# Methylation at CpG islands in intron 1 of *EGR2* confers enhancer-like activity

Motoko Unoki, Yusuke Nakamura*

*Laboratory of Molecular Medicine, Human Genome Center, Institute of Medical Science, The University of Tokyo, 4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639, Japan*

Received 30 July 2003; revised 20 September 2003; accepted 20 September 2003

First published online 7 October 2003

Edited by Horst Feldmann

**Abstract** **We previously demonstrated several lines of evidence indicating that early growth response 2 (EGR2) functions as a tumor suppressor, partly on the basis that its expression was often decreased in human tumors and cancer cell lines. Here we report a possible molecular mechanism to account for down-regulation of *EGR2* in tumor cells. Although no genetic mutations in the gene or alterations in methylation status of its promoter were detected, we found a high degree of methylation at CpG islands in intron 1 of *EGR2* in cell lines that were expressing this gene at a high level. Moreover, reporter gene experiments revealed that methylated intron 1 had somehow conferred enhancer-like activity. The data imply the existence of a previously unsuspected mechanism of gene expression regulation.**



*Key words:* EGR2; DNA methylation; CpG island; Enhancer; Transcriptional regulation

## 1. Introduction

Recently we published evidence that early growth response 2 (EGR2) functions as a tumor-suppressive mediator in the PTEN signaling pathway [1,2]. Our previous reports included findings that expression of endogenous *EGR2* was significantly low in primary ovarian cancers [2] as well as in various cancer cell lines [3].

The*Egr2*/*Krox-20* gene was originally identified as a serum response immediate-early gene [4]. Since Egr2 directly regulates *HoxB2* expression during hindbrain segmentation, Egr2(−/−) mice display disruption of the segmentation and die during the first 2 weeks after birth, and differentiation of Schwann cells in peripheral nervous system is also blocked at an early stage [5–7]. In humans, defects of this gene are responsible for several hereditary neuronal diseases [8].

Here we report a molecular mechanism that might be involved in the reduced expression of *EGR2* that has been observed in cancer cell lines derived from a variety of tissues. Since we were unable to detect any *EGR2* mutations in cells expressing low levels of this gene, we examined its methylation status in 16 representative cancer cell lines and found the obvious tendency that intron 1 was hypermethylated in cells that expressed high levels of *EGR2*, but hypomethylated in lines showing decreased expression. When we examined transcriptional activity of the respective unmethylated or methylated DNA segments, we observed enhancer-like activity in the latter. Our data suggest a novel explanation for the regulation mechanism of *EGR2*.

## 2. Materials and methods

### 2.1. Cell lines and normal tissue samples

Ishikawa3-H-12 was obtained from Tsukuba University (Tsukuba, Japan). Eight cell lines of the HEC series were obtained from Kitasato University (Sagamihara, Japan). AN3CA, KLE, MDAH2774, OV-1063, SW626, NIH:OVCAR-3, SK-OV-3, SW480, LoVo, A172, A549, H1299, LS174T, HT-29, DBTRG-05MG, U373MG, U87MG, HCT116, HepG2, Huh7, Alexander, LNCap.FGC, and PC-3 were obtained from the American Type Culture Collection (Manassas, VA, USA). We obtained Sawano, HHUA, and HOUA-I from the RIKEN Gene Bank (Tsukuba, Japan) and TYK-nu, LU99A, and MKN74 from the Japanese Collection of Research Bioresources (Tokyo, Japan). SNU423 and SNU475 came from the Korean cell line bank. All cell lines used here were maintained as described previously [3]. We obtained mRNA derived from normal tissue samples from BD Bioscience Clontech (Palo Alto, CA, USA).

### 2.2. Real-time quantitative polymerase chain reaction (PCR)

Expression of *EGR2* was analyzed quantitatively by TaqMan-PCR assays [3]. Primer sequences were 5′-TCTTTCCCAATGCCGAACTG-3′ and 5′-GGAGATCCAACGACCTCTTCTCT-3′; the probe was 5′-TTGATCATGCCATCTCCGGCCACT-3′. The PCR products were measured according to the manufacturer's protocol (Applied Biosystems, Foster City, CA, USA). *GAPDH* was used for normalization (4310884E, Applied Biosystems).

### 2.3. Prediction of CpG islands in the EGR2 gene

CpG islands were predicted by means of a CPGPLOT program (EMBOSS, http://bioweb.pasteur.fr/seqanal/interfaces/cpgplot.html).

