FEBS Letters 435 (1998) 138–142 FEBS 20783

Genomic structure and promoter analysis of the p62 gene encoding a non-proteasomal multiubiquitin chain binding protein

Ratna K. Vadlamudi[1], Jaekyoon Shin*

Division of Tumor Virology, Dana-Farber Cancer Institute, Harvard Medical School, 44 Binney Street, Boston, MA 02115, USA

Received 3 June 1998; revised version received 3 August 1998

Abstract p62 is a novel immediate early response gene encoding a ubiquitin chain binding protein. To investigate the mechanism of p62 gene expression, we isolated and characterized the 20 kb long human p62 gene. The p62 gene contains seven introns and eight exons. The splice sites conformed to the GT/AG rule, except introns 6 and 7 which used the unusual GC dinucleotides. The p62 promoter is TATA-less, and 357 nucleotides of the 5′-flanking region contain basic machineries for transcription. A reporter gene linked to 1800 nucleotides of the 5′-flanking region was rapidly activated by various extracellular signals. The presence of a CpG island as well as multiple binding sites for SP-1, AP-1, NF-κB, and Ets-1 family in the promoter region supports the regulated activation of the p62 gene.

© 1998 Federation of European Biochemical Societies.

Key words: p62; Ubiquitin chain; TATA-less; CpG island

1. Introduction

p62 is a cytoplasmic protein originally identified as a phosphotyrosine-independent ligand of the SH2 domain of p56lck [1]. p62 is expressed in most tissues and also binds to ubiquitin and atypical PKC ζ [2,3]. Ubiquitin chain conjugation is critical in regulation of cell proliferation by controlling catabolism of key regulatory proteins at the proteasome or the endosome/lysosome [4,5]. Proteasomal recognition of substrates is initiated by a proteasomal multiubiquitin chain binding protein(s) [6,7]. Interestingly, p62 also has high affinity to multiubiquitin chains (Ko et al., submitted). The molecules to which it binds suggest that p62 is likely involved in a regulation of ubiquitinated protein metabolism.

Recently, we also found that a variety of signals, including those which stimulate proliferation and differentiation, rapidly induces the transcriptional activation of the p62 gene (Lee et al., submitted). Interestingly, the mechanism by which p62 is activated is similar to that of immediate early response genes. Furthermore, a murine p62 homolog A170 gene is activated by oxidative stress in murine intestinal macrophage [8]. Thus, analysis of the genetic regulation of p62 expression would be important for extracellular signal-mediated protein metabolism. To this end, in the present study, we have analyzed the 20 kb long human p62 gene locus, including its genomic organization and promoter analysis. These results will provide the way for further analysis of p62 gene regulation.

*Corresponding author. Fax: (1) (617) 632-2662.
E-mail: shin@mbcrr.harvard.edu

[1]Present address: Cell Growth Regulation Section, University of Texas, M.D. Anderson Cancer Center, Houston, TX, USA.

2. Materials and methods

2.1. Genomic library screening

10^6 plaques of a Lambda FixII genomic library of human placenta (Stratagene, La Jolla, CA) were screened using 5′-end labeled full length p62 cDNA as a probe. Hybridization was carried out at 42°C in formamide buffer [9]. Two positive clones, p62-λ1 and p62-λ2, were hybridized by the 5′ and 3′ halves of p62 cDNA respectively and were further characterized. Restriction digested fragments of the purified λ DNAs were subcloned into pBluescript and analyzed. The 5′-flanking region, exons and their boundaries were directly sequenced on both strands from the clones using Sequenase kit (USB, Cleveland, OH). Introns were amplified from the genomic clones using a long-range PCR kit (Promega, Madison, WI) and primers derived from within the adjacent exons.

