

Human Dbf4/ASK promoter is activated through the Sp1 and MluI cell-cycle box (MCB) transcription elements

Xing Wu¹ and Hoyun Lee^{*.1}

¹Northeastern Ontario Regional Cancer Centre, Sudbury, Ontario P3E 5J1, Canada

Dbf4 is the regulatory subunit of Cdc7 kinase, which is essential for entry into and traversing through S phase. The level of Dbf4, which is critical for the activation of Cdc7, is regulated by transcription and protein degradation. To gain a better understanding as to how the transcription of human Dbf4 (HuDbf4) is regulated, we have cloned and characterized its promoter. We found that HuDbf4 core promoter is localized within –211 to –285 of the translation start-codon. This 75 bp DNA segment contains, among others, a putative MluI Cell-cycle Box (MCB). A point mutation within the MCB dramatically reduced the promoter activity. This is the first example that an MCB element plays an essential role in the activation of a core promoter in mammalian cells. The auxiliary elements required for the full promoter activity are present within 162-bp upstream from the core promoter (i.e., –286/–447). A point mutation within the Sp1 element at –353/–361 resulted in a decrease of promoter activity to the basal level, while the deletion of the putative HES-1 at –326/–331 dramatically increased the promoter activity. Taken together, our data suggests that the MCB element is essential for the core promoter activation, while the Sp1 positive regulator and the HES-1 repressor coordinately determine the efficiency of the HuDbf4 promoter. We have also found: (i) that the major transcription initiations occur at –220, –235 and –245; (ii) that HuDbf4 gene consists of 12 exons, which spread over a 33-kb region.

Oncogene (2002) 21, 7786–7796. doi:10.1038/sj.onc.1205914

Keywords: Dbf4; ASK; promoter; Sp1; MluI cell-cycle box (MCB); gene structure; transcription initiation

Introduction

The Cdc7-Dbf4 serine/threonine kinase is essential for the activation of origins of DNA replication (*oris*) throughout S phase (Bousset and Diffley, 1998;

Donaldson *et al.*, 1998). Although Cdc7-Dbf4 kinase can phosphorylate several replication-related proteins *in vitro* (Masai *et al.*, 2000; Nougarede *et al.*, 2000), its important physiological substrate *in vivo* is thought to be Mcm2 protein (Jares and Blow, 2000; Jiang *et al.*, 1999). The phosphorylation of Mcm2p by Cdc7-Dbf4 kinase is essential for *ori* activation; however, it alone is not sufficient (Jares and Blow, 2000; Mimura and Takisawa, 1998; Owens *et al.*, 1997; Tercero *et al.*, 2000; Treuner *et al.*, 1998; Zou and Stillman, 1998, 2000). Subsequent activation of Cdc45p by Cdc7-Dbf4 kinase appears to be the final trigger for replication initiation (Jares and Blow, 2000; Mimura and Takisawa, 1998; Owens *et al.*, 1997; Tercero *et al.*, 2000; Treuner *et al.*, 1998; Zou and Stillman, 1998, 2000). Thus, Cdc7-Dbf4 kinase may be involved in the activation of chromosome replication at the final two steps, suggesting that it functions as a critical molecular switch for DNA replication (Masai *et al.*, 1999).

The levels of Cdc7 mRNA and protein are relatively constant throughout the cell cycle in the proliferating cell population. However, its kinase activity is cell-cycle dependent, as it is very low in G1 and high throughout the S phase (Ferreira *et al.*, 2000; Jiang *et al.*, 1999; Kumagai *et al.*, 1999). Cdc7 kinase is activated by associating with the Dbf4 regulatory subunit, which generally coincides with high levels of Dbf4 protein (Jackson *et al.*, 1993; Nougarede *et al.*, 2000; Oshiro *et al.*, 1999). Since Dbf4p is critically important for the regulation of Cdc7 protein kinase, cells have developed tight control mechanisms to maintain appropriate levels of Dbf4p. The primary control point of the Dbf4p level is at the initiation of transcription (Chapman and Johnston, 1989; Jackson *et al.*, 1993; Kumagai *et al.*, 1999; Lepke *et al.*, 1999). We have previously demonstrated using Chinese hamster ovary (CHO) cells that there is no detectable level of Dbf4 mRNA in the resting cells (i.e., G0) (Guo and Lee, 2001). When CHO cells in G0 were released into the cell cycle, there was only a slight increase of the Dbf4 mRNA level during G1; however, the levels increased at least threefold as the cells reached the G1/S transition, and maintained high levels during S phase (Guo and Lee, 2001). Although the levels of Dbf4p are also regulated by both ubiquitine-dependent and -independent protein degradation (Ferreira *et al.*, 2000; Oshiro *et al.*, 1999), the protein levels generally

*Correspondence: H Lee, Department of Research, Northeastern Ontario Regional Cancer Centre, 41 Ramsey Lake Road, Sudbury, Ontario, P3E 5J1, Canada; E-mail: hlee@neorcc.on.ca

Received 10 April 2002; revised 24 July 2002; accepted 25 July 2002

mirror those of mRNA (Jiang *et al.*, 1999), at least in part, due to the short half life of Dbf4p (Ferreira *et al.*, 2000; Nougarede *et al.*, 2000; Oshiro *et al.*, 1999; Weinreich and Stillman, 1999).

The initiation of DNA replication in mammalian cells is rapidly and transiently downregulated in response to genotoxic agents such as ionizing radiation or chemotherapeutic drugs (Larner *et al.*, 1999; Lee *et al.*, 1997; Paulovich and Hartwell, 1995). The molecular components that constitute this control mechanism are not well understood; however, an inactivation of Dbf4 may downregulate the initiation of DNA replication through the regulation of the Cdc7 catalytic subunit. Consistent with this expectation, we previously found that the level of Dbf4 mRNA was rapidly and transiently downregulated in response to ionizing radiation (Guo and Lee, 2001). Similarly, Siede *et al.* (1994) found in *Saccharomyces cerevisiae* that Cdc7-Dbf4 kinase is involved in the operation of the G1/S checkpoint. Thus, the control of the Dbf4 transcription initiation appears to be critically important for both the positive and negative regulations of chromosome replication.

To gain a better understanding about how the initiation of Dbf4 transcription is regulated, we have cloned and characterized the human Dbf4 (HuDbf4, or ASK for activator of S-phase kinase) promoter. We found that the HuDbf4/ASK core promoter activity is localized within the nucleotide (nt) 211 to 285 upstream of the translation start-codon (i.e., -211/-285). This 75 bp DNA segment contains putative binding sites for Sp1, TFIIB, and a *Mlu*I Cell-cycle Box (MCB). In addition, a TATA-like element (TTTCAA) is also present in the core promoter region, which may be relevant for the utilization of three major transcription initiation sites in the vicinity of this element. Our data also suggests that the Sp1 transcription activator element at -353/-361 and the HES-1 repressor element at -326/-331 are involved in the control of transcription efficiency. Using PCR and nucleotide sequence analysis, we also determined the organization of HuDbf4/ASK gene.

Results

Identification of DNA segment that contains the HuDbf4/ASK promoter

The nucleotide sequence analysis of the two BAC clones (Figure 1a; GenBank Accession Numbers 005164 and 003083) revealed the presence of an open reading frame, which encodes a putative mitochondrial carrier protein, at 20 kb upstream of HuDbf4/ASK. Considering the opposite transcription polarity of these two genes, it was thought that the transcription regulatory elements for the Dbf4 gene might be located between these two genes. Therefore, we first generated genomic libraries with a DNA segment containing this region as described in Materials and methods. Subsequently, the genomic libraries were pooled and transiently transfected into MDA-MB231 or 293T cells

to identify the DNA fragments containing promoter activity. The promoter activity, which was manifested by the presence of fluorescence, was measured at 48 h post-transfection using fluorescence microscopy, luminescence spectrometry, and flow cytometry. The DNA segment BE5, which spans from -170 to -4919, showed strong promoter activity (Figure 1b,c). In addition to BE5, several smaller DNA fragments also showed a various degree of promoter activity. All of these clones contained the DNA segment -200 to -400 region. Therefore, to precisely map the promoter, subfragments of the 4750 bp BE5 were generated using PCR primers described in Table 1 and examined for their promoter activities.