### 2.4. Sodium bisulfite sequencing

Genomic DNA was treated with sodium bisulfite according to a method described previously [9], with minor modifications to improve the effectiveness of the reaction. Genomic DNA (3.2 μg) was digested overnight at 37°C with 12 U each of *Acc*I, *Rsa*I, *Dra*I, *Hin*cII, *Bst*XI, *Cla*I, *Eco*RV, *Mlu*I, *Nde*I, *Sph*I, and *Xho*I (TaKaRa, Tokyo, Japan). The DNA fragments were purified and dissolved in distilled water, then denatured in 0.3 M NaOH at 42°C for 30 min. A freshly prepared solution containing sodium bisulfite (Sigma, St. Louis, MO, USA; final, 3.1 M, pH 5.0) and hydroquinone (Sigma; final 0.5 mM, pH 5.0) was added to each DNA and the solutions were incubated at 55°C for 18 h.

For DNA sequencing, the treated DNA fragments were amplified using a PCR primer set designed for region IVS1+114 to +272 (forward: 5′-GATTGTATTGGTTTTGAAGATTG-3′, reverse: 5′-CTCTTCCACCCCCATCCCT-3′). Thermal cycling was performed in 35 cycles of 94°C for 30 s, 55°C for 30 s, and 72°C for 2 min. The PCR products were cloned into pCR2.1 vector (Invitrogen, Carlsbad, CA, USA), and 12 clones were sequenced for each cell line.

*Corresponding author. Fax: (81)-3-5449 5433.
*E-mail address:* yusuke@ims.u-tokyo.ac.jp (Y. Nakamura).

0014-5793/03/$22.00 
doi:10.1016/S0014-5793(03)01092-5

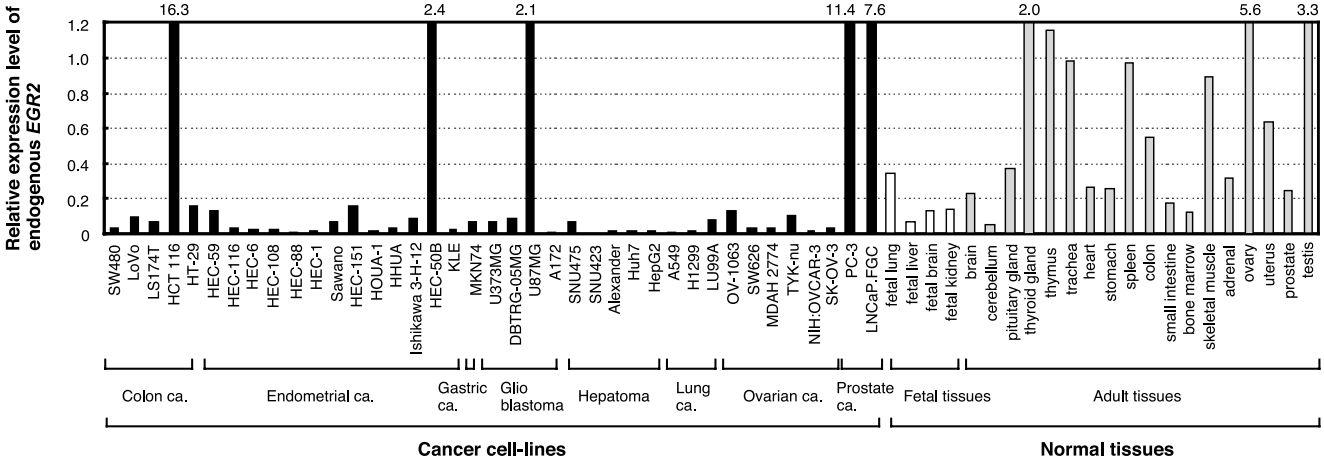


Fig. 1. Expression of endogenous *EGR2* examined by TaqMan-PCR analysis using cDNAs from 39 cancer cell lines, four fetal tissues, and 18 normal adult tissues. The gene expression in the cancer lines was measured in three separate experiments, and the average was compared with the average level of expression in the 18 adult tissues. The integrity of each template was controlled through amplification of *GAPDH*.

