A Conserved Downstream Element Defines a New Class of RNA Polymerase II Promoters*

(Received for publication, August 22, 1995, and in revised form, October 5, 1995)

Tan A. Ince and Kathleen W. Scotto‡
From the Memorial Sloan-Kettering Cancer Center, New York, New York 10021

Although many TATA-less promoters transcribed by RNA polymerase II initiate transcription at multiple sites, the regulation of multiple start site utilization is not understood. Beginning with the prediction that multiple start site promoters may share regulatory features and using the P-glycoprotein promoter (which can utilize either a single or multiple transcription start site(s)) as a model, several promoters with analogous transcription windows were grouped and searched for the presence of a common DNA element. A downstream protein-binding sequence, MED-1 (Multiple start site Element Downstream), was found in the majority of promoters analyzed. Mutation of this element within the P-glycoprotein promoter reduced transcription by selectively decreasing utilization of downstream start sites. We propose that a new class of RNA polymerase II promoters, those that can utilize a distinctive window of multiple start sites, is defined by the presence of a downstream MED-1 element.

Promoters transcribed by RNA polymerase II are divided into two classes: those that contain a canonical TATA box and those that do not. TATA-containing promoters usually direct transcription from a single initiation point, the location of which is determined by the position of TATA (1). In promoters that lack a TATA box, start site selection is not as well understood and has been investigated primarily in genes that use a single transcription start site, where the presence of an “initiator” element at or near the start site appears to be responsible for localizing the preinitiation complex (2). However, despite the fact that many TATA-less promoters utilize multiple start sites, there is little information as to how this multiple selection process occurs. One hypothesis is that the utilization of many start sites is a random or default response to the lack of a strong “selector” such as the TATA box (3). Another possibility is that each start site is independently regulated by a separate initiator-type element; apropos of this, the role of initiators in multiple start site selection within the thymidine synthase promoter was investigated, but multiple initiators were not identified (4).

We have previously shown that transcription from the TATA-less P-glycoprotein (pgp1) promoter can either begin at a single site (+1) or can include multiple downstream start sites within a −70-nucleotide window (5), suggesting that +1 and the downstream sites are independently regulated. In our efforts to understand the activation of the additional downstream start sites within the pgp1 promoter, we have investigated the possibility that the utilization of multiple start sites in TATA-less promoters is neither random nor mediated by independent initiator elements but rather that TATA-less promoters with a similar “window” of start sites may share a common element that regulates their selection and/or activation. In this report we show that: 1) as opposed to being “random,” the size and arrangement of the multiple start site “window” is quite similar in many promoters; 2) a conserved sequence motif (MED-1) can be found in the majority of these promoters downstream of the initiation window; and 4) mutation of this motif within the pgp1 promoter decreases transcription from the downstream start site cassette. We therefore propose that the P-glycoprotein gene is a member of a subclass of TATA-less promoters, which can be classified according to a characteristic transcription “window” and the presence of a common downstream regulatory element.

MATERIALS AND METHODS

Computer Search for the MED-1 Element—The following criteria were imposed upon selection of promoters to be included in this study: 1) they had to contain multiple start sites with a distribution similar to what was found in the pgp1 promoter (i.e. undistributed and spanning less than ~100 bp); and 2) the authenticity of the start sites had to be determined by both nucleotide protection and primer extension assays. ~410 bp of each promoter sequence extending to −70 nucleotides downstream of the initiation window were aligned using the CLUSTAL program (PC/GENE, Intelligenetics) with the following parameters: k-tuple value, 3; gap penalty, 10; window size, 40; filtering level, 5; open gap cost, 10; unit gap cost, 10, with transitions weighted twice as likely as transversions (Fig. 1A).