The relative promoter activity of each subclone was semi-quantified by flow cytometry as described in the legend to Figure 1b. We assumed in this quantification that the promoter activity is generally proportional to the fluorescence intensity. The minor cell populations that were separated from the major group were not included in our calculation, since they could be due to high plasmid copy numbers. According to these specifications, 'endpoints' on the x-axis of the subclones -516/-993, -211/-1890, -211/-993, -170/-4919, and the pEY vector (negative control) were 73, 146, 141, 173, and 75, respectively (Figure 1b). This set of experiments was repeated four times and the mean value of the relative fluorescence intensity for each clone was illustrated in a bar as shown in Figure 1c. Since the promoter activities of the DNA segments -170/-4919 (BE5) and -170/-1628 were essentially the same, all the essential promoter elements are likely present within the 1628 bp upstream of the huDbf4/ASK translation start-codon (Figure 1c). We found that none of the clones representing the upstream region from 516 to 3405 (i.e., -516/-993, -515/-1297, -973/-1645, -1379/-2194, or -2742/-3405) showed any promoter activity, while -279/-437 and -428/-993 showed only very low levels of promoter activity. In contrast, the -211/-285 short segment showed quite strong promoter activity (Figure 1c). As shown in Figure 2a, we further analysed the promoter activity of several subclones that are generated from -170/-1628. We found that the fragment -144/-215 showed very little promoter activity, compared to that of -211/-285 (Figure 2a). It is, thus, likely that the 75 bp DNA segment spanning from nt211 to nt285 upstream of the huDbf4/ASK translation start-codon contains the transcription core elements required for the basal level of the HuDbf4/ASK promoter activity. This DNA segment contains a putative Sp1 element (GGGCGGGG) at -278/-286, a TFIIB element (CGGCGCC) at -262/-268, and an MCB (ACGCGT) at -245/-250 (Figure 2b,c; also see below).

The promoter activity of the DNA segment -211/-447 is significantly higher than that of -211/-285 (Figure 2a). However, the DNA segment -279/-437 showed only very low promoter activity (Figure 1c), suggesting that the nucleotide sequence between -285

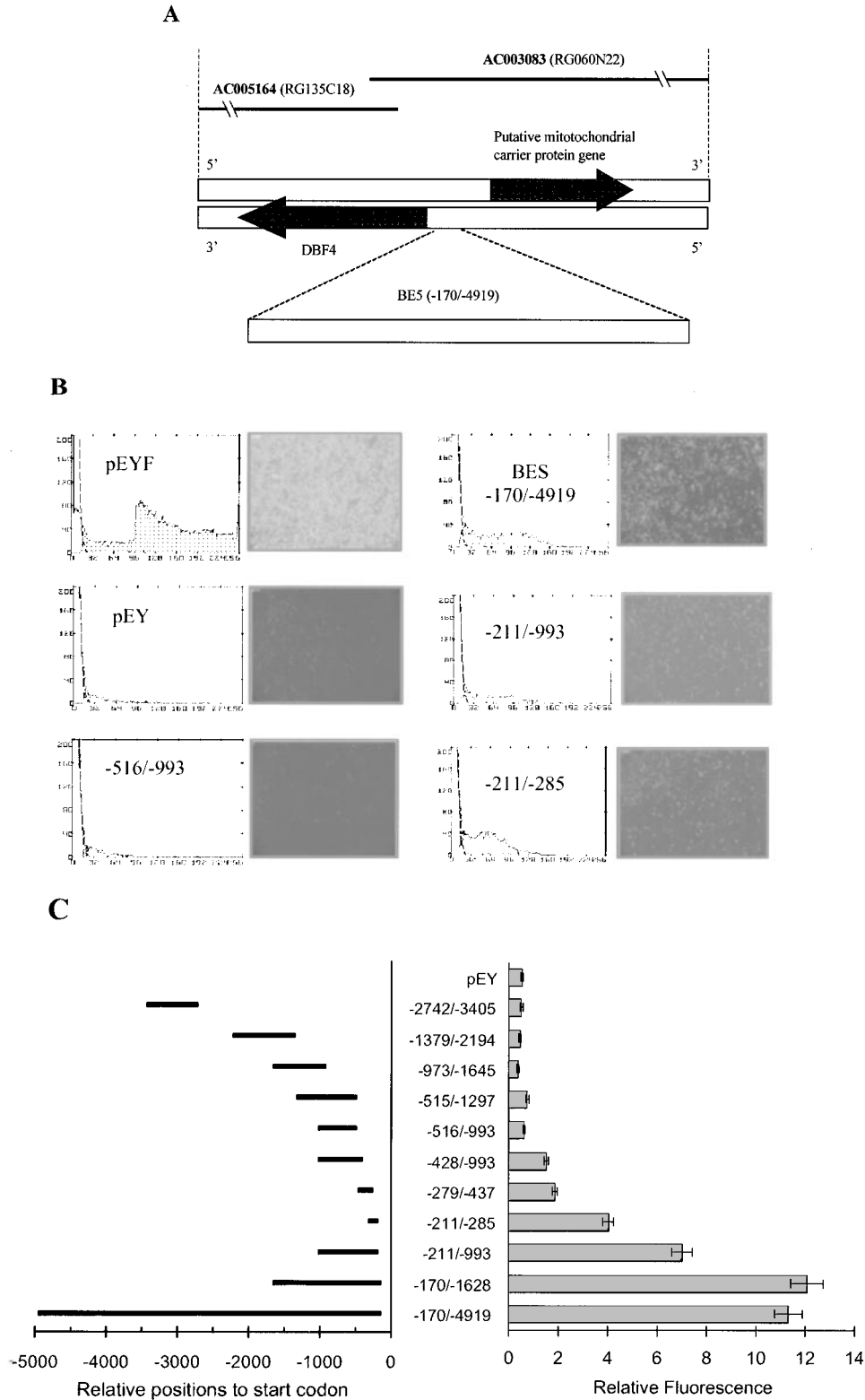


Figure 1 Mapping the HuDbf4/ASK promoter. (a) The relative position of the 4.8 kb BE5 (−170/−4919) in the BAC clones RG060N22 (GenBank Accession Number AC003083) and RG135C18 (Accession number, AC005164) is shown. An open reading frame that encodes putative mitochondrial carrier protein in the opposite direction is located approximately 20-kb upstream of HuDbf4/ASK. (b) Promoter activity was determined by either flow cytometry (left panels) or fluorescence microscopy (right panels) after the plasmid constructs (marked in each panel) were transfected into 293T cells as described in Materials and methods. The numbers on the x-axis of flow cytometry represent arbitrary ‘channel’ numbers. The original acquisition of the fluorescence intensity of each cell was in a log scale. However, the cells with different fluorescence intensity were arbitrarily assigned into 256 channels in a linear fashion. Therefore, the cells in the channel with higher numbers indicate that they emit higher fluorescence signal. However,