### 2.5. *Purification of proteins bound to methylated or unmethylated* EGR2 *intron 1*

Intron 1 was amplified by PCR using biotin-labeled forward primer 5′-(ATAACGCGT)ATACCATCCCAGGCTCAGTC-3′ and non-labeled reverse primer 5′-(TATAGATCT)ATCCCAGTCAGTGCCGTGAT-3′. The amplified, biotinylated PCR product was methylated by *Sss*I (CpG methylase; New England Biolabs, Beverly, MA, USA) under the conditions recommended by the manufacturer and its methylation status was confirmed by digestion with either a methylation-sensitive restriction enzyme, *Hap*II (TaKaRa) or an insensitive enzyme, *Msp*I (TaKaRa). Unmethylated or methylated products were mixed with streptavidin-Sepharose 4B (00560374, Zymed Laboratories, South San Francisco, CA, USA) overnight at 4°C. Binding was confirmed by electrophoresis in 1% agarose gels and free intron 1 fragments were removed.

HCT116 cells were harvested from 20 dishes (150 mm) and their nuclear extracts were prepared in a solution containing 20 mM HEPES (pH 7.6), 20% glycerol, 100 mM NaCl, 1.5 mM $MgCl_2$, 1.2 mM EDTA, 1 mM dithiothreitol, 0.1% NP40, and a protease inhibitor cocktail (Roche, Indianapolis, IN, USA). Appropriate competitors (42 μg of poly d[I-C], 20 μg of intron 1/resin complex, and 60 μg of free pGL3 empty vector) were added to the extracts, mixed overnight at 4°C, and competitive proteins binding to the intron 1/resin were removed by centrifugation. After 50 μg of free competitive intron 1 was mixed with the supernatant at 4°C for 1 h, 16 μg of either methylated or unmethylated intron 1/resin complex was added and mixed at 4°C for 3 h. The protein/intron 1/resin complexes were washed with the same buffer used for binding, and sample buffer was added to adjust the total volume to 125 μl. After the samples were incubated in boiling water for 3 min and centrifuged, the supernatants containing proteins that bound to methylated or unmethylated intron 1 were separated by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE).

### 2.6. *Antibodies for Western blotting*

Antibodies used here included mouse monoclonal antibody to human β-actin (AC-15; Sigma) and YY1 (sc-7341; Santa Cruz Biotechnology, Santa Cruz, CA, USA), goat polyclonal antibody to human MBD1, MBD2, and MBD3 (respectively sc-9395, sc-12444, and sc-9402; Santa Cruz), and rabbit polyclonal antibody to human MBD4 (sc-10753; Santa Cruz) and CTCF (06-917; Upstate, Lake Placid, NY, USA).

### 2.7. *Reporter gene assay using DNA fragments corresponding to intron 1 of the* EGR2 *gene*

Reporter gene plasmids were constructed to represent various lengths of nucleotide sequence in intron 1, each cloned into pGL3 promoter vector. The concentration of each plasmid clone was measured precisely, and DNAs were methylated. After heat inactivation of the enzyme, the methylation status of each plasmid was examined (see Section 2.5). Methylated and unmethylated constructs were separately co-transfected with pRL-TK (Promega, Madison, WI, USA) into HCT116 cells by means of FuGENE™6 Transfection Reagent (Roche). Two days after transfection the reporter assay was carried out using the Dual-Luciferase Reporter Assay system (Promega) according to the manufacturer's protocol.

## 3. Results

### 3.1. *Expression of endogenous* EGR2 *in cancer cell lines*

We determined the relative level of endogenous *EGR2* expression in each of 39 cancer cell lines by quantitative reverse transcriptase PCR, using the average expression level in 18 normal adult tissues as a baseline (Fig. 1). Expression of this gene was lower than the normal average in almost all of the cancer lines regardless of the organ of origin; the relative expression levels of 30 (77%) of the 39 cancer lines were less than 0.1, while those in 20 of the 22 normal tissues were higher than 0.1 ($P < 0.0001$: Fisher's exact probability test).

### 3.2. *Methylation status within* EGR2 *genomic sequences*

We examined the gene and its promoter sequences for screening somatic mutations or altered methylation in several cell lines whose endogenous *EGR2* expression was very low, but observed no changes in any of the cells examined (data not shown). However, by sequencing PCR products that were amplified from genomic DNAs treated with sodium bisulfite, we found that a CpG island present in intron 1 (IVS+114 to IVS+272; Fig. 2A,B) tended to be highly methylated in cell lines where endogenous *EGR2* expression level was relatively high, i.e. where the relative expression by TaqMan assay was $> 0.1$ (Fig. 2C). The same region was hypomethylated in cell lines where relative expression was $< 0.1$ (Fig. 2D). This correlation was observed in 13 of the 16 cell lines examined in this experiment; the three exceptional cases are documented in Fig. 2E. The methylation status of 18 additional sites in a CpG island that partially included exon 1 (+383 to IVS1+114) and 31 sites in another CpG island (IVS1+531 to IVS1+790) revealed similar correlations in several representative cell lines examined (data not shown).

