Biochemical and Biophysical Research Communications **259,** 206–211 (1999)
Article ID bbrc.1999.0732, available online at http://www.idealibrary.com on IDEAL®

Cloning and Characterization of the Promoter Region of the Wilson Disease Gene[1]

Won Jun Oh,*,[2] Eun Kyung Kim,*,[2] Kee Don Park,† Si Houn Hahn,‡ and Ook Joon Yoo*,†,[3]
**Department of Biological Sciences, Korea Advanced Institute of Science and Technology (KAIST), Taejon 305-701, Korea; †BioMedical Research Center, KAIST, Taejon, Korea; and ‡Department of Pediatrics, Ajou University School of Medicine, Suwon, Korea*

Received April 9, 1999

Wilson disease (WD), an autosomal recessive disorder of copper transport, is marked by impaired biliary excretion and incorporation of copper into ceruloplasmin. Molecular mechanism regulating the expression of the *WD* gene was studied. We isolated, sequenced, and characterized ~1.3 kb of the 5′-flanking region of the *WD* gene from the human genomic library. The ~1.3 kb of the *WD* sequence directed high level of luciferase activity in HepG2 cells. Interestingly, the 5′-flanking region contained four metal response elements (MREs) and six MRE-like sequences (MLSs), usually found in the metallothionein genes. It also contained a number of putative regulatory elements such as Sp1, AP-1, AP-2, and E-box, but lacked TATA box. The transcription start site was located at 335 base pairs upstream of the translation initiation site. Successive 5′-deletion analyses suggested that the 159-base pair region from −811 to −653, which includes MLS2 (−802 to −796) and MLS3 (−785 to −779), contained one or more positive regulatory element(s). A negative element was also identified at region −1038 to −812. A protein-MLS complex was identified through electrophoretic mobility shift and competition assay using MLS2/MLS3 and HepG2 cell nuclear proteins. © 1999 Academic Press

[1] This work was supported by grant Molecular Medicine Research Group Program (98-MM-01-01-A-01) from the Ministry of Science and Technology through the BioMedical Research Center at KAIST.
[2] These authors contributed equally to this work.
[3] To whom correspondence should be addressed: Department of Biological Sciences, Korea Advanced Institute of Science and Technology, 373-1 Kusong-dong, Yusong-gu, Taejon 305-701, Korea. Fax: 82-42-869-8160. E-mail: ojyoo@sorak.kaist.ac.kr.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank/EBI Data Bank with accession number(s) AF034838.

Abbreviations used: WD, Wilson disease; MNK, Menkes disease; OHS, Occipital horn syndrome; MRE, metal response element; MLS, MRE-like sequence; PCR, polymerase chain reaction; RPA, ribonuclease protection assay; bp, base pair; kb, kilobase.

Wilson disease (WD) is an autosomal recessive disorder of copper transport. Hepatic cirrhosis and neuronal degeneration are major symptom caused by the impairment of biliary copper excretion and toxic effects of accumulated copper (1). Three genetic diseases are known in copper metabolism: Wilson disease, Menkes disease (MNK), and Occipital horn syndrome (OHS) (2-4). Menkes disease, an X-linked disorder of copper uptake, is lethal and caused by a deficiency of copper-dependent enzymes. The *WD* gene encodes a copper-transporting P-type ATPase with 60% similarity of the MNK protein and is primarily expressed in liver (5), while the *MNK* gene is almost ubiquitously expressed in a variety of tissues including heart, lung, pancreas, lymphoblasts, and skeletal muscle, but not in liver (3, 6). The 5′-flanking region of the *MNK* gene has been cloned and sequenced (7). Deletion of one of the three repeats was found in OHS, an allelic form of Menkes disease (7). However, its molecular structure that may give a clue to understanding the regulatory mechanism of expression of the *MNK* gene has not been reported.

Several mutations have been reported in WD and MNK patients since their structural genes were identified in 1993 (8-10). Characterization of the promoter region of the *WD* gene in depth should reveal whether the *WD* gene is also regulated by iron metabolism.

In this report, we describe our studies on the isolation and functional characterization of the 5′-flanking region of the*WD* gene. Our results demonstrate that the *WD* gene contains a single transcription site and lacks TATA box near the initiation site. Functional analysis of the regulatory region of the *WD* gene shows strong cis-acting elements, at −811 to −653 from the transcription initiation site, for high level expression of the *WD* gene.