Table 1 The PCR primers that were used to generate DNA segments

Constructs (segments)	PCR primers (5' to 3')		Nucleotide number on AC003083	Relative positions to start codon
	Forward	Reverse		
-1397/-2194	TCAGTGGATCCACGCAGAG AACAGGTTTGC	CATACGAATTCGGATACCT GAAACGCCAAAA	2984-3799	-1397/-194
-973/-1645	TCAGGATCCGTAGTTGGCTA TTTGGCGCTGT	CGGAATTCCTCCACCACTC AAGATCTGTTAAC	2578-3250	-973/-1645
-515/-1297	TGCGGATCCTCTGTGTTGA CCAAACCTAG	CGGAATTCCTTCCTTGAGC TACCAGGTAAA	2119-2902	-515/-1297
-515/-993	TGCGGATCCTCTGTGTTGA CCAAACCTAG	CGGAATTCACAGCGCAA TAGCCAACCTAC	2119-2598	-525/-993
-428/-993	TCAGGATCCCTCCATTCCG TTGCTCTCTC	CGGAATTCACAGCGCAA TAGCCAACCTAC	2033-2598	-428/-993
-279/-437	TGCGGATCCCGCCCCGTCT CACGCCA	CGGAATTCGGAATGGAGG CGGGGTAGAGGC	1884-2042	-279/-437
-211/-993	TGCGGATCCAAACGAGTGG GCTGCGGCG	CGGAATTCACAGCGCAA TAGCCAACCTAC	1816-2598	-211/-993
-211/-535	TGCGGATCCAAACGAGTGG GCTGCGGCG	CGGAATTCACTAGGTTTGG TCAACACAGAG	1816-2140	-211/-535
-211/-447	TGCGGATCCAAACGAGTGG GCTGCGGCG	CGGAATTCGAGAGAGCAA CGGAATGGAGGC	1816-2052	-211/-447
-211/-285	TGCGGATCCAAACGAGTGG GCTGCGGCG	CGGAATTCGGGCGGGGCG CGCGTATCGGCG	1816-1890	-211/-285

The nucleotide sequences underlined are *Bam*HI or *Eco*RI restriction sites that were included in the primers

and -447 does not have an essential core promoter element, but is likely to contain important auxiliary elements that can further increase the basal promoter activity. This DNA segment contains three putative Sp1 elements -422/-428, -353/-361, and -307/-313) and one putative HES-1 (-326/-331) element (Figures 2b,c and 3a).

The Sp1 element at -353/-361 is critical for an efficient promoter activity

Since the DNA segment -211/-447 showed almost full promoter activity (Figure 2a), it was thought that both the core elements and the binding sites for most of the auxiliary transcription factors are present within this DNA segment. Therefore, we generated several mutants within this DNA segment and determined the promoter activity of each mutant. To our surprise, the construct containing a deletion from -302 to -352 (i.e., the -211/-447_{del} construct in Figure 3) showed significantly increased promoter activity, compared to the undeleted control (Figure 3b). A DataBank analysis suggested that a putative binding site for the HES-1 repressor is present within the deleted region. Although we cannot rule out other possibilities, a simple, straightforward interpretation of our data is that the deletion of the putative HES-1 element could have resulted in the increase of the promoter activity.

To determine the importance of the putative Sp1 element at -353/-361 (Box 3), we altered the

nucleotide sequence -359GGC-357 to TTG and determined its promoter activity. This mutant showed approximately 50% lower promoter activity compared to the -211/-447 'wild-type' construct (Figure 3b), which was comparable to that of the basal promoter activity.

Our preliminary EMSA experiments using agarose gel and the entire DNA fragment from -211 to -447 (construct p1816-2052) suggested that there was a DNA band mobility shift, which was inhibited by an oligo nucleotide that includes the Sp1 Box 3 (data not shown). Therefore, we determined by EMSA as to whether this putative Sp1 element is actually bound by the transcription factor Sp1. As shown in Figure 4a, the ³²P-end-labeled 23-mer DNA migrated much more slowly when mixed with nuclear extracts, compared to the free 'probe' that is shown at the bottom of the gel, suggesting the 23-mer DNA oligo formed a complex with protein(s) (lane 1). This electric mobility shift was not shown when the 23-mer unlabeled oligo (Figure 4b) or the 29-mer 'wild-type' Sp1 DNA (Figure 4c) was added in the reaction mixture (Figure 4a, lanes 2 and 3). However, the mutant Sp1 oligo DNA that contains a change of three nucleotides from GGC to TTG (Figure 4c) did not inhibit the formation of the DNA-protein complex (Figure 4a, lane 4), suggesting that the protein binding to the putative Sp1 Box 3 is Sp1 transcription factor. This was confirmed by the fact that anti-Sp1 antibody binding to the DNA-protein complex resulted in a 'supershift' of the complex in

the numbers themselves do not represent the original fluorescence intensity of each cell in a linear or a log scale but relative fluorescence intensity, thus relative promoter activity. (c) Relative promoter activities estimated, as above, are illustrated as bars. All the experiments were carried out in duplicate and repeated four times. The numbers at the center panel are DNA segments that were cloned into the pEY vector. For example, -170/-4919 is the plasmid construct containing the DNA segment from nt170 to nt4919 upstream of the HuDbf4/ASK translation start-codon

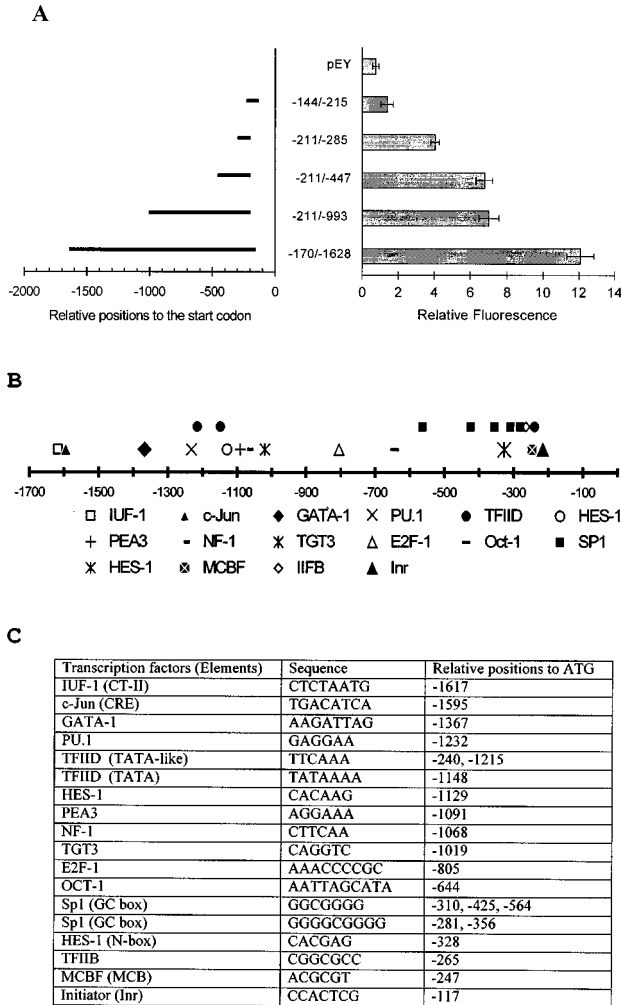


Figure 2 Mapping of the core promoter and identification of putative binding sites for transcription factors. (a) To localize the DNA segment containing a core promoter, the segment from -170 to -1628 was further subcloned and examined for promoter activity of each subfragment. (b and c) Putative binding sites for transcription factors are identified using computer-based analysis as described in Materials and methods. The numbers are relative nucleotide positions from the HuDbf4/ASK translation start-codon

migration rate, which was inhibited by an Sp1 peptide (Figure 4a, lanes 5 and 6). Taken together with the data shown in Figure 3, the element at -353/-361 (i.e., Sp1 Box 3) is bound by Sp1 transcription factor and is involved in the increase of the huDbf4/ASK promoter activity.