### 3.3. *Is* EGR2 *intron 1 an insulator?*

At first, we investigated whether intron 1 is an insulator

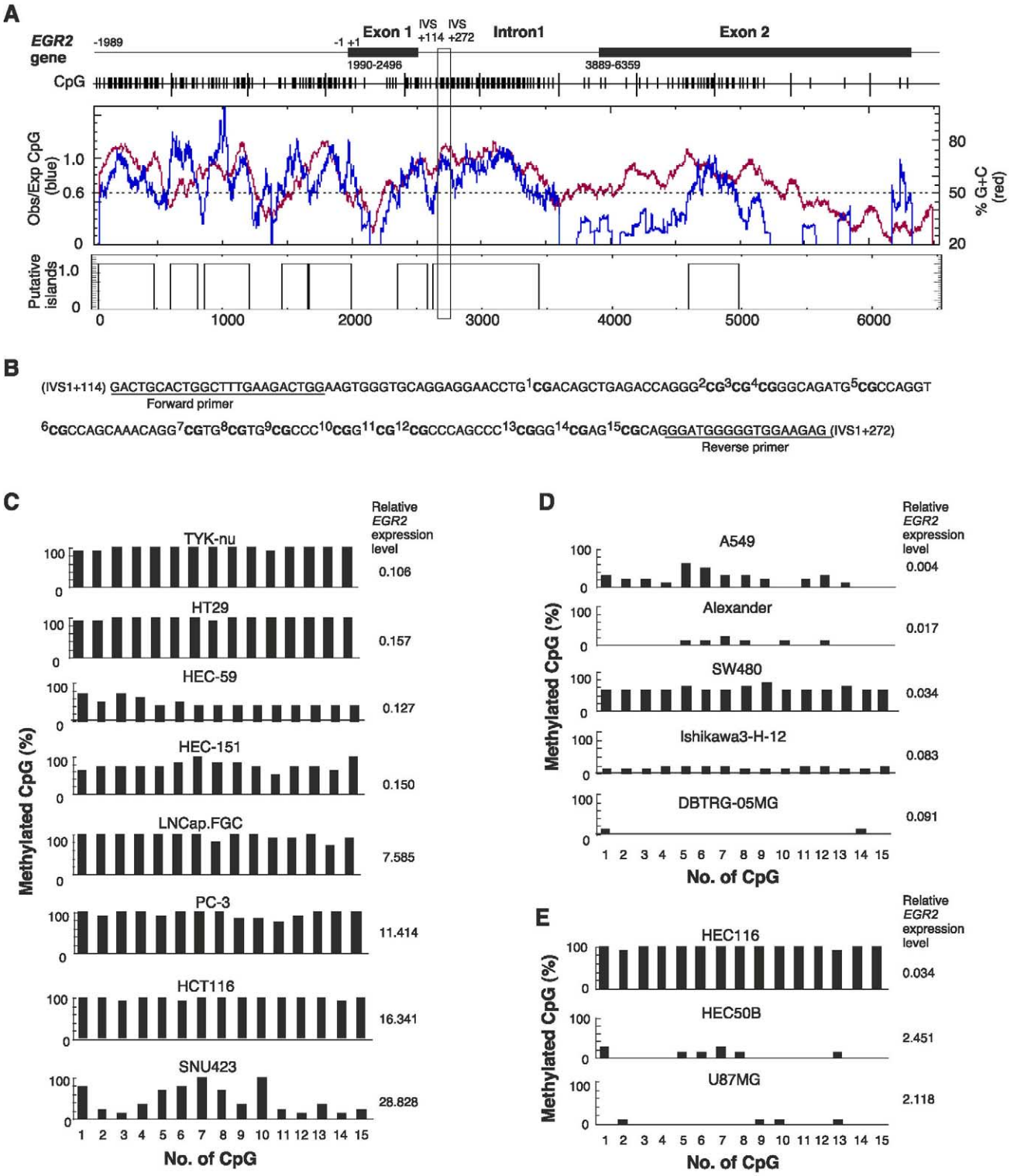


Fig. 2. Predicted CpG islands within the *EGR2* gene and methylation status of intron 1 in cancer cell lines. A: Map of predicted CpG islands. Moving averages of %G+C (red) and Obs/Exp CpG (blue) were calculated using the CPGPLOT (EMBOSS) program. An area surrounded by a vertical rectangle (IVS+114 to IVS+272) indicates the region described in B. B: Sequence of IVS+114 to IVS+272 in intron 1, with 15 CpG dinucleotides indicated in boldface and underlined. The primer set indicated here was used for subsequent sodium bisulfite sequencing analysis. C: Hypermethylation in eight cell lines showing relatively high expression of *EGR2*. The 'relative expression' levels indicate results of TaqMan-PCR experiments (see Fig. 1). D: Hypomethylation in intron 1 of *EGR2* in five low-expression cell lines. E: Results in three exceptional cancer cell lines.

