

## Characterization of Human Glycoprotein VI Gene 5' Regulatory and Promoter Regions

Kenichi Furihata, Thomas J. Kunicki

**Objective**—Platelet glycoprotein VI is a collagen receptor belonging to the immunoglobulin-like protein family that is essential for platelet interactions with collagen and is exclusively expressed in the megakaryocytic lineage. The objective of this study was to characterize the human glycoprotein VI gene (*GP6*) 5' regulatory and promoter regions.

**Methods and Results**—We first used 5' RACE to establish experimentally that the major transcription start site lies 28 bp upstream from the start codon. We next subcloned the 5' regulatory region of *GP6* into pGL3-basic [pGL3(−1576)] and used deletion mutagenesis to identify important regulatory regions, comparing the activity of transiently expressed promoter-luciferase constructs in Dami and HeLa cells. We found that megakaryocyte lineage-specific transcription is largely controlled within the segment −191/−39. By site-directed mutagenesis, we confirmed that a GATA-1 site at −176 and an Ets-1 site at −45 play important roles in the regulation of *GP6* transcriptional activity.

**Conclusions**—We have determined that the *GP6* sequence −191 to −39 represents the core promoter and that transcription is driven largely by GATA-1 (−176) and c-Ets-1 (−45) sites within this segment. (*Arterioscler Thromb Vasc Biol.* 2002;22:1733-1739.)

**Key Words:** gene regulation ■ platelets ■ signal transduction

Platelet glycoprotein VI (GPVI) is a member of the immunoglobulin-like protein family that is an essential platelet receptor for collagens<sup>1</sup> and is exclusively expressed in the megakaryocytic lineage.<sup>2,3</sup> We recently reported a variation in platelet GPVI content among normal healthy individuals that correlated well with convulxin-induced platelet prothrombinase activity.<sup>4</sup> Although platelet GPVI levels are most likely under genetic control, the responsible mechanism has yet to be determined. A requisite first step in understanding the regulation of expression of this gene is a characterization of its 5' regulatory region and promoter properties. Considering the critical role of GPVI in collagen-initiated signal transduction and platelet procoagulant activity, the observed variation in GPVI content may influence risk for hemorrhagic or thromboembolic disorders. From a clinical standpoint, it would be important to ascertain the mechanism(s) that controls GPVI content and to screen individuals for predisposition to thromboembolic disorders.

The promoter regions of megakaryocyte-specific genes, such as the integrin subunit  $\alpha_{IIb}$ ,<sup>5,6</sup> platelet factor-4,<sup>7</sup> platelet glycoprotein V,<sup>8</sup> and platelet glycoprotein Ib $\alpha$ ,<sup>9</sup> share a number of common characteristics. In each case, the core promoter is generally short ( $\leq 500$  bp), it lacks CCAAT or TATA consensus sequences, and it has functional GATA, Sp-1, and Ets *cis*-acting elements.<sup>10,11</sup> GATA and Ets factors are not exclusive to the megakaryocyte lineage but are also

critical for expression of erythroid-restricted and other unrelated genes, such as platelet and endothelial cell adhesion molecule-1<sup>12</sup> and P-selectin.<sup>13</sup>

Among megakaryocyte-specific genes, the  $\alpha_{IIb}$  promoter has been the best studied. The human  $\alpha_{IIb}$  promoter has a positive regulatory element in its 5' region (from −598 to −400), wherein there are an Ets binding motif at −515 and a GATA-1 binding motif at −463. Both of these elements seem to play important roles in the expression of  $\alpha_{IIb}$ .<sup>6</sup> Three groups have postulated that there is a silencing element in the  $\alpha_{IIb}$  promoter that suppresses expression of the  $\alpha_{IIb}$  promoter in all cells except those of the megakaryocyte lineage.<sup>5,14,15</sup> However, Era et al,<sup>6</sup> using *in vitro* differentiation of murine D3 embryonic stem cells, recently refuted this observation and observed that solely a positive regulatory element in the region from −598 to −400 is sufficient for megakaryocyte lineage-specific expression of the  $\alpha_{IIb}$  gene.<sup>6</sup> Likewise, Albanese et al<sup>16</sup> reported that the GATA element at −463 and a flanking region of the human  $\alpha_{IIb}$  promoter were sufficient for restoring megakaryocyte lineage-specific expression of the murine  $\alpha_{IIb}$  gene in human cell lines. On the basis of these observations, it seems likely that positive regulatory elements play a prominent role in defining the lineage specificity of the human  $\alpha_{IIb}$  promoter.

To elucidate the mechanism of megakaryocyte lineage-specific expression of GPVI, we first excised a 1.5-kb

Received July 10, 2002; revision accepted July 31, 2002.

From the Roon Research Center for Arteriosclerosis and Thrombosis, Division of Experimental Hemostasis and Thrombosis, Department of Molecular and Experimental Medicine, The Scripps Research Institute, La Jolla, Calif.