MluI Cell-cycle Box (MCB) is essential for the basal promoter activity of HuDbf4/ASK

The 75 bp core promoter -211/-285 contains a putative MCB element (Figure 5). Since the involvement of an MCB element in the regulation of mammalian transcription is currently unknown, we examined the effects of an MCB mutation on promoter activity. As shown in Figure 5b, a 'C' to 'T' transition

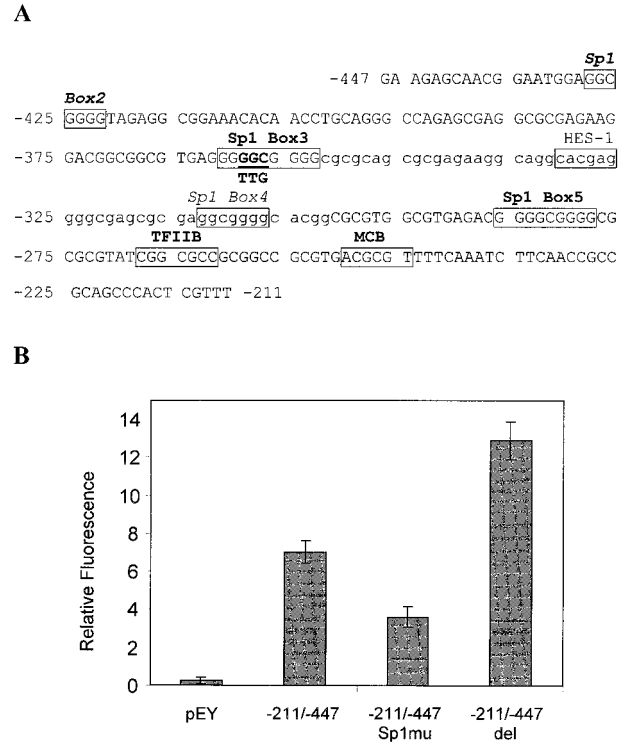


Figure 3 Promoter activity was decreased when a mutation was introduced within the Sp1 Box 3, but dramatically increased when the putative HES-1 element was deleted. (a) The nucleotide sequence of the putative transcription-factor binding sites of the DNA segment -211/-447 is shown. The mutant, -211/-447 Sp1_{mut} was generated by substitution of GGC (underlined) with TTG within the Sp1 Box 3. The -211/-447_{del} mutant was generated by deleting 51 bp spanning from -302 to -352 (marked as lower case) from the -211/-447 construct. Putative Sp1 sites written in italic may not be functional. (b) Relative promoter activity of each construct was estimated as per Figure 1 legend

mutation at -249 within the putative MCB dramatically reduced the promoter activity.

Initiator (Inr) and a TATA-like element in the core promoter

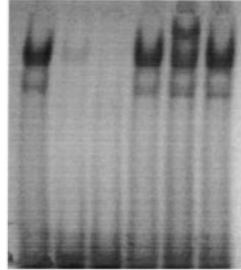
A putative Initiator element, CCACTCG, was found at -214/-220 within the core promoter region (Figure 5a). However, no typical TATA box is found in the vicinity of the transcription initiation sites. Interestingly, a TATA-like sequence, TTTCAAA, is found at nt238 to nt244 upstream of the ATG start-codon (Figures 5 and 6). A sequence motif search of the Eukaryotic Promoter Databank (Perier *et al.*, 2000) showed that there are at least seven eukaryotic genes that contain 5'-(T)TTTCAA(A)-3' near transcription initiation sites (Figure 6a).

The initiation of huDbf4/ASK transcription occurs at multiple sites

To determine the transcription initiation site of the huDbf4/ASK gene, primer extension analysis was carried out using the primer -106, the 5' end of

A

Treatments \ Lanes	1	2	3	4	5	6
Unlabeled oligo competitor		+				
Wild-type Sp1 oligo competitor			+			
Mutant Sp1 oligo (GGC to TTG)				+		
Anti-Sp1 antibody					+	+
Sp1 peptide						+
Nuclear extracts	+	+	+	+	+	+



B Human Dbf4 promoter

-375 gacggCGGCG TGAGGGGCG GGC CGCGCag cgcgagaagg caggcacgag
 Sp1 Box3
 TTG

C Competitor oligos

Wild-type Sp1 oligo competitor: 5'-CCCTTTGTGGGGCGGGGCGCTAAGCTGCG-3'
 Mutant Sp1 oligo competitor: 5'-CCCTTTGTGGGTTGGGGCGCTAAGCTGCG-3'

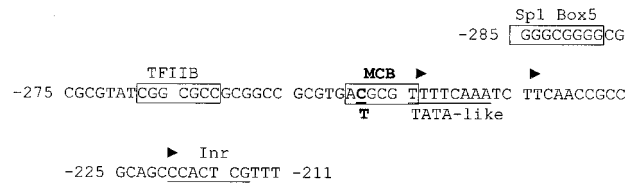
D Competitor Sp1 peptide (peptide DNA binding epitope domain): CKDSEGRSGSDPGKK

Figure 4 The putative Sp1 element at $-353/-361$ is bound by Sp1 transcription factor. (a) EMSA was carried out using HeLa cell nuclear extracts and the double-stranded DNA probe from -347 to -370 (bold in b). In the competition experiments, unlabeled otherwise identical oligo probe, wild-type or mutant Sp1 oligo that was generated by substituting GGC (underlined) with TTG (c), or Sp1 peptide (d) was added into the reaction. Rabbit anti-Sp1 polyclonal antibody was used for the supershift assay (lanes 5 and 6)

which is corresponding to nt106 upstream of the translation start-codon. As shown in Figure 7, data obtained by the primer extension method using mRNAs isolated from both MCF7 and 293T suggests that there are six potential transcription initiation sites (-220 , -235 , -245 , -305 , -323 and -332) within 350 nucleotides upstream of the huDbf4/ASK translation start-codon. Among the six, initiation at -235 appears to occur most frequently. We repeated the experiment four times and found the initiation patterns were always similar. Since the HuDbf4/ASK cDNA that was reported by Kumagai *et al.* (1999, GenBank Accession Number AB028069) contains 504 nucleotides upstream of the translation start-codon, we also carried out primer extension experiments using the primers derived from -402 and -796 . The data from these experiments suggests that transcription initiations also occur at -452 , -742 , and -847 (Figure 7a).

To confirm the data obtained from the primer extension approach, we carried out a Rapid Amplification of 5' cDNA End (5' RACE) assays. A cDNA strand was generated from the mRNA-anchor RNA fusion template (see Materials and methods) using the primer F276-1571. Subsequently, the 5' region of the intact mRNA was amplified by PCR using the 'nested primer' -36 (F1641-60) or -106 (F1711-30) and the 'anchor primer' (complementary to the T7 DNA that

A



B

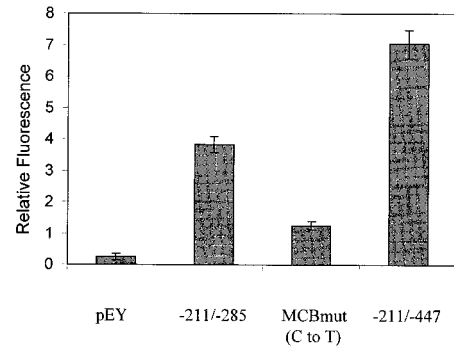


Figure 5 A point mutation within the putative MCB element dramatically reduced promoter activity. (a) A point mutation was introduced by substituting 'C' at -249 (underlined) to 'T' and examined its promoter activity. Arrowheads denote the transcription initiation sites that were mapped by both the primer extension and 5'RACE methods. (b) The promoter activity of this transition mutant (MCB_{Mut}) was compared to those of vector alone (pEY) and the wild-type $-211/-285$ construct. The promoter activity of the $-211/-447$ construct, which is thought to include most of the essential promoter elements, is also included as reference

was added at the 5' end of the full-length mRNAs). Similar experiments were also carried out using primers -402 and -796 . Finally, the nucleotide sequences of the PCR products were determined. As expected, the 5'RACE confirmed that transcription starts at position -220 , -235 or -245 . However, the transcription initiations at other sites mapped by primer extension were not reproducible by 5'RACE in spite of several repeats.