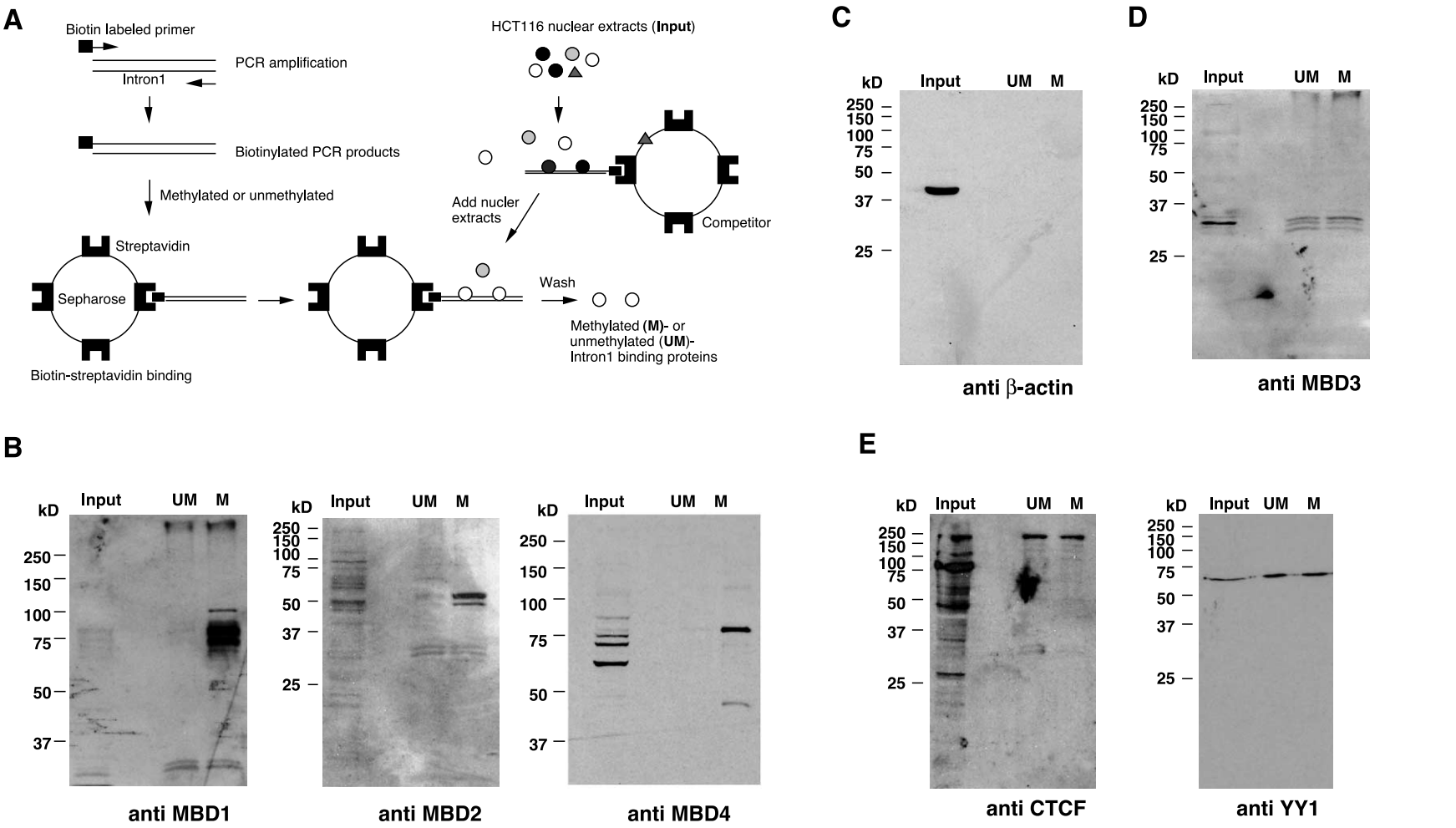


Fig. 3. Examination of binding affinity of insulator-binding proteins to *EGR2* intron 1. A: Schematic representation of the steps used to purify these proteins (see Section 3.3). B: Methyl-CpG-binding proteins MBD1, MBD2, and MBD4 were detected only in the M lanes. 'Input' indicates lanes containing untreated nuclear extracts from HCT116 cells; M and UM indicate lanes containing methylated and unmethylated intron 1-binding proteins, respectively. C: β-Actin was removed by the protocol. D: MBD3, which has little specificity for methyl-CpG, bound equally to unmethylated and methylated intronic sequences. E: Insulator-binding proteins CTCF and YY1 were detected in both lanes, at equal levels.