2.2. Primer extension

Poly(A) mRNA was isolated from HeLa cells with mRNA separator kit (Clontech, Palo Alto, CA). An antisense oligonucleotide primer 5′-ggtcccgccggcactctttt-3′ derived from the region +326 to +305 bases to the translation initiation site of p62 cDNA [1] was synthesized and 5′-end labeled with [γ-^{32}P]ATP using T4 polynucleotide kinase. 10 pmol of the end labeled primer was annealed to 5 μg of poly(A) mRNA at 52°C. Extension was performed at 42°C for 1 h with 10 units of AMV reverse transcriptase (Promega) and analyzed on a 5% polyacrylamide gel.

2.3. Reporter plasmid construction and expression

The 5′-flanking region was isolated by *Bam*HI-*Bsi*EI digestion from the 6 kb *Bam*HI fragment of p62λ1. These enzymes liberate a 1818 bp long fragment from the 5′ cloning site (*Bam*HI) to 10 bp downstream of the transcription start site (*Bsi*EI). Both ends of the fragment were blunt ended using T4 polymerase then ligated to upstream of luciferase gene in PGL3 basic vector (Promega) (PGL3.p62:−1808/+10). A construct containing the same fragment with inverse orientation (PGL3.p62:+10/−1808) was also isolated and used as a negative control. PGL3.p62:−1408/+10, PGL3.p62:−873/+10, and PGL3.p62:−357/+10 were similarly constructed by ligating restriction fragments of *Stu*I-*Bsi*EI, *Bst*EII-*Bsi*EI, and *Kpn*I-*Bsi*EI digestion respectively. 6×10^5 NIH3T3 fibroblast cells were transfected using lipofectin (Gibco-BRL, Gaithersburg, MD) with 3 μg each of the various reporter constructs. At 48 h post transfection, cells were washed with PBS, and luciferase activity contained in the cells was measured by following the procedure described by Williams et al. [10]. To measure serum-induced activation, NIH3T3 cells transfected with PGL3.p62:−1808/+10 were serum starved for 48 h, incubated further in the presence of 15% fetal bovine serum for the indicated time, and luciferase activity contained in the cells was measured. Transfection efficiency was normalized by β-galactosidase activity which was expressed by cotransfection with a plasmid containing CMV promoter-linked β-galactosidase gene (pCMV.β-gal). β-Galactosidase activity was measured using Luminescent β-gal detection kit (Clontech, Palo Alto, CA).

3. Results and discussion

3.1. Genomic organization of the p62 gene

p62 cDNA was used to screen 10^6 plaques of a Lambda FixII genomic library of human placenta (Stratagene). Two overlapping positive clones (p62-λ1 and p62-λ2) cover full

0014-5793/98/$19.00 © 1998 Federation of European Biochemical Societies. All rights reserved.
PII: S0014-5793(98)01021-7

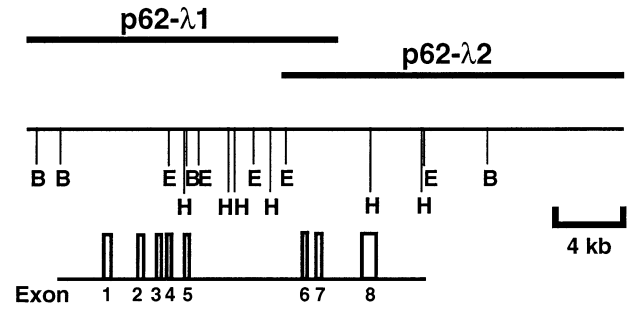

Fig. 1. Schematic representation of p62 genomic organization. Structure of the p62 gene was deduced from analysis of two overlapping clones p62-λ1 and p62-λ2. The confirmed restriction enzyme sites are shown for *Bam*HI (B), *Eco*RI (E) and *Hin*dIII (H). Eight exons are shown at the bottom with open rectangles.

length p62 cDNA (Fig. 1). The p62 gene contains eight exons ranging in size from 81 bp to 872 bp. The exon sequence is identical to the previously reported p62 cDNA sequence [1]. Exon-1 contains the ATG translation initiation codon and the 5′-untranslated region (UTR), while exon 8 has stop codon and 3′-UTR (Fig. 1 and Table 1). The sizes of introns range from 200 bp to 8 kb.