EXPERIMENTAL PROCEDURES

Cloning and sequencing of the 5′-flanking region of the human WD gene. For cloning of a 5′-flanking region of the *WD* gene, a partial exon 1 region was isolated by PCR from human genomic DNA using the published sequence of the *WD* cDNA (5). The PCR product was

0006-291X/99 $30.00

TABLE 1

Upstream Primers Used for Constructing the Deletion Mutants

Name	Sequence
PWD1	5′-TTGGTACCTCGGCCAACAGTGAA-3′
PWD2	5′-AAGGTACCACGACCGGCTGCTCA-3′
PWD3	5′-TTGGTACCGTGCGTTACTATTGG-3′
PWD4	5′-AAGGTACCAGTGTTCGGCGTGGC-3′
PWD5	5′-TTGGTACCGACAGCCGTCGCTCC-3′
PWD6	5′-TTGGTACCCGTTGGAGGCCATTG-3′
PWD7	5′-AAGGTACCCTCTTGCCGCGGTTG-3′
PWD8	5′-TTGGTACCAGAGCGGACCCGACG-3′
PWD9	5′-AAGGTACCGAGTCTGCGGTCCGG-3′
PWD10	5′-TTGGTACCGGAGCGCACCAGCGC-3′
PWD11	5′-AAGGTACCGACATTGTGGCACTG-3′

subcloned into M13mp19, which named M13mp19/PWD. M13mp19/PWD was used as a template for synthesizing a radioactive single-stranded DNA probe, PWD probe, which was prepared as described previously (11). The PWD probe was labeled to a specific activity of 1.57×10^7 cpm/μg and used to screen 3×10^5 plaques of a human genomic library (EMBL3, a gift from Dr. Y. S. Kim at KRIBB). Southern hybridization analysis was performed to isolate small fragments of the 5′-flanking region with six positive clones (PWD1-6). A *Hin*dIII fragment of ~2.7-kb in PWD5 was subcloned into pUC19, which was designated pPWD5H and used for further study. For determining the entire 2.7 kb sequence, pPWD5H was mapped with various restriction enzymes and sequenced by a dideoxy chain termination method (12).

Primer extension analysis. Total RNA was isolated from HepG2 cells using RNeasy kit (Qiagen). A 26-base primer (5′-CTCTTCACACGGATGATTCAAAGTTG-3′) corresponding to the non-coding strand at positions +113 to +138 was end-labeled with γ-^{32}P-ATP (3000 Ci/mmol) using T4 kinase. This labeled primer was hybridized with 20 μg of total RNA from HepG2 cells and yeast RNA at 40°C and 50°C for 12 hrs. The cDNA was synthesized with M-MLV reverse transcriptase (Gibco BRL) at 37°C for 1 hr. The extended products were analyzed on a 6% polyacrylamide/urea gel in parallel with a DNA sequencing ladder.

Ribonuclease protection analysis (RPA). For *in vitro* transcription, pPWD6-LUC was digested with *Kpn*I and *Sma*I. A 296-bp *Kpn*I-*Sma*I fragment was inserted downstream from a T7 promoter in pBluescript KS (Stratagene). The constructed plasmid was used as a template for *in vitro* transcription of a labeled antisense RNA probe using α-^{32}P-ATP (3000 Ci/mmol) in the presence of T7 RNA polymerase (Promega). RPA was performed with the purified RNA probe as described (11). The RNA size was compared with a DNA sequencing ladder.

Construction of plasmids. A 1,600-bp insert from pPWD5H was amplified by PCR using primers with *Kpn*I and *Xho*I sites at 5′ end, 5′-ACCGGTACCAAGCTTTATATTTAAGT-3′ and 5′-GCACTCGAG-CGTCCCGCACGGACACC-3′. A 1.6-kb *Kpn*I-*Xho*I fragment spanning from −1265 to +335 was subcloned upstream from a luciferase gene in the plasmid pGL2-Basic (Promega), pPWD-LUC. Also, the eleven serial deletion mutants of the *WD* promoter, pPWD1-LUC (−1038), pPWD2-LUC (−811), pPWD3-LUC (−652), pPWD4-LUC (−579), pPWD5-LUC (−493), pPWD6-LUC (−401), pPWD7-LUC (−340), pPWD8-LUC (−265), pPWD9-LUC (−200), pPWD10-LUC (−125), and pPWD11-LUC (−50), were constructed by PCR with eleven sets of primers (Table 1) and pPWD-LUC as a template. Upstream primers used for constructing the deletion mutants were represented in Table 1. A common downstream primer, 5′-CTC-GAGCGTCCCGCACGGACACC-3′, was used.