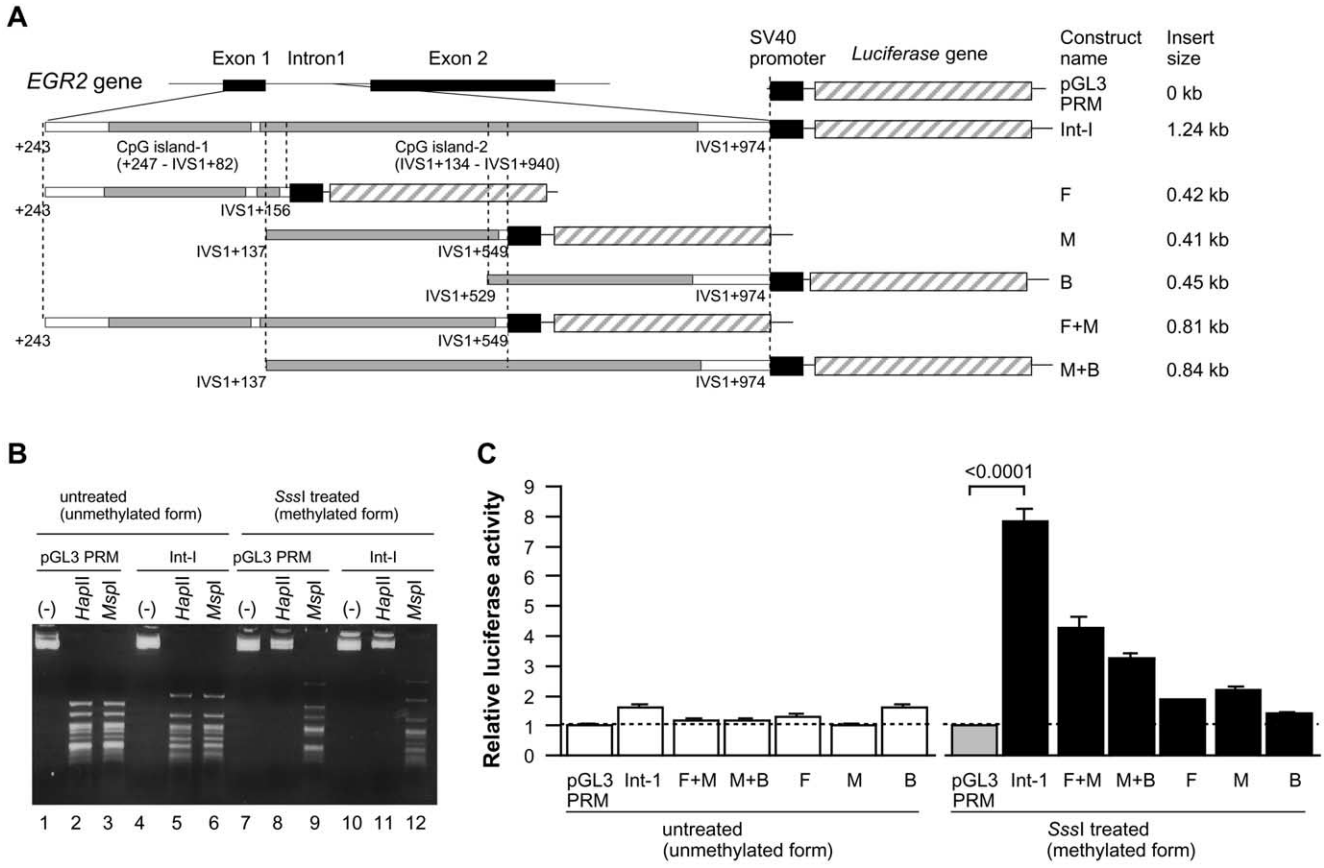


Fig. 4. Enhancer-like activity of methylated intron 1 of *EGR2*. A: Construction of reporter gene plasmids representing various lengths of the intron 1 sequence sub-cloned into pGL3 promoter vector. B: Confirmation of methylation status of the constructs using a methylation-sensitive enzyme, *Hap*II, or an insensitive enzyme, *Msp*I. C: Each methylated or unmethylated construct was co-transfected with pRL-TK into HCT116 cells. Luciferase activity was measured 48 h later, and relative luciferase activity was calculated based on the activity of pGL3 PRM as 1.0 (respectively unmethylated and methylated). Each data bar represents the mean of values from three separate experiments; error bar, S.D. (Scheffé's *F* test).