Organization of the HuDbf4/ASK gene

By comparing the nucleotide sequence of HuDbf4 cDNA and that of the genomic DNA reported by the two BAC clones, RG135C18 and RG060N22 (Figure 1a), we were able to identify the putative positions of exons and introns of the HuDbf4/ASK gene. To confirm this gene structure, we first analyzed the sizes of DNA fragments amplified by PCR using appropriate sets of primers and cDNA or genomic DNA templates as shown in Figure 8. Subsequently, part or the whole nucleotide sequence of each DNA fragment isolated from the gel was determined to confirm the exon-intron junctions. As summarized in Figure 9, huDbf4/ASK gene spans 33 kb and consists of 12 exons. The sizes of the HuDbf4/ASK exons are

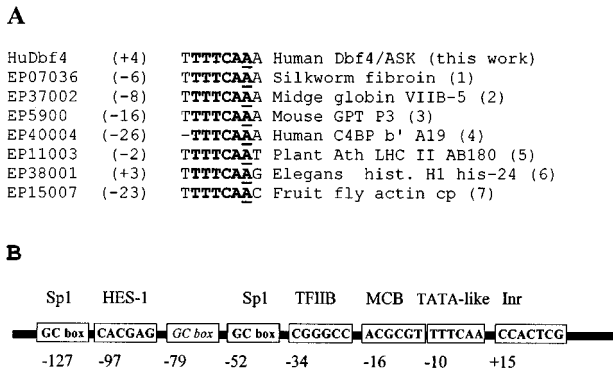


Figure 6 TATA-like sequence, TTTCAA, is often found in the vicinity of transcription initiation sites. (a) A promoter DataBank search showed that at least eight promoters contain a TATA-like, TTTCAA sequence near transcription initiation sites. The numbers in brackets are the relative positions of the TATA-like sequence to the transcription start-sites ('-' and '+' are the relative positions up- and down-stream of the transcription initiation sites from the underlined 'A', respectively). The promoter accession numbers were taken from EPD. The references are as follows: (1) Tsujimoto and Suzuki (1979); (2) Trewitt *et al.* (1988); (3) Rajput *et al.* (1994); (4) Hillarp *et al.* (1993); (5) Leutwiler *et al.* (1986); (6) Sanicola *et al.* (1990); (7) Couderc *et al.* (1987). (b) Organization of the HuDbf4/ASK core promoter is summarized. The numbers are relative positions to the major transcription start site at nt235 upstream of the ATG start-codon. The GC box at 79 (italic) may not function

generally small, ranging from only 36 (exon 7) to ~280 nucleotides (exon 1), except exon 12 (1188 nucleotides). Only a small part of exon 1 (45 nucleotides) encodes protein while 976 nucleotides of exon 12 is within the ORF. The actual size of exon 1 may vary depending on the transcription initiation site. Initiation at -504 (Kumagai *et al.*, 1999) would increase the size of first exon; however, it does not include another splicing site. The sizes of introns vary considerably ranging from 412 nucleotides (intron 9) to 8414 nucleotides (intron 6).

Discussion

We found that the DNA segment from 211 to 285 bp upstream (-211/-285) of the HuDbf4/ASK translation start-codon contains basal promoter activity (Figures 1c, 2a and 5). This 75 bp DNA segment contains putative Sp1 and TFIIB binding sites, an MCB, a TATA-like sequence, and a potential Inr element (Figures 5 and 6). A change of the 'C' residue within the putative MCB to 'T' nearly eliminated the promoter activity (Figure 5), suggesting that this element is essential for the HuDbf4/ASK core promoter activity. As far as we know, this is the first report that an MCB element plays such an important role for the activation of transcription in mammalian cells. In *S. cerevisiae*, an MCB element is often involved in the promoter activation of the genes that

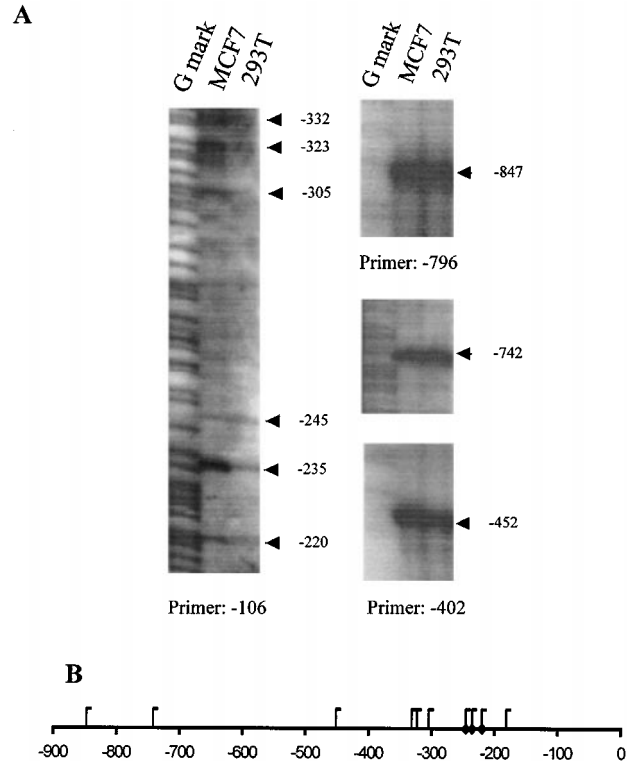


Figure 7 Transcription of the HuDbf4/ASK gene utilizes multiple initiation sites. (a) To determine transcription initiation sites, we used three different primers that are widely separated. The primer -106 is complementary to the nucleotide sequence from 106 to 125 upstream of the HuDbf4/ASK translation start-codon. The primers -402 and -796 are complementary to the sequences from -402 to -422 and from -796 to -815, respectively. MCF7 and 293T above each panel are mRNAs that are isolated from these cells, respectively. The transcription initiation sites are determined by comparing to the 'G' marker and the nucleotide sequence of the GenBank (Accession Number, AC003083). The nucleotide numbers expressed in negative ('-') are the relative upstream positions from the HuDbf4/ASK translation start-codon. (b) A schematic presentation of the transcription start-sites obtained by the primer extension method. The three initiation sites with diamonds were reproducible by 5'RACE while other initiation sites were not

are turned on at the G1/S transition, including Dbf4, Pol1, CDC9, TMP1, CLN1 and 2 (Masai and Arai, 2002; Gordon and Campbell, 1991; McIntosh, 1993; Verma *et al.*, 1991, 1992). Since the mammalian Dbf4 gene is also activated at the G1/S transition (Guo and Lee, 2001), our data raises the possibility that the MCB element in mammalian cells functions similarly to that in *S. cerevisiae*. The functionality of the Sp1 box 5 and TFIIB in the core promoter is yet to be experimentally confirmed.

The 237 bp DNA segment -211/-447 showed almost full promoter activity. An addition of 547 bp from -448 to -994 did not further increase promoter activity (Figure 2a). Therefore, -211/-447 may contain core elements, which are present within -211/-285, as well as essential auxiliary elements for the full promoter activity. Three putative Sp1 elements (boxes 2, 3, and 4) and one HES-1 binding

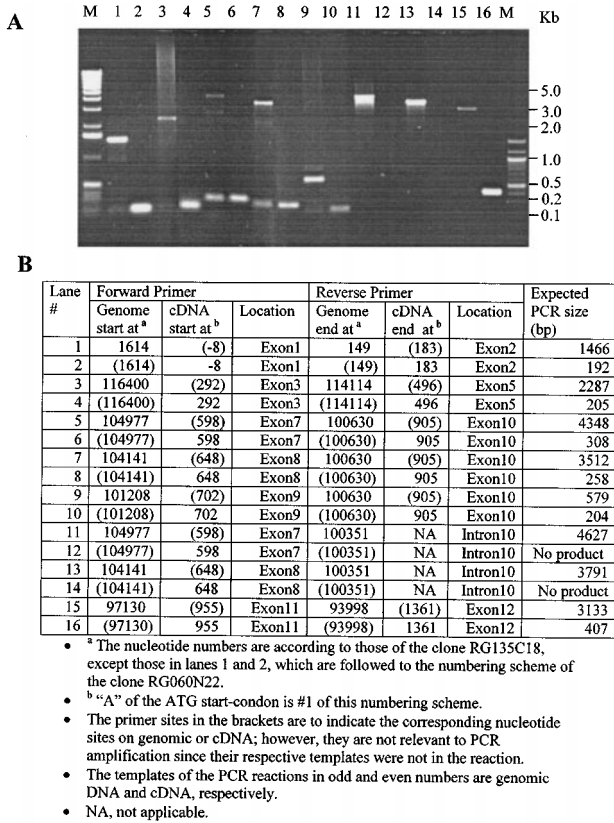


Figure 8 Mapping the exons and introns of HuDbf4/ASK. (a) The PCR products generated using the primer sets shown in panel b were separated by agarose gel electrophoresis. (b) Template for each PCR reaction was either genomic DNA (odd number lanes) or cDNA (even number lanes). The nucleotide numbers in brackets are the corresponding positions on either cDNA (odd number lanes) or genomic DNA (even number lanes)