**Gel Shift Assays—DC-3F/ADII cells were cultured as described previously (5). Nuclear minipreparates were prepared from monolayer cells (6), with the addition of the following protease inhibitors: pepstatin (0.7 μg/ml), leupeptin (0.5 μg/ml), aprotonin (2 μg/ml), E-64 (0.3 μg/ml), bestatin (40 μg/ml), and antipain/dihydrochloride (50 μg/ml). A double-stranded oligonucleotide probe including the pgp1 MED-1 sequence (+63 5'-GGTGAATCAGGACACGGTTCACCAGCCTGCTCCCTACATCCCCAGGCCC-3' +113) was synthesized in the presence of BSA and used in gel shift analyses with one of four double-stranded oligonucleotide competitors: wild type pgp1 MED-1 (+78 5'-CGCTCCAGGACCTGCTCCCTACATCCCCAGGCCC-3' +113), mutant pgp1 MED-1 (+81 5'-TTCAGAGGACCTCAGGATCTCCAGGCCC-3' +113), HMG-CoA reductase MED-1 (7) (+101 5'-GGCGGCGAACCCGTCCCGTCCGGCAGGCCC-3' +136), and nonspecific (5'-CATGCACTTTGGTTAAACTGGTTCG-3'). Competitors and extract were preincubated for 10 min at 25 °C in 20 μl of 20 mM Hapes (pH 8.0), 1.25 mM KCl, 5 mM MgCl₂, 50 μM EDTA, 0.1 mg/ml poly(dI·dC). 0.1 ng of probe was added and incubation continued for an additional 10 min, followed by ultraviolet cross-linking for 5 min in a Stratalinker (Stratagene). While the specific DNA-protein interactions were not dependent on UV cross-linking (data not shown), visualization of the upper complex (Fig. 2A) following

*This work was supported by National Cancer Institute Grant P30-CA-08748 (Memorial Sloan-Kettering Cancer Center), National Cancer Institute Grant RO1-CA57307 (to K. W. S.), and the Samuel and May Rudin Foundation, the Wendy Will Case Cancer Fund, and the Frank L. Horsfall, Jr. Fellowship (to T. A. I.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

†To whom correspondence should be addressed: Memorial SloanKettering Cancer Center, 1275 York Ave., New York, NY 10021. Tel.: 212-639-8972; Fax: 212-639-2787; E-mail: k-scotto@mskcc.org.

‡The abbreviations used are: bp, base pair(s); HMG-CoA, 3-hydroxy-3-methylglutaryl coenzyme A.
**Transcriptional Regulation of Multiple Start Site Promoters**

**A**

**FIG. 1.** Multiple start site TATA-less promoters containing the MED-1 element. A, alignment of promoters with multiple start sites. Promoters were chosen and aligned as described under “Materials and Methods.” Only the 3’-half of the alignment is shown. Identity among all five promoters is indicated by closed circles. **B**, transcription initiation window size and position relative to MED-1. Arrows indicate transcription start sites, as described in references noted (5, 7, 10–21). The size of the arrow does not necessarily correspond to the relative strength of the particular transcription start site. Asterisk indicates values for Go(α) were approximated from the data available. Due to inherent artifacts in assays used to identify mRNA 5’-ends, genes in which start sites were not confirmed by at least two independent methods were excluded from these comparisons. Only promoters with complete identity to the MED-1 consensus are shown.

ACBP, bovine acyl-CoA-binding protein (13); WT1, human Wilms’ tumor (14); MHC-A, human nonmuscle myosin heavy chain (15); N-RAS, mouse N-ras (16); CATL, rat catalase (17); Gαo, mouse G protein (18); AK2, bovine adenylate kinase isozyme (19); GHRH, rat growth hormone-releasing hormone (20); AAT, rat aspartate aminotransferase (21).

**pgp1 Reporter Constructs and Transfections—**A 433-bp pgp1 promoter fragment between +256 and +177 was subcloned into BlueScript KS II (Stratagene) at the BamHI and SalI sites (5). The promoter insert was released with SacI and XhoI, subcloned into the luciferase vector pGL2-Basic (Promega), and designated pgpLuc-B. The mutant MED-1 construct (pgpLuc-Bm) was created by site-directed mutagenesis, using the mutant MED-1 oligonucleotide described for gel shift analyses.