Transfection and luciferase assay. DNA transfections were performed according to the procedure provided by Life Technology. HepG2 cells were transfected with 1.5 μg of serially deleted lucif-

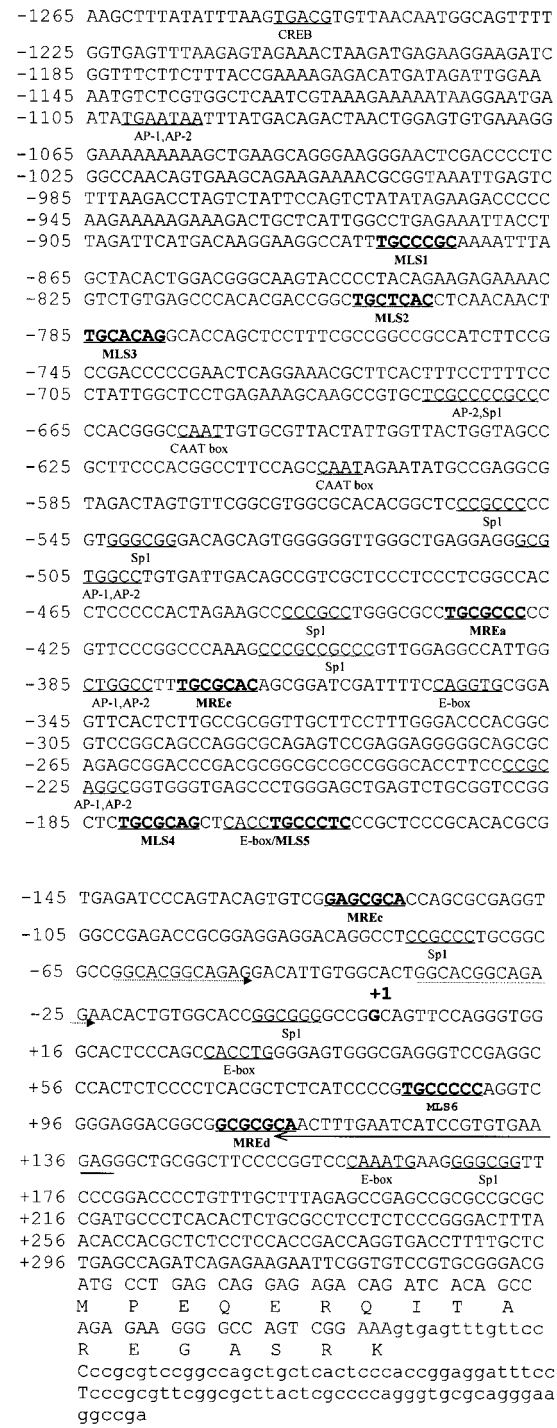

FIG. 1. The nucleotide sequence of the 5′-flanking region of the *WD* gene. The transcription start site is designated as +1. The negative numbers indicate upstream sequence relative to the transcription start site. The putative transcription factor-binding sites are indicated. The oligonucleotide used for primer extension is underlined with an arrow. Two 12 tandem repeat sequence are represented by dotted lines.

