, 7739–7743
- Nirula, A., Moore, D.J., and Gaynor, R.B. (1997) Constitutive binding of the transcription factor interleukin-2 (IL-2) enhancer binding factor to the IL-2 promoter. *J. Biol. Chem.* **272**, 7736–7745
- Gaudreault, I., Guay, D., and Lebel, M. (2004) YB-1 promotes strand separation in vitro of duplex DNA containing either mispaired bases or cisplatin modifications, exhibits endonucleolytic activities and binds several DNA repair proteins. *Nucleic Acids Res.* **32**, 316–327
- Ise, T., Nagatani, G., Imamura, T., Kato, K., Takano, H., Nomoto, M., Izumi, H., Ohmori, H., Okamoto, T., Ohga, T., Uchiumi, T., Kuwano, M., and Kohno, K. (1999) Transcription factor Y-box binding protein 1 binds preferentially to cisplatin-modified DNA and interacts with proliferating cell nuclear antigen. *Cancer Res.* **59**, 342–346
- Izumi, H., Imamura, T., Nagatani, G., Ise, T., Murakami, T., Uramoto, H., Torigoe, T., Ishiguchi, H., Yoshida, Y., Nomoto, M., Okamoto, T., Uchiumi, T., Kuwano, M., Funai, K., and Kohno, K. (2001) Y box-binding protein-1 binds preferentially to single-stranded nucleic acids and exhibits 3'→5' exonuclease activity. *Nucleic Acids Res.* **29**, 1200–1207
- Moldovan, G., Pfander, B., and Jentsch, S. (2007) PCNA, the maestro of the replication fork. *Cell* **129**, 665–679
- Marenstein, D.R., Ocampo, M.T.A., Chan, M.K., Altamirano, A., Basu, A.K., Boorstein, R.J., Cunningham, R.P., and Teebor, G.W. (2001) Stimulation of human endonuclease III by Y box-binding protein 1 (DNA-binding protein B). *J. Biol. Chem.* **276**, 21242–21249
- Das, S., Chattopadhyay, R., Bhakat, K.K., Boldogh, I., Kohno, K., Prasad, R., Wilson, S.H., and Hazra, T.K. (2007) Stimulation of NEIL2-mediated oxidized base excision repair via YB-1 interaction during oxidative stress. *J. Biol. Chem.* **282**, 28474–28484
- Durocher, D., and Jackson, S. P. (2002) The FHA domain. *FEBS Lett.* **513**, 58–66
- Katoh, M., and Katoh, M. (2004) Human FOX gene family. *Int. J. Oncol.* **25**, 1495–1500