site are present within the nucleotide sequence from -286 to -447, where the auxiliary elements are thought to be located. Theoretically, the presence of multiple Sp1 elements can further increase the basal promoter activity. However, we did not find any evidence showing that the Sp1 boxes 2 and 4 are bound by transcription factors (data not shown), suggesting that these putative elements are not functional. In contrast, our EMSA data showed that the Sp1 element at -353/-361 (i.e., box 3) is bound by the Sp1 transcription factor in a sequence-specific manner (Figure 4). Thus, only the box 3 is functional among the three putative Sp1 elements within -286 to -447. Consistent with this idea, the change of GGC to TTG within the Sp1 box 3 reduced the promoter activity to that of the basal level shown by the DNA segment -211/-285 (-211/-447 Sp1mu in Figure 3b vs -211/-285 in Figure 5b). Taken together, the Sp1 element at -353/-361 appears to be the critical auxiliary element that is required for an efficient promoter activity.

The deletion of the putative HES-1 and the putative Sp1 Box 4 dramatically increased the promoter activity (Figure 3b). This result suggests that the putative HES-

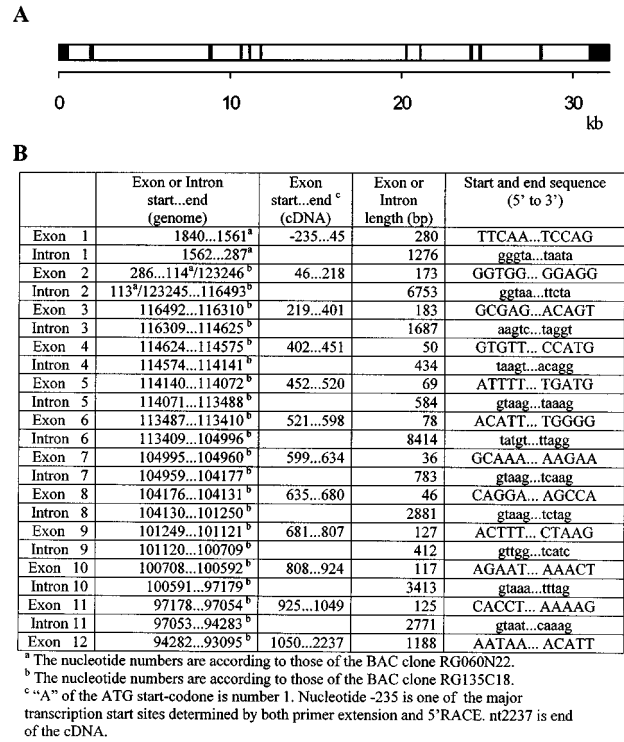


Figure 9 HuDbf4/ASK gene structure. (a) Filled and open boxes are the positions of exons and introns, respectively. (b) The sizes of exons and introns as well as their boundaries are shown

1 repressor element is functional but, again, the putative Sp1 Box 4 is not functional. The HES-1 was previously shown to be involved in promoter repressing in human and rat (Gordon and Campbell, 1991; McIntosh, 1993; Verma *et al.*, 1991). It is interesting to note that the Sp1 Box 3, which is required for efficient promoter activity, is located within 30 bp from the putative HES-1 repressor element. It seems possible that a further increase or decrease of the basal promoter activity is regulated through a delicate coordination of the Sp1 activator element at -353/-361 and the HES-1 repressor element at -326/-331.

Although almost full promoter activity was observed with the 237 bp DNA segment -211/-447, the complete promoter activity was seen with the DNA segment -170/-1628 (Figures 1c and 2a). Since no promoter activity was found within the -144/-215 DNA segment (Figure 2a), all the essential promoter elements including enhancer are likely present between -217 and -1628. The promoter activity was significantly decreased when the DNA between -994 and -1628 was removed (i.e., -211/-993 vs -170/-1628 in Figure 2a). However, no DNA fragment within the -516 to -3405 segment showed promoter activity (Figure 1c). Therefore, the 865 bp DNA segment from -994 to -1628 is likely to contain elements that can enhance promoter activity, but not the core elements. A computer-based DataBank search suggests that there are several potential transcription

activator elements within this putative 865-bp promoter enhancer segment, including one each of IUF-1, c-Jun, GATA-1, GCN4, PEA3, NF-1, Zeste, TGT3, and two Pit-1a as well as three TFIID recognition sites (Figures 2b, c). Further studies are required to confirm as to whether any or all of these putative transcription elements function as a transcription enhancer.

The DNA segment between -448 and -993 contains putative binding sites for E2F-1, OCT-1, and Sp1 (Figures 2b, c). However, this DNA segment does not appear to increase or decrease the promoter activity, since the constructs -211/-447 and -211/-993 showed essentially the same promoter activity (Figure 2a). Nevertheless, we cannot rule out if these putative transcription elements within this DNA segment are involved in other regulations such as cell-cycle related control mechanisms *in vivo*.

Transcription initiations at nt220, nt235, and nt245 upstream of the translation start-codon were confirmed by two independent mapping approaches, the primer extension and 5'RACE methods (Figure 7 and data not shown). Among these three sites, the transcription initiation at -235 appears to be the strongest one (Figure 7). The 5' end of the HuDbf4 cDNA reported by Jiang *et al.* (GenBank Accession Number AF160249) is -217, while that reported by Hollingsworth (GenBank Accession Number AF160876) is -199. These sequence data, therefore, are in line with our observation that the major transcription initiation occurs within -220 to -245. However, the 5' end of the cDNA sequence reported by Kumagai *et al.*, 1999 (GenBank Accession Number AB028069) is -504, which is located further upstream of the major initiation sites that we have mapped by 5'RACE. Nevertheless, the data obtained from our primer extension experiments using three different primers suggests that transcription initiation also occurs at -305, -323, -332, -452, -742 and -847 (Figure 7). Thus, the cDNA reported by Kumagai *et al.* (1999) is in line with some of the transcription initiation sites mapped by the primer extension approach. However, in spite of many repeats, we could not reproduce these initiation sites by the 5'RACE mapping method. Since we also did not find core promoter activity further upstream of nt279 (Figure 1c and data not shown), transcription initiation may occur only infrequently further upstream of -245.

As summarized in Figure 6b, the distribution of transcription regulatory elements in the HuDbf4 core promoter is similar to that of a typical TATA-less promoter (Carey and Smale, 2000). However, there is a TATA-like sequence at 21 bp upstream of the Inr, which may be relevant to the fact that the three major transcription initiation sites are concentrated within a 25 bp region from the -220 to -245 DNA segment.

The presence of MCB in the core promoter region is extremely interesting as the element plays an important regulatory function for the activation of the genes that are involved in the G1/S transition in yeast (Gordon and Campbell, 1991; McIntosh, 1993; Verma *et al.*, 1991). Given the important role of Dbf4/ASK in the

activation of DNA replication in mammalian cells, further studies on the roles of MCB, Sp1, and HES-1 elements of this gene may shed new light on the G1/S and S phase control mechanism in mammalian cells.

Materials and methods

Cell culture

The adenovirus-transformed human embryonic kidney cell line 293T and human breast cancer cell lines MCF7 and MDA-MB231 were cultured in a 37°C incubator with 5% CO₂ in Dulbecco's modified Eagle's minimal essential medium supplemented with 10% heat-inactivated fetal bovine serum (Sigma).