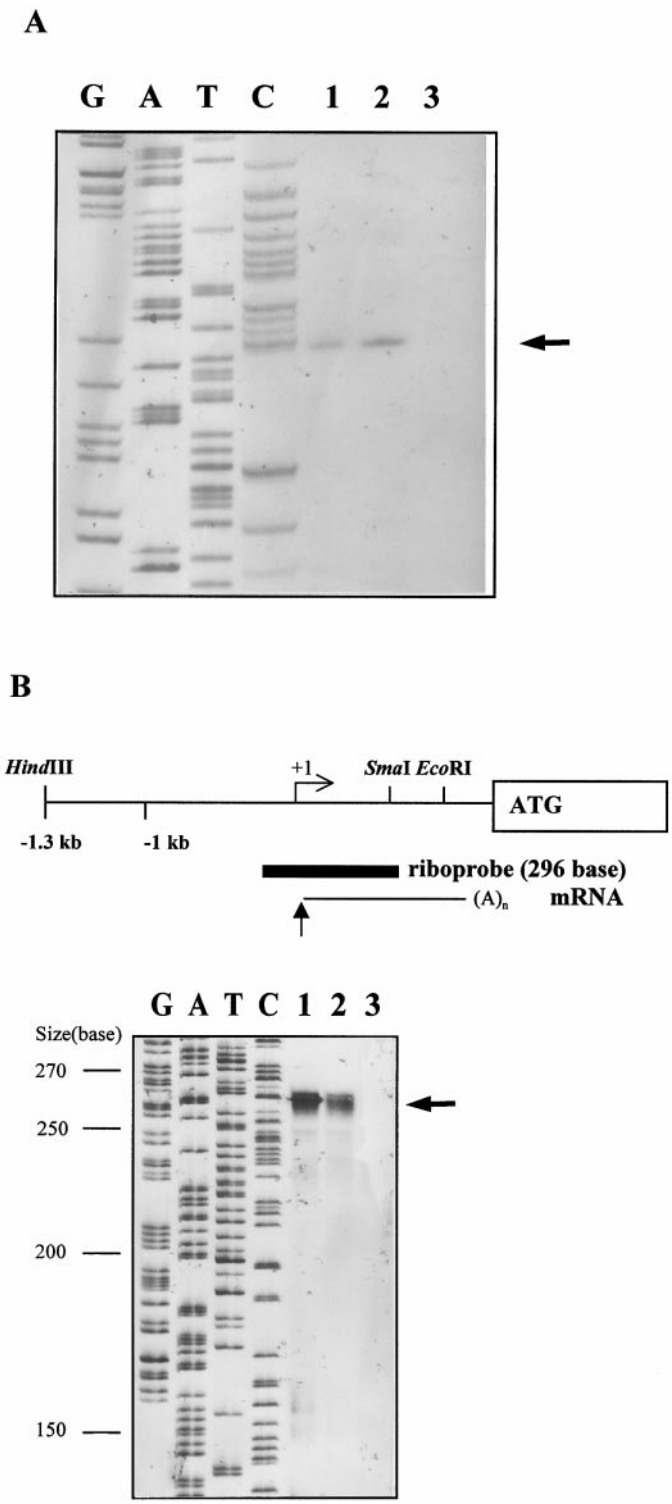

FIG. 2. Determination of the *WD* gene transcription initiation site by primer extension and ribonuclease protection analyses. (A) Primer extension analysis was performed using an oligonucleotide primer corresponding to positions +113 to +138 and total RNA of HepG2 cells. A radiolabeled oligonucleotide was hybridized with 20 μg of total RNA from HepG2 cells and Yeast RNA as a negative control. M13mp18 DNA sequencing ladder was used as a size marker. Arrow indicates the position of the putative transcription start site relative to the ATG translation initiation codon. (B) Ribonuclease protection analysis was carried out using a radiolabeled anitisense RNA probe *in vitro*-transcribed from a 296 bp *Kpn*I-*Sma*I fragment of the *WD* gene. A radiolabeled RNA was hybridized with 20 μg of total RNA from HepG2 cells and control yeast RNA. The protected RNA fragments were resolved juxta-posed to a M13mp18 DNA sequencing ladder.

erase reporter DNAs and 1.5 μg of pSVβ-gal, which served as an internal control. The transfected cells were further cultured for 48 hrs, and expression levels of luciferase and β-galactosidase were determined. Luciferase activity was analyzed with the Promega's luciferase assay kit. Chemiluminescence was measured with the luminometer (Berthold) and the β-galcatosidase activities were determined as described (11). Luciferase activity analysis using all above constructs were repeated three or more times, and the average result was presented.

Electrophoretic mobility shift assay (EMSA). Nuclear extracts were prepared with slight modification as described (13). 2-6 μg of HepG2 nuclear proteins was mixed with an end-labeled double stranded oligonucleotide (30,000 cpm/reaction) in 20 μl of binding reaction buffer (17 mM HEPES, pH 7.9, 32 mM Tris-Cl, pH 7.8, 13% glycerol, 25 mM KCl, 0.8 mM DTT, and 4 μg Poly(dI-dC)), and then incubated for 30 min at 25°C. For competition experiments, the molar excess of unlabeled oligonucleotide was added prior to the addition of labeled probe as specified. The oligonucleotide sequence (−811 to −761) contains the 2nd and 3rd MLS sites (underlined bases denote the MLS functional core; 5′-ACGACCGGCTGCTCAC-CTCAACAACTTGCACAGGCACCAGCTCCTTTCGCC-3′). Protein-DNA complexes were separated electrophoretically in a 4% polyacrylamide gel with 0.5 × TBE.