that has an enhancer-inhibiting activity when the site is unmethylated. So far, two insulator-binding proteins, CTCF and YY1, have been isolated [10–13]. To investigate the possibility that these proteins are associated with *EGR2* expression, we purified proteins that have binding affinity for either methylated or unmethylated DNA (Fig. 3A). Each form of the 1.2-kb intronic fragment was biotinylated during PCR amplification, and the *Sss*I-treated or -untreated PCR product was bound to streptavidin-Sepharose (resin). The fragment–resin complex was mixed with nuclear extracts from HCT116 cells and appropriate competitors; proteins that bound to one or the other DNA fragment were purified and separated by SDS–PAGE. Fig. 3B–D shows results of experiments to validate our protocol. Proteins known to have higher binding affinity for methyl-CpG, such as MBD1, MBD2, and MBD4 [10], were detected only in the lanes containing proteins bound to the methylated intronic sequence (Fig. 3B); no β-actin, the negative control, was detected (Fig. 3C). MBD3, which has no specific binding affinity for methyl-CpG [10], bound equally to both the unmethylated and methylated intron 1 sequences (Fig. 3D). Similar to MBD3, no specific binding was observed when we examined binding of two insulator-binding proteins, CTCF [11–13] and YY1 [14], to intron 1 (Fig. 3E).

### 3.4. *Transcriptional activity of* EGR2 *intron 1*

We then investigated unmethylated or methylated forms of a 1.2-kb intronic segment that included two CpG islands by means of a luciferase reporter assay. Reporter plasmids (pGL3 promoter vector) with basal luciferase activity were constructed to include various parts of this intronic DNA segment (Fig. 4A), and enzymatically methylated. The methylation status was confirmed by digestion with two restriction enzymes (Fig. 4B). Fig. 4C represents results of the assay using HCT116 cells, which expressed *EGR2* at a high level and had a methylated intron 1, after transfection with *Sss*I-treated or -untreated plasmids. The methylated 1.2-kb intronic fragment revealed the highest enhancer-like activity among the fragments tested, but the corresponding unmethylated form showed none at all. Either of the two 0.8-kb fragments indicated in Fig. 4A as F+M and M+B, as well as each of the three 0.4-kb fragments (F, M, and B), revealed respectively 40–60% and 20–35% enhancer-like activity compared with the 1.2-kb fragment.

## 4. Discussion

Endogenous expression of *EGR2* in most of the cancer cell lines we examined was less than in normal tissues. Somatic

mutations or aberrant methylation in the *EGR2* promoter region were ruled out by sequencing and by analysis of methylation status of the promoter by sodium bisulfite sequencing (data not shown). However, CpG islands in intron 1 tended to be highly methylated in cancer cell lines that were expressing the gene abundantly, and to be hypomethylated in low-expressing lines. A few other reports have indicated that DNA methylation does not always confer a gene-silencing effect [11–15]. Therefore on the basis of our findings we hypothesized that (1) an insulator-binding protein such as CTCF [11–13] and YY1 [14] may bind to the unmethyl-CpG and inhibit enhancer activity; (2) a methylation-sensitive trans-suppressive protein may down-regulate gene expression by binding to the unmethylated intron; (3) a trans-activative methyl-CpG-binding protein can bind to the methylated intron; (4) a methylation-sensitive covering protein may inhibit the binding of trans-activating factors to the unmethylated intron. To investigate these possibilities, at first we examined the possibility that the *EGR2* intron 1 was an insulator. We attempted to purify proteins that bound specifically to the methylated or unmethylated forms of intron 1 and examined binding affinity of CTCF and YY1 to intron 1. As a result, they showed no specific binding to the unmethylated intron and were not likely to be the regulator of *EGR2* expression. MBD1, 2, and 4 used here as positive control are also considered to be possible candidates for the *EGR2* expression regulator. Since MBD2 forms complexes with a variety of proteins and might have several functions [16–18], it is possible that other members of the MBD protein family would be able to form complexes with unreported proteins and have unknown functions to regulate the gene expression.