All intron and exon splice sites follow the AG/GT rule [11] except introns 6 and 7 which have GC in the 5′ splice donor site (Table 1). Although rare, this functional splice sequence has been previously noted in several genes [12] and implicated for alternative splicing [13]. However, the alternatively spliced p62 message was not detected in any tissues [1].

3.2. Transcription start site

The transcription start site of p62 gene was estimated using primer extension analysis with an oligonucleotide complementary to the region +326/+305 bp of the translation initiation site of p62 cDNA [1]. As shown in Fig. 2, primer extension yielded a major band running between 311 nt and 427 nt size markers (lane 2), while no extension products were observed when mRNA was omitted (lane 1) or substituted with yeast transfer RNA (lane 3). Based on relative mobility measured in three independent experiments, the size of the extended primer was estimated to be about 380 nt. This result indicates that the transcription start site of the p62 gene is likely located between −50 nt and −60 nt from the ATG translation initiation site. In the previous report, HeLa p62 cDNA started with a GCGCGGCGGCTG sequence which is 52 nt upstream of the ATG translation initiation site [1]. Furthermore, 5′-RACE using HeLa poly(A) mRNA also resulted in the same 5′-end sequence (data not shown). Thus, it is tentatively concluded that the transcription start site of the p62 gene is likely located at −52 nucleotides from the translation initiation site (Fig. 3).

3.3. Characterization of the 5′-flanking region of p62 gene

A 6 kb *Bam*HI fragment of p62-λ1 covers the 1.8 kb long 5′-flanking region in addition to the first five exons of p62 gene (Fig. 1). Interestingly, the 5′-flanking region of the p62 gene lacks both TATA and CAAT boxes (Fig. 3). Instead, immediately upstream (within 100 bp 5′) of the transcription start site there is a very G/C rich region which contains three consensus sequences for SP1 transcription factor binding. Sp1 is a ubiquitous zinc finger containing transcription factor, and, in many cases, responsible for transcription of genes containing TATA-less promoters [14].

In order to analyze the functional promoter activity of the cloned 5′-flanking region of the p62 gene, various lengths of the 5′-flanking region were subcloned into PGL3 basic vector upstream of the luciferase gene. NIH3T3 cells were transfected with these constructs and activation of the luciferase gene was measured. Significant increase of luciferase activity was observed in cells transfected with PGL3.p62[−357/+10] (Fig. 4A). Transfection of PGL3.p62[−873/+10] induced luciferase activity similar to that of PGL3.p62[−357/+10]. However, addition of 535 bp (PGL3.p62[−1408/+10]) or 935 bp (PGL3.p62[−1808/+10]) further upstream induced much increased luciferase activity, fourfold higher than that by PGL3.p62[−357/+10] or PGL3.p62[−873/+10]. Under the same condition, luciferase activity was minimal in cells transfected with PGL3 basic vector or with PGL3.p62[+10/−1808] in which the 5′-flanking region of p62 is inversely linked to the luciferase gene. These results suggest that the basic machinery for p62 gene transcription lies within 357 nucleotides 5′ of the transcription start site, and that a strong enhancer element is contained in the region −873 to −1408.

3.4. The p62 promoter contains elements for the induced expression

Highly elevated transcript and protein levels of p62 have been observed during G0 exit in peripheral blood mononucleocytes and serum-starved NIH3T3 fibroblasts, during early G1 progression in a pre-B cell line Ba/F3, during monocytic differentiation of a promyelocyte U937 (Lee et al., submitted), and upon oxidative stress in murine intestinal macrophages [8].