RESULTS AND DISCUSSION

Isolation and Characterization of the 5′-Flanking Region of the WD Gene

We obtained six positive clones through screening the human genomic DNA library. The positive clones were mapped for restriction enzyme sites, subcloned, and sequenced. Figure 1 shows the sequence of the 1.3 kb of the promoter region. The sequence was run through the 'Signal scan' web program (http://bimas.dcrt.nih.gov:80/molbio/signal/) and other transcription factor searching programs. The putative transcription factor binding sites were also presented in Fig. 1. The 5′-flanking sequence of the *WD* gene had high G and C residues, and contained two putative CAAT boxes at positions −605 and −657. The TATA consensus sequence was not found near the transcription start site. The three 98 bp tandem repeat sequences in the promoter region of the *MNK* gene did not exist in this case (7). Instead, two 12 bp tandem repeat sequences were observed at positions −36 and −62. However, they seemed to have no significant effect on the promoter activity of the gene. Sp1, AP-1, and AP-2 consensus sequences were found at various locations along the 5′-flanking region. Three E-boxes were also identified at positions −355, −173, and +27. Interestingly, four metal response element (MRE) sites, found in metallothionein (MT) promoters (14), were identified. Similar to the *MT* promoters that contain multiple MRE-

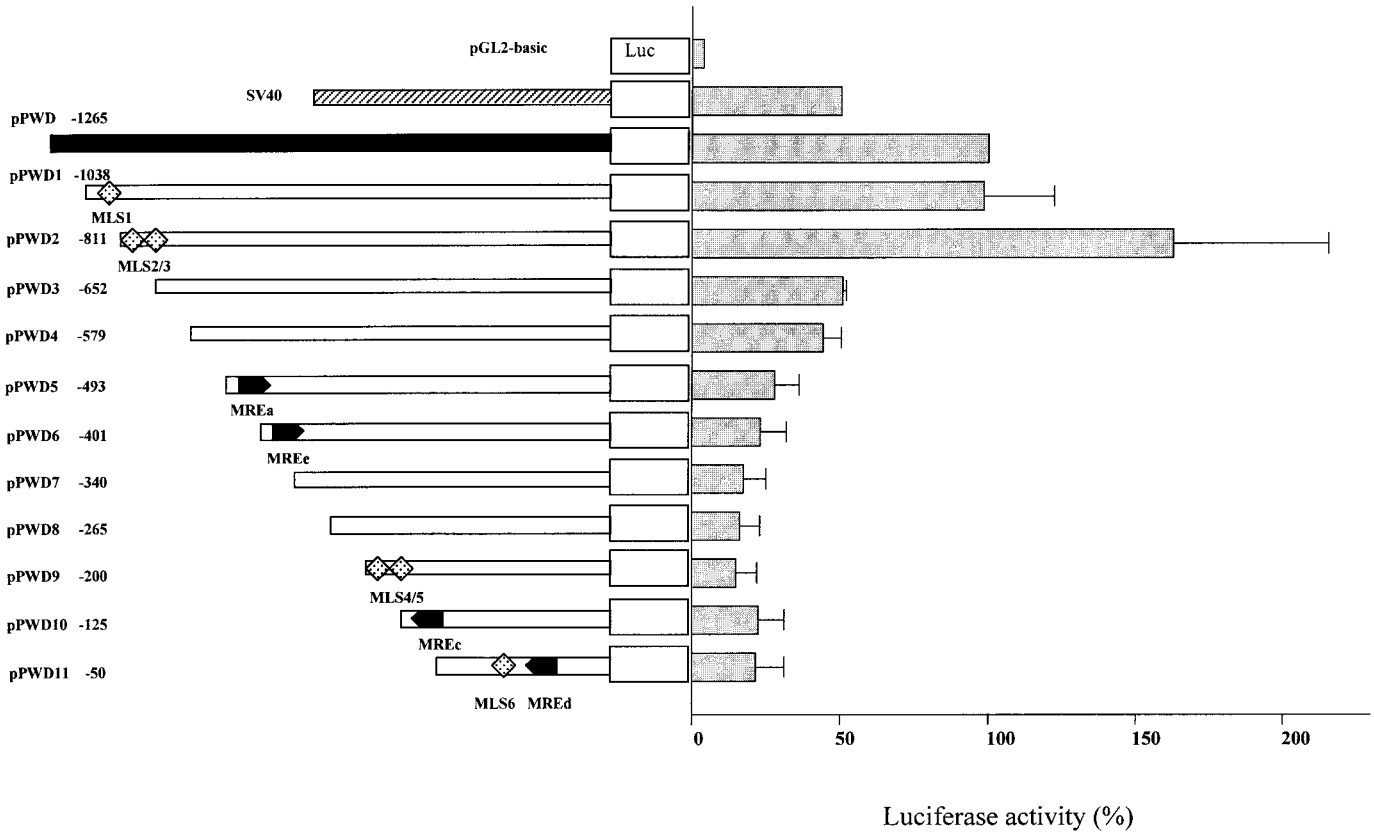

FIG. 3. 5′-deletion analysis of the promoter region of the *WD* gene in HepG2 cells. The structures of the *WD* promoter-luciferase (Luc) reporter deletion series are shown schematically (left) and their relative luciferase activities are shown to the right of each construct. ▰, metal response element (MRE); ◇, MRE-like sequence (MLS). The result is represented as mean with standard deviation.