To construct pgp1/globin reporters, the unique BamHI site of the pgpUC-B plasmid was first converted into an Apal site by site-directed mutagenesis; the resulting plasmid was designated pgpUC-Ba. pgp1GL was created by cloning a B-globin insert (isolated from PTAG-1 (8) by Apal/HindIII digestion) into pgpUC-Ba, replacing the luciferase gene and the SV40 3’-untranslated region. pgpGLm was created by the same approach, using pgpUC-Bm as vector. 6 × 10⁵ cells were co-transfected by the calcium phosphate method with 12 μg of reporter plasmid and 0.25 μg of the neomycin resistance plasmid, p308 (ATCC). After 36 h, cells were split into dishes containing medium supplemented with 400 μg/ml G-418 (Life Technologies, Inc.). A typical experiment yielded 300–400 neomycin-resistant clones. Individual clones were isolated after 15–18 days. The presence and integrity of the luciferase constructs were confirmed by Southern blot analyses (data not shown). Luciferase assays were performed using the Promega luciferase reporter assay system, as recommended by the vendor. Protein concentrations were determined using the bicinchoninic acid assay kit (Pierce) using microtiter plates (9). For analysis of pgp1/globin constructs, resulting clones were pooled prior to RNA isolation.

Riboprobe Plasmids and Nucleic Protection Assays—pgp1 reporter/globin inserts were released from pgpGL and pgpGLm by digestion with SpeI and BamHI and cloned into pGEM7zf (+) (Promega) between the XbaI and BamHI sites. Constructs were digested with BspMI and EcoT22I to remove intron I and exon II of β-globin, blunt ended with T4 DNA polymerase, and recircularized. Resulting plasmids (R-GL and R-GLm) were used to generate riboprobes for nuclease protection assays (5). Protected fragments were quantitated using a Fuji PhosphoImager.
A search of the literature identified 14 promoters (7, 10–22) similar to the pgp1 promoter with respect to the distribution of multiple start sites within the transcription initiation window (see “Materials and Methods” for search strategy). We began with the assumption that a DNA element involved in multiple start site selection would be common to all these promoters and, analogous to a TATA box, would lie within a conserved distance from the window. In order to test this hypothesis, several of the promoters were aligned and analyzed for such an element (Fig. 1A). A hexanucleotide sequence, GCTCC(C/G), which we have designated MED-1 (Multiple start site Element Downstream), was identified as the only element common to these promoters (Fig. 1A, outlined). The relationship of this element to the initiation window is shown in Fig. 1B. MED-1 was present in 14 out of 15 promoters (it was not found in the human Ha-ras promoter (22)) and lies 20–45 bp downstream of the 3′-end and a maximum of ~110 bp downstream of the 5′-end of the transcription initiation window.

The striking conservation of MED-1 in multiple start site promoters suggested a role for this element in multiple start site selection and/or activation. In order to test this hypothesis, similar experiments were performed using pgp1/globin reporter constructs (luciferase RNA was undetectable in the previous experiments). Following stable transfection of these reporter constructs into DC-3F/ADII cells, two significant observations were made. First, the endogenous multiple start site pattern was recapitulated (Fig. 4B, lane 1), confirming our initial observation (5) that the selection of multiple start sites is not simply a result of a mutation in the endogenous promoter. Second, mutation of the MED-1 element resulted in a ~3-fold reduction in utilization of the downstream start sites relative to +1 (Fig. 4B and C), indicating that the downstream cassette can be regulated independently.

### RESULTS

A specific oligonucleotide (within the HMG-CoA reductase promoter (A), except competed with an oligonucleotide representing sequences which we have designated MED-1 (Multiple start site Element Downstream), a hexanucleotide sequence, GCTCC(C/G), distance from the window. In order to test this hypothesis, and, analogous to a TATA box, would lie within a conserved start site selection would be common to all these promoters (see “Materials and Methods” for search strategy). We began with the assumption that multiple start site promoters suggested a role for this element in multiple start site pattern was recapitulated (Fig. 4B, lane 1), confirming our initial observation (5) that the selection of multiple start sites is not simply a result of a mutation in the endogenous promoter. Second, mutation of the MED-1 element resulted in a ~3-fold reduction in utilization of the downstream start sites relative to +1 (Fig. 4B and C), indicating that the downstream cassette can be regulated independently.