Thus, signal-induced activation of p62 promoter was examined. NIH3T3 cells transfected with PGL3.p62[−1808/+10] were incubated in the presence or absence of 10% fetal bovine serum (FBS) for 48 h. About fourfold higher luciferase activity driven by the p62 promoter was observed in the cells

Table 1
Exon/intron organization of the p62 gene

Exon number	Exon length (nt)	5′-Splice donor (nt)	Intron length	3′-Splice donor
1	286	ACGGGG/**gt**gagc	1500	atctag/ACTTGG
2	84	TTAAAG/**gt**gagc	900	ctgtag/AGAAAA
3	230	TCTGAG/**gt**gagc	200	ccccag/GGCTTC
4	142	AATCAG/**gt**gagg	800	tcaaag/CTTCTG
5	81	CTCTGG/**gt**gatg	8000	ctctag/GCATTG
6	216	CCTGAG/**gc**aagc	500	ttccag/GAACAG
7	195	CGCCAG/**gc**aagt	2200	cggcag/AGGCTG
8	872			

Exon sequences are in uppercase letters and intron sequences in lowercase.

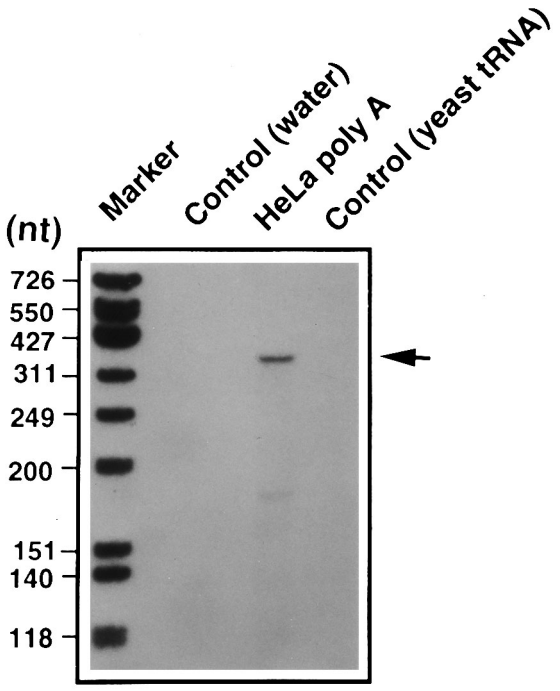

Fig. 2. Transcription initiation site analysis by primer extension. A primer corresponding to the +326 to +305 bp region of p62 cDNA was annealed to 5 μg of HeLa poly(A) mRNA, and extended by AMV reverse transcriptase (lane 3). As negative controls, the same primer extension was performed either in the presence of 5 μg yeast tRNA instead of HeLa poly(A) mRNA (lane 4) or in the absence of any RNA (lane 2). Molecular size markers in nucleotides (nt) are shown on the left (lane 1).
←

cultured in the presence of 10% FBS compared to that in the serum-starved cells (Fig. 4B). Furthermore, when the serum-starved cells were stimulated with 15% FBS, luciferase activity was rapidly increased to submaximal level 1 h after serum treatment (Fig. 4C). The induced activity was decreased to some extent for the next 5 h, increased again to 2.5-fold by 9 h post stimulation, and then maintained the activity for the next 15 h. On the other hand, without serum stimulation, the luciferase activity was maintained at relatively low, albeit significant, levels.