like sequence (MLS), the promoter region of the *WD* gene contained six additional MLSs that only have one or two bases different from that of MRE consensus sequence (TGCRCNC). In the case of metal-regulated genes, the response to heavy metal is generally independent to the positions or orientation of the MRE, and requires at least two copies of the MRE sequence (15). Since this also seems to be the case in the *WD* gene, the MRE sites in the 5′-flanking region of the *WD* gene may provide as binding sites for metal-dependent transcription factors that play important roles in the control of transcription activity of the *WD* gene.

Identification of the Transcription Start Site of the WD Gene in HepG2 Cells

Primer extension analysis in HepG2 cells using a complementary nucleotide probe corresponding to +113 to +138 (underlined with arrow, Fig. 1) was performed to identify the transcription start site of the *WD* gene. A single transcription start site, corresponding to a G nucleotide located at 335 nucleotides upstream of the translation initiation site, was identified. Nanji *et al.* previously reported that transcription started at three different sites in human liver RNA (16). However, we found that transcription of the *WD* gene started at 122 bp further upstream position from the reported start site of the *WD* gene in human liver (Fig. 2A). Therefore, we designated this nucleotide as +1 for our remaining experiments involving the promoter region of the *WD* gene. To confirm the result from the primer extension analysis, we performed ribonuclease protection assay (RPA) using an antisense RNA probe consisting of a 296 bp *Kpn*I-*Sma*I fragment. The single fragment corresponding in length to the distance between the 3′ *Sma*I and the transcription start site was identified (Fig. 2B). Since the size of the RNA fragment obtained in RPA is generally estimated to be 5-10% smaller than that of the DNA fragment on a polyacrylamide gel (17), the band size agreed well with the result obtained by primer extension experiments. It should be noted that MLS6 and MREd site were located at the downstream region of the transcription start site in the reverse orientation, and further investigation will elucidate whether this site is relevant to the transcriptional or post-transcriptional activity of the *WD* gene.