Database searching

Databank search was carried out using the BLAST program (Altschul *et al.*, 1997) through the NCBI web site (www.ncbi.nlm.nih.gov). The putative transcription factor binding sites were identified using the Searching Transcription Factor Binding Sites (TFSEARCH) software Version 1.3 (Heinemeyer *et al.*, 1998) (www.cbrc.jp/research/db/TFSEARCH) and the Transcription Element Search System (www.cbil.upenn.edu/tess). Searching the Eukaryotic Promoter Database (EPD) (Perier *et al.*, 2000) was carried out using the search engine at www.epd.isbsib.ch.

Transfection, fluorescence microscopy and flow cytometry

Lipofectamine (Invitrogen)-based transfection of MCF7, MDA-MB231, and 293T cells in a 10-cm dish was carried out as described previously (Guo *et al.*, 2001). A calcium phosphate precipitation-based transfection using a 96-well plate was performed as follows: An appropriate amount of cells was seeded one day prior to transfection so that cell density at the time of experiment could be ~70% confluent. To prepare transfection, 5 µl of plasmid DNA (0.09 pmol DNA in 293 mM CaCl₂) was mixed with an equal volume of 2× HEPES buffer (2× buffer is 54.5 mM HEPES, 274 mM NaCl, 1.5 mM Na₂HPO₄, pH 7.0). The DNA preparation (10 µl) was then slowly added to each well that contains 100 µl of complete medium. The transfection mixture was replaced with fresh medium 24 h after transfection. Fluorescence microscopy was carried out using a Zeiss Axiovert 100 inverted microscope with Chroma Endow GFP Bandpass (exciter HQ470/40x and emitter HQ525/50m) 48 h after transfection.

To prepare samples for flow cytometric analysis, the transfected cells in a 10-cm dish were trypsinized, spun down, and then suspended in 1×PBS. For each sample, approximately 3×10⁴ cells were analysed using a Beckman Coulter Epics[®] Elite flow cytometer. Fluorescence intensity upon stimulation with an argon-ion laser at 488 nm was measured between 500 and 540 nm and plotted against cell number. The fluorescence emissions of mock-transfected and plasmid-transfected cells were compared using the Phoenix Flow System's Multiplus AV software (San Diego, CA, USA). The relative promoter activity was determined by measuring the fluorescence intensity captured on the x-axis as described in the legend to Figure 1b.

Plasmid constructs

The promoterless pEY reporter vector was constructed by removing the CMV promoter-containing DNA segment

between the *AseI* and *NheI* sites from pEYFP-N1 (Clontech, Palo Alto, CA, USA). The 136 kb BAC clone RG060N22 (GenBank Accession Number 003083), which contains part of huDbf4/ASK coding and the upstream noncoding regions, was purchased from Research Genetics (Huntsville, AL, USA). To make genomic DNA libraries containing a potential HuDbf4/ASK transcription promoter, the 136 kb BAC clone was partially digested with *BamHI* and *EcoRI* and then the resultant DNA fragments were ligated with the pEY vector that was digested with either *BamHI*, *EcoRI*, or *BamHI/EcoRI*. Subsequently, the hybrid DNAs were transformed into DH5 α . To subclone the positive BE5 (–170/–4919), this 4.8 kb DNA fragment was partially digested with *Sau3A* and the subfragments were cloned into the *BamHI* site of pEY. Alternatively, subfragments were generated by PCR using the primer pairs listed in Table 1 and cloned into the *BamHI*–*EcoRI* site of pEY.

Isolation of total RNA and DNA

Total RNA was isolated from 293T, MCF7 or MDA-MB231 cells using Tri-Reagent (Molecular Research Center, Cincinnati, OH, USA), which is based on the acid guanidinium thiocyanate-phenol/chloroform extraction method originally described by Chomczynski and Sacchi (1987). Isolation of genomic or BAC DNA was according to a standard protocol (Sambrook and Russell, 2001).

Primer extension

The three primers used for the primer extension experiment were F1711 or –106 (5'-CTTCCGCCAGCTACGGCCTC-3'), –402 (5'-AGGTTGTGTTTCCGCCTCTAC-3'), and –796 (5'-CAGCAAGCGGGGTTTTCCCG-3'), which are complementary to the nucleotide sequences between 106 to 125, 402 to 422, and 796 to 815, respectively, upstream of the HuDbf4/ASK translation start-codon. The primers were labeled at their 5' ends with digoxigenin (Integrated DNA Technologies, Coralville, IA, USA). For each primer extension experiment, 20 μ g of total RNA isolated from 293T, MCF7 or MB231 cells was mixed with digoxigenin labeled primers and incubated at 70°C for 10 min in 30 μ l 'First-strand' buffer (50 mM Tris-HCl (pH 8.3), 75 mM KCl, and 15 mM MgCl₂). A first cDNA strand was generated at 43°C using SuperScript II reverse transcriptase (Invitrogen, Burlington, ON, Canada) according to the supplier's specifications. Subsequently, the resultant mRNA-cDNA hybrid was denatured by incubating it at 90°C for 3 min. The sample was cooled in ice, mixed with 20 μ l of loading buffer (98% formamide, 10 mM EDTA, 0.1% xylene cyanol, 0.1% bromophenol blue), and then separated by sequencing gel electrophoresis (5% polyacrylamide-7.7 M urea in 0.5 \times TBE). The labeled cDNA was then transferred onto a Hybond N⁺ membrane (Amersham Pharmacia Biotech, Baie d'Urfe, QC, Canada) using a semidry sample transfer apparatus (Bio-Rad, Mississauga, ON, Canada). Finally, DNA bands were detected using a DIG Luminescence Nucleic Acid Detection Kit (Roche Diagnostics, Laval, QC, Canada). The 'G' size marker was prepared by the digoxigenin 5' end-labelling and dimethyl sulphate (DMS) treatment followed by piperidine cleavage as described by Maxam and Gilbert (1980).

Rapid amplification of 5' cDNA end (5' RACE)

Messenger RNA was extracted using the Oligotex mRNA spin-column (Qiagen, Mississauga, ON, Canada). The 5'

RACE assay was performed as described by Shaefer (1995). Briefly, mRNA without a 5' cap was dephosphorylated using calf intestinal phosphatase to inhibit ligation of degraded mRNA molecules in the subsequent 5' RACE procedure. The full-length mRNA was decapped using tobacco acid pyrophosphatase (Epicentre Technologies, Madison, WI, USA), resulting in the exposure of a phosphate group at the 5' end of the full-length mRNA molecule. The decapped mRNA was then ligated using T4 RNA ligase to a 21-nucleotide anchor RNA (5'-GGGCGAAUUGGAGCUC-CACCG-3') (Integrated DNA Technology, Coralville, IA, USA). A first strand cDNA was generated by reverse transcription using the anchor RNA-mRNA template and the primer F276-1571 (5'-TGGATTCCACCCTGGAAATGTC-3') that spans nucleotides 35–45/1319–1329 upstream of the translation start-codon, where the slice, '/', is the location of an intron. Since the nucleotide sequence of this primer is derived from two exons flanking a large intron from nt46 to nt1318, it can reduce unwanted amplification of the genomic DNA contaminant in the mRNA preparation. The 5' region was amplified by PCR using the anchor primer (5'-GGGCGAATTGGAGCTCCACCG-3') and gene-specific primers from the first strand cDNA template. To increase specificity, four different 'nested' gene-specific primers were used: A primer F1640-60 or –36 (5'-AAAGTGTCGCACCGGCAGC-3' at –36 to –55) and three primers, –106, –402, and –796. Subsequently, the nucleotide sequences of the PCR products were determined after they were isolated from agarose gel.

Mutagenesis

Generation of point and deletion mutants were carried out by PCR as described previously (Ho *et al.*, 1989; Tomic-Canic *et al.*, 1996). Three mutations that we examined for their promoter activities are: (1) Sp1_{mut} which was altered –359GGC-357 to TTG within the Sp1 Box3; (2) a 51 bp deletion from –302 to –352; (3) alteration of 'C' to 'T' at –249 within the MCB element. All the mutants generated in our experiments were verified by nucleotide sequencing.