Although the possibility that intron 1 functions as an insulator is not completely excluded, we suggest here other possibilities by detecting transcriptional activity of unmethylated or methylated fragments of the intron 1 sequence. A methylated 1.2-kb fragment showed the highest enhancer-like activity among the regions tested, while the corresponding unmethylated fragment lacked any such activity supporting the third or fourth hypothesis. Since shorter fragments of intron 1 revealed lower enhancer-like activity compared with the 1.2-kb sequence, structural conformation is probably an important feature of the transcription-enhancing activity we observed, in conjunction with protein interactions involving this genomic segment.

We think our data strongly imply the existence of a novel mechanism that up-regulates gene expression through a combination of intronic DNA methylation and structural conformation. We think elucidation of this mechanism will not only lead to an understanding of the *EGR2* transcriptional regulation mechanism but also shed light on the additional mechanism of the gene expression regulation. Although promoter regions of some tumor suppressors are reportedly hypermethylated in cancer, genomic DNA from cancers generally tends to be totally hypomethylated [19]. One of the keys to solving this discrepancy might be found in the phenomenon reported here.

*Acknowledgements:* We thank Drs. Takashi Shimokawa, Masashi Morita and Hidewaki Nakagawa for helpful advice and discussion. This work was supported in part by Research for the Future Program Grant #00L01402 from the Japan Society for the Promotion of Science (JSPS).

## References

[1] Matsushima-Nishiu, M., Unoki, M., Ono, K., Tsunoda, T., Minaguchi, T., Kuramoto, H., Nishida, M., Satoh, T., Tanaka, T. and Nakamura, Y. (2001) Cancer Res. 61, 3741–3749.
[2] Unoki, M. and Nakamura, Y. (2001) Oncogene 20, 4457–4465.
[3] Unoki, M. and Nakamura, Y. (2003) Oncogene 22, 2172–2185.
[4] Chavrier, P., Zerial, M., Lemaire, P., Almendral, J., Bravo, R. and Charnay, P. (1988) EMBO J. 7, 29–35.
[5] Sham, M.H., Vesque, C., Nonchev, S., Marshall, H., Frain, M., Gupta, R.D., Whiting, J., Wilkinson, D., Charnay, P. and Krumlauf, R. (1993) Cell 72, 183–196.
[6] Schneider-Maunoury, S., Topilko, P., Seitandou, T., Levi, G., Cohen-Tannoudji, M., Pournin, S., Babinet, C. and Charnay, P. (1993) Cell 75, 1199–1214.
[7] Topilko, P., Schneider-Maunoury, S., Levi, G., Baron-Van Evercooren, A., Chennoufi, A.B., Seitanidou, T., Babinet, C. and Charnay, P. (1994) Nature 371, 796–799.
[8] Warner, L.E., Mancias, P., Butler, I.J., McDonald, C.M., Keppen, L., Koob, K.G. and Lupski, J.R. (1998) Nat. Genet. 18, 382–384.
[9] Rein, T., Zorbas, H. and DePamphilis, M.L. (1997) Mol. Cell. Biol. 17, 416–426.
[10] Ballestar, E. and Wolffe, A.P. (2001) Eur. J. Biochem. 268, 1–6.
[11] Bell, A.C., West, A.G. and Felsenfeld, G. (1999) Cell 98, 387–396.
[12] Bell, A.C. and Felsenfeld, G. (2000) Nature 405, 482–485.
[13] Hark, A.T., Schoenherr, C.J., Katz, D.J., Ingram, R.S., Levorse, J.M. and Tilghman, S.M. (2000) Nature 405, 486–489.
[14] Kim, J., Kollhoff, A., Bergmann, A. and Stubbs, L. (2003) Hum. Mol. Genet. 12, 233–245.
[15] Mager, J., Montgomery, N.D., de Villena, F.P. and Magnuson, T. (2003) Nat. Genet. 33, 502–507.
[16] Ng, H.H., Zhang, Y., Hendrich, B., Johnson, C.A., Turner, B.M., Erdjument-Bromage, H., Tempst, P., Reinberg, D. and Bird, A. (1999) Nat. Genet. 23, 58–61.
[17] Tatematsu, K.I., Yamazaki, T. and Ishikawa, F. (2000) Genes Cells 5, 677–688.
[18] Bhattacharya, S.K., Ramchandani, S., Cervoni, N. and Szyf, M. (1999) Nature 397, 579–583.
[19] Feinberg, A.P. and Vogelstein, B. (1983) Nature 301, 89–92.