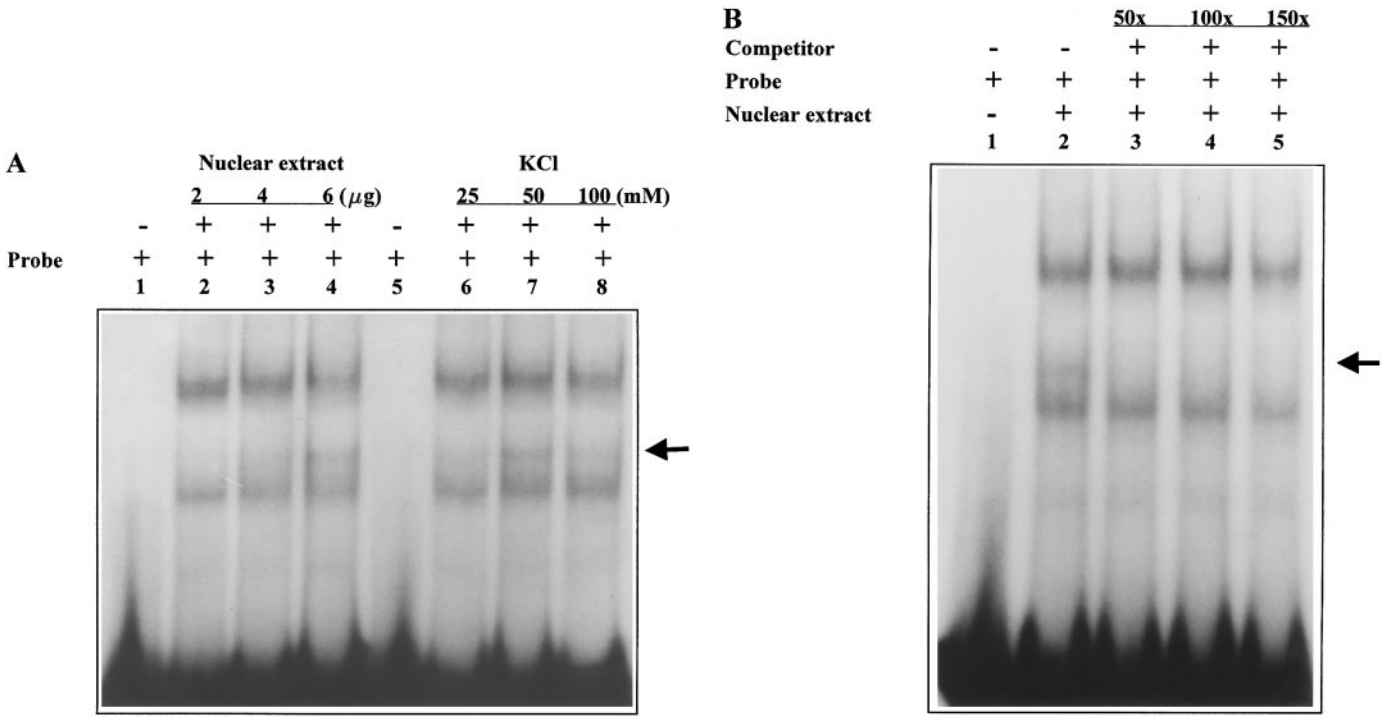

FIG. 4. Electrophoretic mobility shift assay of nuclear proteins interacting with the MLS2/3 sequence. (A) A radiolabeled double-stranded oligonucleotide containing MLS2/3 sequence was incubated without (lane 1) or with 2, 4, and 6 μg of HepG2 nuclear proteins (lanes 2-4). The efficient salt concentration for binding nulcear protein factors was also examined. The radiolabed oligonucleotide was incubated with HepG2 nuclear extracts in the absence (lane 5) or presence of 25, 50, and 100 mM KCl (lanes 6-8). The position of the MLS1-specific band is indicated by an arrow. (B) The radiolabeled duplex probe was incubated with 4 μg of HepG2 nuclear extracts with increasing amouts of competitors. Competition experiment was performed in the absence (lane 1) or presence of 50-, 100-, and 150-fold molar excess of unlabeled MLS2/3 competitors (lanes 2-5). The competed band is indicated by an arrow.

Deletion Analysis of the Promoter Region of the WD Gene

To analyze the promoter function of the 1.3 kb fragment, the activity of pPWD-LUC was measured using pGL2-Promoter as a positive control. 5′-flanking region of the *WD* gene showed high level of luciferase activity in HepG2 cells, indicating the importance of the 1.3 kb 5′-flanking region for the promoter activity of the *WD* gene (Fig. 3). Cis-acting elements important for the expression of the *WD* gene were located by serial deletion of the 1.3 kb *WD* promoter, construction of clones containing various length of the promoter region, and measuring luciferase activity in HepG2 cells. Figure 3 shows the luciferase activity of each constructs. Interestingly, pPWD2-LUC where −1038 to −812 was deleted showed more than 1.5 times higher activity compared that of the pPWD-LUC, suggesting presence of a negative element in this region. On the other hand, pPWD3-LUC (−811 to −653) lacking MLS2, MLS3, Sp1, and 1st CAAT box showed a remarkable decrease (only 50% of normal promoter activity). Further deletion of the region showed gradual decrease in luciferase activity with pPWD9-LUC showing the lowest activity (only 15% of normal activity).

These results indicated that at least one positive regulatory element existed in between −811 and −653 region. Especially, the MLS2 and MLS3 in this region are considered as strong candidate for maintaining the promoter activity of the *WD* gene.

Binding of Specific Proteins to the MLS2/MLS3

Mobility shift assay using the DNA fragment containing the MLS2/MLS3 and HepG2 cell nuclear extracts was performed. Figure 4A shows the results of the assay in different salt and nuclear extract concentrations. Several band shifts were observed. When competed with poly(dI-dC), no significant change was observed with the two major bands, although some of the minor bands disappeared as indicated with arrows (Fig. 4B). There has been a report that MRE site, especially the MREd, contains an Sp1 site and that Sp1 and MTF-1 bind in an overlapping manner (18). However, since our DNA fragment did not contain an Sp1 binding site (CGCC), the major band is thought to be a non-specific band rather than the Sp1-DNA complex. Gunnar *et al.* reported that the binding of transcription factors to the MRE site was dependent on metal ions such as Zn^{2+} and Cd^{2+} (19). We tested whether zinc ion

had any effect on the protein binding *in vitro,* but no significant effect was observed (data not shown). Also, Marta *et al.* reported that addition of metal ions showed significant effect on MRE-protein binding only *in vivo* (20). Further characterization of the protein factors that bind to the MLS2 and MLS3 should lead to a better understanding of the precise mechanism of regulation of the *WD* gene.