In U937 cells, under normal growing condition, luciferase activity expressed by PGL3.p62[−1808/+10] maintained its relatively low level (Fig. 4D). Upon PMA treatment, an increase of luciferase activity was observed in 1 h and a sharp surge of the activity followed at 3–4 h post stimulation. After

```
-1808 ccctattacg acagcggtca tgggacgctg actcactgcc gccagaccac ctgacctccg cggcgggagg agagggccct gccagggggt tcccgccgcg
                                  AP-1 / TRE                                                                NFkB
-1708 ccttgtttac cctccgggag gcctcggcct cgcgtggggg cagggcggcc gctggggccg caaggcgtgc ggggaagggc cagagccggg tgtccacccc

-1608 agcttcccaa agactccctc ttctgtgctt ccttctcccc tccccgcccc cccccagtct cttcacatgc ccctagcccc cgcggaaact tcccgcgatc
                                  PEA3              SP-1
-1508 ccaaacgggc caaatggcga gaaaagcaaa ggagctcctt cttgggggtg agtggggcgc cttgagcgct tcctcaaagc tatgttccca gagccacagg
                                                                                  PEA3 / Pu.1
-1408 ccttccttgt gtccctcacc ctgctcagac cgggccatag ccgggggctg gggcaggaaa gccggccctc ggcgggggcc acgtggctct caggcgcctg
       PEA3                                                                                       c-myc
-1308 ggctgctgag tcacgcttgg ccagcacctg tctgtaggcc acagcctctg ccagcacgcc cctctgtgtc ccctgcccct gtctgcaagg cagtggctcc
          AP-1/ TRE
-1208 agcaggccct ggggcatttt ccactctcca ccgccggatg cagggagagg cctgaaccct ctccacaggg ctgctctggg cagggtggaa gccttgccca
                                              ETS-1
-1108 cttcggagcc ctccgggaag gatcattcac acctgtggac cagcccctgc tgtgcgcaca cccacacatc accttcgcac ctgactggcc ccatccagcc

-1008 actctggcct ccctctgggt ttcctccct gggaggtttc tccagctcct gcaagccctg ggctgaaatg gcatgagttg gacccagcag gttctgacct
                                  Pu.1
 -908 cctactcaca ggaccttgcc tgggaggctc cagagggtga ccactcgtcc tgcccctcgt tccttgcccc agttctggcg gacaggttac tctggtggca

 -808 taaagcagtg tttcttcctt cctagctgaa ggaggaggct gttggctgac ccccttggct gcccacaagg ccaacgggcc tgagccccca cagggccatg
                PEA3 PEA3
 -708 ggcattacct gctgaattga ggagcccata aggagtcact tggaccacag tgaacacttg gcgaccactg acactcagga gaccttagct ggtcctccag

 -608 cacctctcaa ctccactcct actaaactgg aacttctgtg ctcaggccag agtcgggtcc gtcaccgagt atgttatgcg ctgcccatca ccgaggatgc
                                                                                                    ETS-1
 -508 catgcgctgt aagagggctg ccaccgcggc aggctgacca tggcagggtc ggaacgcaac ctgagagcca gcttgttctg gccagcagtg cccactgggc

 -408 gacctagcag cctcctgata tggggctgtg tccccctctc ccctgcactg ggtacccccа actgaggata ttgctgagtc atggccaggc ccaagcctgg
                                                                                  AP-1 / TRE
 -308 gaggggggcga ggggctggac ccccgccagt accctgatcc caggtgcaga ggctggagcc caggcctgta tgagtgccag ggccggtttc ctggggtcct
                                                                                                    PEA3
 -208 tggtgcaccg gggcaatgaa gagaggggtc agcaatgagg ggggccggga gacctggagc gaggggtagc ggggaagggg agagtagtga aggggcctct

 -108 gcagggcggc tctcgcgccg cgacgacggt ggcgggggcg gggagggcgc gagagactcc gcccctctcg aggcgggggcg gggcctccgc gttcgctaca
                                  SP-1 / AP-2                 SP-1                SP-1 / AP-2
   -8 aaagccgcgc ggcggctgcg accgggacgg cccatttcc gccagctcgc cgctcgctat ggcgtcgctc accgtgaagg cctaccttct gggcaaggag
                                                                 M   A   S   L   T   V   K   A   Y   L   L   G   K
  +93 gacgcggcgc gcgagattcg ccgcttcagc ttctgctgca gccccgagcc tgaggcggaa gccgaggctg cggcgggtcc gggaccctgc gagcggctgc
       D   A   A   R   E   I   R   R   F   S   F   C   C   S   P   E   P   E   A   E   A   E   A   A   A   G   P   G   P   C   E   R   L
 +193 tgagccgggt ggccgccctg ttccccgcgc tgcggcctgg cggcttccag gcgcactacc gcgatgagga cgggg*** INTRON I ***
       L   S   R   V   A   A   L   F   P   A   L   R   P   G   G   F   Q   A   H   Y   R   D   E   D   G
```