### DISCUSSION

Previous efforts to understand the regulation of promoters containing multiple start sites have focused on individual genes, and the results have been largely inconclusive (4). We therefore began with the assumption that multiple start site promoters share common regulatory features and that any sequence that is involved in start site selection would be at a conserved position relative to the transcription initiation window. It is important to emphasize that the alignment shown in Fig. 1 preceded the functional evaluation of the MED-1 element, thereby reducing the bias that can be associated with data base searches for a DNA element following its identification in a single gene. Therefore, our analysis of the role of MED-1 in pgp1 transcription has strong predictive value relative to its function in the other genes in which it has been

---

**Fig. 2.** Gel shift analyses of protein complexes interacting with MED-1. A, 1 ng of 32P-labeled MED-1 oligonucleotide was incubated with nuclear extracts prepared from DC-3F/ADII cells without competitor (lane 1), with cold wild-type (WT) oligonucleotide (lanes 2–4), with cold mutant (MUT) oligonucleotide (lanes 5–7), or with a nonspecific (NS) oligonucleotide (lanes 8–10). 200 and 200 ng of competitor oligonucleotide were added. Arrows designate specific complexes. B, same as A, except competed with an oligonucleotide representing sequences within the HMG-CoA reductase promoter (lanes 2–4) or a nonspecific oligonucleotide (lanes 5–7). 100 and 200 ng of competitor oligonucleotide were added.

**Fig. 3.** Mutation of MED-1 reduces pgp1 promoter activity. Each bar indicates an independent clone stably expressing a pgp1/luciferase reporter construct: wild type (hatched), n = 16, X̄ = 2966 ± 1497; MED-1 mutant (white), n = 14, X̄ = 731 ± 682; promoterless (black), n = 11, X̄ = 119 ± 110. Arbitrary luciferase units were obtained after normalization of the values against the amount of protein in each extract. The clones were rank ordered within each group according to activity for ease of visual comparison. Statistical significance was determined by independent t test (α = 0.05).
Fig. 4. Role of MED-1 in start site utilization. A, endogenous pgp1 start sites were determined by nuclease protection analysis as described previously (5, 23). Lane 1, single start site selection in DC-3F cells; lane 2, multiple start site selection in DC-3F/ADII cells. Reproduced from Ref. 5. B, total RNA from DC-3F/ADII cells stably transfected with wild type (W) or MED-1 mutant (M) pgp1 promoter/globin constructs (30 and 60 μg, respectively) was analyzed by nuclease protection (S), using a riboprobe derived from the pgp1/globin construct. Lane 1, untransfected control (U), 30 μg of RNA. Position of start sites is indicated. C, quantitation of data presented in B, represented as the ratio of transcripts initiating within the downstream cassette (DSC)/transcripts initiating at +1. W, wild type; M, MED-1 mutant.

identified. Apropos of this, it is interesting to note that in earlier studies deletion of downstream sequences in both MHC-A (15) and N-ras (16) promoters significantly reduced expression from these genes; we now know that these deletions included the MED-1 element.

Whether MED-1 and its cognate binding proteins act as selectors or activators of multiple start sites is not yet known. However, it is clear that the mere presence of MED-1 is not sufficient for activation of multiple start sites since 1) we have already shown that the same pgp1 promoter that supports multiple start sites in some cells uses only the +1 site in others (5) and 2) the protein binding activity shown in Fig. 2 is also present in cells which only utilize +1 (data not shown). Therefore, we suggest that MED-1 is necessary but not sufficient for multiple start site utilization and that other, likely trans-acting factor(s) impose a higher order of regulation on the recognition of this element.

In conclusion, we propose that a new class of RNA polymerase II promoters can be defined by 1) the size of the transcription initiation window and the arrangements of start sites therein and 2) the presence of a downstream MED-1 element. Since the criteria imposed upon selection of the promoters included in Fig. 1B were quite stringent (requiring verification of start site position by both nuclease protection and primer extension analyses, complete homology with the MED-1 element defined in Fig. 1A, as well as the spatial restrictions suggested by the initial alignment), we predict that as more is known about the spatial and sequence requirements for the MED-1 element, additional promoters will be included in this class.