The 5′-flanking region of the *WD* gene shares certain common sequences such as E-box and Sp1 sites with that of the *MNK* gene, the most closely related gene to the *WD* gene. However, the presence of MRE and MLS in the *WD* gene and not in the *MNK* gene is unusual, considering the close similarity of the functional effect in copper transport. It should be noted that the *WD* gene contains several MRE and MLS sequences that seem to play important roles during the *WD* gene expression. This was in part supported by the fact that the −811 to −653 region is essential for the promoter activity. Further deletions and mutation studies in this region should elucidate the precise role of the sequences included in this region.

REFERENCES

1. Danks, D. M. (1995) *in* The Metabolic and Molecular Bases of Inherited Disease (Scriver, C. R., Beaudet, A. L., Sly, W. S., and Valle, D. Eds.), 7th ed., pp. 2211–2235, McGraw–Hill, New York.
2. Bull, P. C., Thomas, G. R., Rommens, J. M., Forbes, J. R., and Cox, D. W. (1993) *Nat. Genet.* **5,** 327–337.
3. Vulpe, C., Levinson, B., Whitney, S., Packman, S., and Gitschier, J. (1993) *Nat. Genet.* **3,** 7–13.
4. Levinson, B., Gitschier, J., Vulpe, C., Whitney, S., Yang, S., and Packman, S. (1993) *Nat. Genet.* **3,** 6.
5. Petrukhin, K., Lutsenko, S., Chernov, I., Ross, B. M., Kaplan, J. H., and Gilliam, T. C. (1994) *Hum. Mol. Genet.* **3,** 1647–1656.
6. Chelly, J., Tumer, Z., Tønnesen, T., Petterson, A., Ishikawa-Brush, Y., Tommerup, N., Horn, N., and Monaco, A. P. (1993) *Nat. Genet.* **3,** 14–19.
7. Levinson, B., Conant, R., Schnur, R., Das, S., Packman, S., and Gitschier, J. (1996) *Hum. Mol. Genet.* **5,** 1737–1742.
8. Thomas, G. R., Forbes, J. R., Roberts, E. A., Walshe, J. M., and Cox, D. W. (1995) *Nat. Genet.* **9,** 210–217.
9. Kim, E. K., Yoo, O. J., Song, K. Y., Yoo, H. W., Choi, S. Y., Cho, S. W., and Hahn, S. H. (1998) *Hum. Mut.* **11,** 275–278.
10. Tumer, Z., Lund, C., Tolshave, J., Vural, B., Tønnesen, T., Horn, N. (1997) *Am. J. Hum. Genet.* **60,** 63–71.
11. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd ed. Cold Spring Harbor Laboratory.
12. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467.
13. Dignam, J. D., Lebovitz, R. M., and Roeder, R. G. (1983) *Nucleic Acids Res.* **11,** 1475–1489.
14. Carter, A. D., Felber, B. K., Walling, M. J., Jubier, M. F., Schmidt, C. J., and Hamer, D. H. (1984) *Proc. Natl. Acad. Sci. USA* **81,** 7392–7396.
15. Searle, P. F., Stuart, G. W., and Palmiter, R. D. (1985) *Mol. Cell. Biol.* **5,** 1480–1489.
16. Nanji, M. S., Nguyen, V. T., Kawasoe, J. H., Inui, K., Endo, F., Nakajima, T., Anezaki, T., and Cox, D. W. (1997) *Am. J. Hum. Genet.* **60,** 1423–1429.
17. Chu, Y. Y., Tu, K. H., Lee, Y. C., Kuo, Z. J., Lai, H. L., and Chern, Y. (1996) *DNA Cell Biol.* **15,** 329–337.
18. Peterson, M. G., Tanese, N., Pugh, B. F., and Tjian, R. (1990) *Science* **248,** 1625–1630.
19. Westin, G. and Schaffner, W. (1988) *EMBO J.* **7,** 3763–3770.
20. Czupryn, M., Brown, W. E., and Vallee, B. L. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 10395–10399.