Fig. 3. Nucleotide sequence of the 5′-flanking region and the first exon of the p62 gene. The putative transcription start site identified by primer extension is indicated by an arrowhead, and the ATG translation initiation site by a box. The numbers on the left are relative to the guanosine of the putative transcription start site. Exon 1 is shown in boldface letters, and deduced amino acid sequences in boldface capital letters. Consensus binding sequences for the known transcription factors are underlined. The nucleotide sequence was deposited in the GenBank under accession number AF060494.

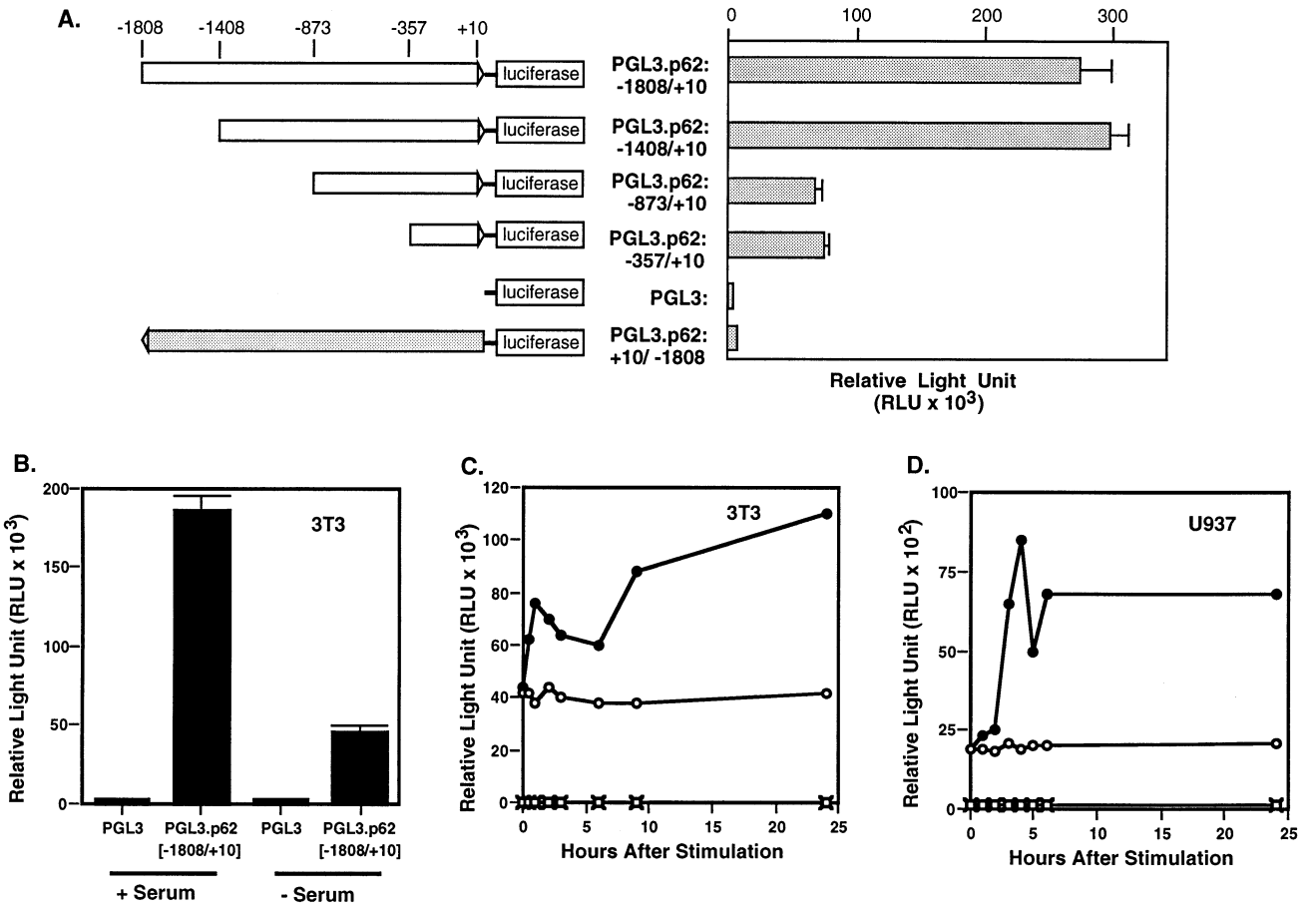

Fig. 4. Activation of the p62 promoter. A: Promoter activity of the 5′-flanking region of the p62 gene. Constructed luciferase reporter plasmids are illustrated on the left (see Section 2). The reporter plasmids were cotransfected with pCMV.β-gal into NIH3T3 cells. The effects of the 5′-flanking sequences on luciferase expression were normalized to β-galactosidase activity, and are expressed as relative light units (RLU). B: p62 promoter activity in NIH3T3 cells in the presence and absence of 10% serum. C, D: Time-dependent changes in the p62 promoter activity of serum-starved NIH3T3 (C) and U937 (D) cells after stimulation with 15% serum and 20 ng/ml PMA respectively. Luciferase activities driven by PGL3.p62[−1808/+10] in stimulated (closed circles) and unstimulated (open circles) cells and those by PGL3 (open square and cross) are shown. The results are representative of more than three independent experiments.

a temporary 40% reduction, the induced luciferase activity upregulated again and was maintained thereafter. These results confirm that the cloned 5′-flanking region contains *cis*-element(s) responsible for activation of the p62 gene during cell proliferation and differentiation.