Acknowledgments—We thank K. Pabon for excellent technical assistance, J. L. Fridovich-Keil for PTAG-1, and J. R. Bertino, N. Rosen, S. Swendeman, and Z. Altun-Gultekin for critical reading of the manuscript.

REFERENCES
1. Kollmar, R., and Farnham, P. J. (1993) Proc. Soc. Exp. Biol. Med. 203, 127–139
2. Weis, L., and Reinberg, D. (1992) FASEB J. 6, 3300–3309
3. Takadera, T., Leung, S., Gernone, A., Koga, Y., Takihara, Y., Miyanoto, N. G., and Mak, T. K. (1989) Mol. Cell. Biol. 9, 2173–2180
4. Geng, Y., and Johnson, L. F. (1993) Mol. Cell. Biol. 13, 4894–4903
5. Ince, T. A., and Scotto, K. W. (1995) Gene (Amst.) 156, 287–290
6. Lee, A. W., and Green, M. R. (1990) Methods Enzymol. 181, 20–30
7. Reynolds, G. A., Basu, S. K., Osborne, T. F., Chinn, D. J., Gil, G., Brown, M. S., Goldstein, J. L., and Luskey, K. L. (1984) Cell 36, 275–285
8. Fridovich-Keil, J. L., Gudas, J. M., Bovard Bryan, I., and Pardee, A. B. (1991) BioTechniques 11, 572–579
9. Hinson, D. L., and Webber, R. J. (1988) BioTechniques 6, 14–16
10. DeWille, J. W., J, en, C-H., Deng, T., Harendza, C. J., and Johnson, L. F. (1988) J. Biol. Chem. 263, 84–91
11. Gudas, J. M., Fridovich-Keil, J. L., Datta, M. W., Bryan, J., and Pardee, A. B. (1992) Gene (Amst.) 110, 205–216
12. Patel, P. I., Franson, P. E., Caskey, C. T., and Chauvin, A. C. (1986) Mol. Cell. Biol. 6, 393–403
13. Mandrup, S., Hummel, R., Ravn, S., Jensen, G., Andersen, P., Gregersen, N., Knudsen, J., and Kristiansen, K. (1992) J. Mol. Biol. 228, 1011–1022
14. Fraizer, G. C., Wu, Y-J., Hewitt, S. M., Myatt, T., Tan, C. C. T., Huf, V., and Saunders, G. F. (1994) J. Biol. Chem. 269, 8892–8900
15. Kawamoto, S. (1994) J. Biol. Chem. 269, 15101–15110
16. Jeffers, M., and Pellicer, A. (1994) Biochim. Biophys. Acta 1219, 623–635
17. Nakashima, H., Yanamato, M., Goto, K., Osumi, T., Hashimoto, T., and Endo, H. (1989) Gene (Amst.) 79, 279–288
18. Li, Y., Mortensen, R., and Neer, E. J. (1994) J. Biol. Chem. 269, 27589–27594
19. Tanaka, H., Yamada, M., Kishi, F., and Nakazawa, A. (1994) Gene (Amst.) 139, 221–227
20. Gonzalez-Crespo, S., and Boronat, A. (1991) Proc. Natl. Acad. Sci. U. S. A. 88, 8749–8753
21. Toussaint, C., Bouquet-Lemercier, B., Garlatti, M., Hanouve, J., and Barouki, R. (1994) J. Biol. Chem. 269, 13318–13324
22. Lu, J., Lee, W., Jiang, C., and Keller, E. B. (1994) J. Biol. Chem. 269, 5391–5402
23. Saconamna, C. F., Bordonaro, M., Chen, J. S., and Nordstrom, J. L. (1992) BioTechniques 13, 846–850
A Conserved Downstream Element Defines a New Class of RNA Polymerase II Promoters
Tan A. Ince and Kathleen W. Scotto

J. Biol. Chem. 1995, 270:30249-30252.
doi: 10.1074/jbc.270.51.30249

Access the most updated version of this article at http://www.jbc.org/content/270/51/30249

Alerts:
• When this article is cited
• When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 23 references, 10 of which can be accessed free at http://www.jbc.org/content/270/51/30249.full.html#ref-list-1