Electrophoretic mobility shift assay (EMSA)

EMSA was performed using an Sp1 Nushift kit (Geneka Biotechnology, Montreal, QC, Canada) according to a protocol provided by the supplier. The 25-bp long DNA probe, F1952-75 (i.e., –347 to –370) was made by hybridization of two synthetic complementary strands. The double strand probe was then end-labeled with γ -³²P-ATP (3000 Ci/mmol, MEN Life Science Products, Boston, MA, USA) using T4 polynucleotide kinase (United States Biochemicals, Cleveland, OH, USA). In the competition experiments, the unlabeled oligo that contains either a 'wild-type' Sp1 binding site (5'-CCCTTTGTGGGGCGGGCCTAAGCTGCG-3'), an altered Sp1 binding site (5'-CCC-TTTGTGGGttgGGGCCTAAGCTGCG-3', where the lower case is mutated sequence), or Sp1 peptide (CKDS-EGRGSDPGKK, Geneka, Biotechnology, Montreal, QC, Canada) was added to the HeLa cell nuclear extracts and incubated for 20 min at 4°C prior to addition of the radiolabeled probes. In the 'supershift' experiments, rabbit polyclonal antibody against Sp1 (Geneka Biotechnology, Montreal, QC, Canada) was added to the reaction mixture and incubated at 4°C for 20 min prior to addition of the labeled probes. The resultant DNA-protein complexes were separated by 4.5% polyacrylamide gel electrophoresis.

Acknowledgements

We thank the members of the Lee lab, particularly, B Guo and A Pearce for many helpful discussions. This work was supported by the operating grants from the Canadian

References

- Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W and Lipman DJ. (1997). *Nucleic Acids Res.*, **25**, 3389–3402.
- Bousset K and Diffley JF. (1998). *Genes Dev.*, **12**, 480–490.
- Carey M and Smale ST. (2000). *Transcriptional Regulation in Eukaryotes*. Cold Spring Harbor Laboratory Press: Cold Spring Harbor, New York.
- Chapman JW and Johnston LH. (1989). *Exp. Cell Res.*, **180**, 419–428.
- Chomczynski P and Sacchi N. (1987). *Anal. Biochem.*, **162**, 156–159.
- Couderc JL, Hilal L, Sobrier ML and Dastugue B. (1987). *Nucleic Acids Res.*, **15**, 2549–2561.
- Donaldson AD, Fangman WL and Brewer BJ. (1998). *Genes Dev.*, **12**, 491–501.
- Ferreira MF, Santocanale C, Drury LS and Diffley JF. (2000). *Mol. Cell Biol.*, **20**, 242–248.
- Gordon CB and Campbell JL. (1991). *Proc. Natl. Acad. Sci. USA*, **88**, 6058–6062.
- Guo B and Lee H. (2001). *Gene*, **264**, 249–256.
- Guo B, Pearce AG, Traulsen KE, Rintala AC and Lee H. (2001). *Biotechniques*, **31**, 314–321.
- Heinemeyer T, Wingender E, Reuter I, Hermjakob H, Kel AE, Kel OV, Ignatieva EV, Ananko EA, Podkolodnaya OA, Kolpakov FA, Podkolodny NL and Kolchanov NA. (1998). *Nucleic Acids Res.*, **26**, 362–367.
- Hillarp A, Pardo-Manuel F, Ruiz RR, Rodriguez de Cordoba S and Dahlback B. (1993). *J. Biol. Chem.*, **268**, 15017–15023.
- Ho SN, Hunt HD, Horton RM, Pullen JK and Pease LR. (1989). *Gene*, **77**, 51–59.
- Jackson AL, Pahl PM, Harrison K, Rosamond J and Sclafani RA. (1993). *Mol. Cell Biol.*, **13**, 2899–2908.
- Jares P and Blow JJ. (2000). *Genes Dev.*, **14**, 1528–1540.
- Jiang W, McDonald D, Hope TJ and Hunter T. (1999). *EMBO J.*, **18**, 5703–5713.
- Kumagai H, Sato N, Yamada M, Mahony D, Seghezzi W, Lees E, Arai K and Masai H. (1999). *Mol. Cell Biol.*, **19**, 5083–5095.
- Larner JM, Lee H, Little RD, Dijkwel PA, Schildkraut CL and Hamlin JL. (1999). *Nucleic Acids Res.*, **27**, 803–809.
- Lee H, Larner JM and Hamlin JL. (1997). *Proc. Natl. Acad. Sci. USA*, **94**, 526–531.
- Lepke M, Putter V, Staib C, Kneissl M, Berger C, Hoehn K, Nanda I, Schmid M and Grummt F. (1999). *Mol. Gen. Genet.*, **262**, 220–229.
- Leutwiler LS, Meyerowitz EM and Tobin EM. (1986). *Nucleic Acids Res.*, **14**, 4051–4064.
- Institutes of Health Research (MT-15015), Northern Cancer Research Foundation, and the Northern Ontario Heritage Fund Corporation (NOHFC #50070). H Lee is a recipient of the Premier's Research Excellence Award.
- Masai H and Arai K-I. (2002). *J. Cell. Physiol.*, **190**, 287–296.
- Masai H, Matsui E, You Z, Ishimi Y, Tamai K and Arai K. (2000). *J. Biol. Chem.*, **275**, 29042–29052.
- Masai H, Sato N, Takeda T and Arai K. (1999). *Front Biosci.*, **4**, D834–D840.
- Maxam AM and Gilbert W. (1980). *Methods Enzymol.*, **65**, 499–560.
- McIntosh EM. (1993). *Curr. Genet.*, **24**, 185–192.
- Mimura S and Takisawa H. (1998). *EMBO J.*, **17**, 5699–5707.
- Nougarede R, Della SF, Zarzov P and Schwob E. (2000). *Mol. Cell Biol.*, **20**, 3795–3806.
- Oshiro G, Owens JC, Shellman Y, Sclafani RA and Li JJ. (1999). *Mol. Cell Biol.*, **19**, 4888–4896.
- Owens JC, Detweiler CS and Li JJ. (1997). *Proc. Natl. Acad. Sci. USA*, **94**, 12521–12526.
- Paulovich AG and Hartwell LH. (1995). *Cell*, **82**, 841–847.
- Perier RC, Praz V, Junier T, Bonnard C and Bucher P. (2000). *Nucleic Acids Res.*, **28**, 302–303.
- Rajput B, Ma J and Vijay IK. (1994). *J. Biol. Chem.*, **269**, 9590–9597.
- Sambrook J and Russell DW. (2001). *Molecular Cloning: A Laboratory Manual*. (3rd ed). Cold Spring Harbor Laboratory Press: Cold Spring Harbor, New York.
- Sanicola M, Ward S, Childs G and Emmons SW. (1990). *J. Mol. Biol.*, **212**, 259–268.
- Schaefer BC. (1995). *Anal. Biochem.*, **227**, 255–273.
- Siede W, Friedberg AS, Dianova I and Friedberg EC. (1994). *Genetics*, **138**, 271–281.
- Tercero JA, Labib K and Diffley JF. (2000). *EMBO J.*, **19**, 2082–2093.
- Tomic-Canic M, Sunjevaric I and Blumenberg M. (1996). *Methods Mol. Biol.*, **57**, 259–267.
- Treuner K, Eckerich C and Knippers R. (1998). *J. Biol. Chem.*, **273**, 31744–31750.
- Trewitt PM, Saffarini DA and Bergtrom G. (1988). *Gene*, **69**, 91–100.
- Tsujimoto Y and Susuki Y. (1979). *Cell*, **16**, 425–436.
- Verma R, Patapoutian A, Gordon CB and Campbell JL. (1991). *Proc. Natl. Acad. Sci. USA*, **88**, 7155–7159.
- Verma R, Smiley J, Andrews B and Campbell JL. (1992). *Proc. Natl. Acad. Sci. USA*, **89**, 9479–9483.
- Weinreich M and Stillman B. (1999). *EMBO J.*, **18**, 5334–5346.
- Zou L and Stillman B. (1998). *Science*, **280**, 593–596.
- Zou L and Stillman B. (2000). *Mol. Cell Biol.*, **20**, 3086–3096.