Nevertheless, the promoter of the p62 gene likely has multiple regulatory features for responding to divergent signals in various tissues. Indeed, the p62 promoter sequence shows the presence of a number of potential binding sites for known transcription factors (Fig. 3). These include three TPA-responsive element (TRE) sites which are known to be high affinity binding sites for heterodimers between Fos- and Jun-related bZip proteins or Jun homodimers [15,16]. Another characteristic of the p62 promoter region is the presence of multiple binding sites for Ets-1 family transcription factors, including two Ets-1 sites, six PEA3 sites, and two Pu.1 sites (Fig. 3, and [17]). PEA3 is induced by TPA, serum, EGF, v-*src*, v-*raf*, or Ha-*ras*, thus representing a primary target of mitogenic signal transduction pathways [17,18]. Pu.1 is another Ets-1 family member known to be involved in monocytic differentiation of bone marrow progenitor cells [19]. An NF-κB and a c-*myc* binding site are also present at −1724 and −1329 respectively. Thus, these *cis*-elements would be responsible, at least partly, for activation of the p62 gene upon signals for proliferation, differentiation, and oxidative stress.

Interestingly, the proximal promoter region and extending into the first exon (−950 to +267) has a high G/C content up to 73%, and is enriched with CpG dinucleotides; the ratio of CpG to GpC in this region is 0.93. Such high G/C content and high frequency of CpG dinucleotides fulfil the definition of CpG islands [20]. Genes containing a CpG island in their promoter region including housekeeping genes, transcription factors, and oncogenic proteins are essential for cell vitality and their activities are highly regulated. One determinant for the relative activity of a promoter containing CpG islands is the state of methylation of cytosine residues in 5′-CpG dinucleotides [20,21]. Further analysis is required to evaluate the significance of CpG island and various *cis*-elements present in the promoter region of the p62 gene.

p62 interacts with multiubiquitin chains. However, unlike proteasomal multiubiquitin chain binding protein Mcb1/S5a, p62 is localized only in the cytoplasm and is not physically associated with the proteasome (Ko et al., submitted). Thus, it is likely that the immediate early response of the p62 gene is essential in cell growth, differentiation, and oxidative stress through a novel mechanism regulating protein metabolism.

Information provided in this article will be useful in further studies to provide more insight into the mechanisms regulating the p62 gene.

Acknowledgements: This work was supported by National Institute of Health Research Grant GM48961 and American Cancer Society PRG-97-018-01CN.

References

[1] Joung, I., Strominger, J.L. and Shin, J. (1996) Proc. Natl. Acad. Sci. USA 93, 5991–5995.
[2] Vadlamudi, R.K., Joung, I., Strominger, J.L. and Shin, J. (1996) J. Biol. Chem. 271, 20235–20237.
[3] Pulse, A., Schmidt, S., Grawe, F. and Stabel, S. (1997) Proc. Natl. Acad. Sci. USA 94, 6191–6196.
[4] Ciechanover, A. (1994) Cell 79, 13–21.
[5] Hicke, L. (1997) FASEB J. 11, 1215–1226.
[6] Deveraux, Q., Ustrell, V., Pickart, C. and Rechsteiner, M. (1994) J. Biol. Chem. 269, 7059–7061.
[7] van Nocker, S., Deveraux, Q., Rechsteiner, M. and Vierstra, R.D. (1996) Proc. Natl. Acad. Sci. USA 93, 856–860.
[8] Ishii, T., Yanagawa, T., Kawane, T., Yuki, K., Sieta, J., Yoshida, H. and Bannai, S. (1996) Biochem. Biophys. Res. Commun. 226, 456–460.
[9] Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
[10] Williams, T.M., Buerlein, J.E., Kricka, L.J. and Kant, J.A. (1989) Anal. Biochem. 176, 28–32.
[11] Padgett, R.A., Grabowski, P.J., Konarska, M.M., Seiler, S. and Sharp, P.A. (1986) Annu. Rev. Biochem. 55, 1119–1150.
[12] Shapiro, M.B. and Senapathy, P. (1987) Nucleic Acids Res. 15, 7155–7173.
[13] Dierick, H.A., Ambrosini, L., Spencer, J., Glover, T.W. and Mercer, J.F.B. (1995) Genomics 28, 462–469.
[14] Azizkhan, J.C., Jensen, D.E., Pierce, A.J. and Wade, M. (1993) Crit. Rev. Eukaryot. Gene Express. 3, 229–254.
[15] Angel, P., Imagawa, M., Chiu, R., Stein, B., Imbra, R.J., Rahmsdorf, H.J., Jonat, C., Herrlich, P. and Karin, M. (1987) Cell 49, 729–739.
[16] Smith, S.E., Papavassiliou, A.G. and Bohmann, D. (1993) Nucleic Acids Res. 21, 1581–1585.
[17] Karim, F.D., Urness, L.D., Thummel, C.S., Klemsz, M.J., McKercher, S.R., Celada, A., Van Beveren, C., Maki, R.A., Gunther, C.V., Nye, J.A. and Graves, B.J. (1990) Genes Dev. 4, 1451–1453.
[18] D'Orazio, D., Besser, D., Marksitzer, R., Kunz, C., Hume, D.A., Kiefer, B. and Nagamine, Y. (1997) Gene 201, 179–187.
[19] Rosmarin, A.G., Caprio, D., Levy, R. and Simkevich, C. (1995) Proc. Natl. Acad. Sci. USA 92, 801–805.
[20] Bird, A.P. (1986) Nature 321, 209–213.
[21] Bird, A.P. (1992) Cell 70, 5–8.