Correspondence to Thomas J. Kunicki, PhD, The Scripps Research Institute, Department of Molecular and Experimental Medicine, MEM-150, 10550 North Torrey Pines Rd, La Jolla, CA 92037. E-mail tomk@scripps.edu

© 2002 American Heart Association, Inc.

*Arterioscler Thromb Vasc Biol.* is available at <http://www.atvbaha.org>

DOI: 10.1161/01.ATV.0000034493.76465.FF

fragment of the GPVI gene (*GP6*) 5' regulatory region from a BAC clone harboring the human gene (Clone 211(12) Incyte Genomics) and inserted this segment into pGL3-basic, upstream from a luciferase reporter gene. By luciferase assays and DNase I footprinting, we determined that a region of *GP6* spanning from -191 to -39, which contains 1 GATA-1 and 1 Ets-1 binding site, is necessary for the transcriptional activity of *GP6* in transfected cell lines. From a comparison of genomic DNA of 17 normal volunteers, 3 dimorphisms in the *GP6* 5' regulatory region were found to be in linkage disequilibrium, defining 2 major haplotypes.

## Methods

### Cell Lines and Blood Samples

The human cell line HeLa was purchased from the American Type Culture Collection (Manassas, Va). A Dami cell clone was received as a gift from Dr David Wilcox (Medical College of Wisconsin, Milwaukee). These cell lines were grown in DMEM supplemented with 10% FBS (HeLa) or Iscove's modified Dulbecco's medium supplemented with 10% horse serum (Dami). Blood samples were obtained by medical personnel of The Scripps Clinic General Clinical Research Center from normal human volunteers, as described previously,<sup>4</sup> with informed consent.

### 5' RACE

Online Table I (please see <http://www.ahajournals.org>) summarizes the *GP6*-specific oligonucleotide primers that are referred to in the present study. Total RNA was extracted from washed platelets with the use of an RNAqueous RNA extraction kit (Ambion) according to the manufacturer's instructions. First-strand cDNA was synthesized by using a RETROscript reverse transcription-polymerase chain reaction (PCR) first-strand synthesis kit (Ambion) and primer GSP1 (5'-CCATGATCCCTCCCTTGGAT-3'). The remaining RNA was digested with a combination of RNase H and RNase T1 (GIBCO-BRL), and then the *GP6*-specific first-strand cDNA was purified with the use of a GENECLEAN III kit (BIO 101) and tailed with dCTP by use of terminal deoxynucleotidyl transferase (Roche Molecular Biochemicals). The tailed cDNA was amplified by PCR with the use of 5' RACE anchor primer (GIBCO-BRL, 5'-GGCCACGCGTCTGACTAGTACGGGIIIGGGIIIGGGIIG-3') and primer GSP2 (5'-CCTGCTGGAGAAGCTGTAGCA-3'), followed by a second round of PCR with the use of Abridged Universal Amplification Primer (GIBCO-BRL, 5'-GGCCACGCGTCTGACTAGTAC-3') and primer GSP3 (5'-GAGGCTTCCGTTCTGGTAGGAGCA-3'). The PCR product was gel-purified, subcloned into pGEM-T Easy vector (Promega), and sequenced.

### Comparison of Human and Murine *GP6* Sequence

The human and murine *GP6* sequences were retrieved from the GenBank high-throughput gene sequence (htgs) database (human clone RP11-700B5 and mouse clone RP23-106A10, respectively) by using the appropriate cDNA as query sequence and the BLAST algorithm (available at [http://www.ncbi.nlm.nih.gov/BLAST/blast\\_references.html](http://www.ncbi.nlm.nih.gov/BLAST/blast_references.html)). The sequences of these BAC clones are not yet fully refined, and contigs within the obtained sequences are unordered. On the basis of the corresponding cDNA sequence, we established the correct order of the genomic contigs before the corresponding sequences from each species were compared with one another with use of the Vista software package (available at <http://www-gsd.lbl.gov/VISTA>).

### Generation of Human *GP6* Promoter Constructs

Online Table II (please see <http://www.ahajournals.org>) summarizes the *GP6* promoter constructs created for the present study. A BAC clone harboring human *GP6* was identified at Incyte Genomics, Inc, by PCR-based screening of a human genomic library with the use of primer 6.1 (5'-GAGAAGCCAGTGACCCCTCCGGTG-3') and

primer GSP3, generated in our laboratory. The proximal 5' region of *GP6* (-1576 to 75) was then amplified from this BAC clone DNA by PCR with the use of primers GP6P1F (5'-CGGGTACCTGGAGACAGAATCTCGCTCTG) and GP6PR (5'-CTCGAGCCTCAGGACTACCAAGACAG-3'). The amplified PCR product was subcloned into pGEM-T easy vector (Promega) and sequenced. A 1651-bp *KpnI-XhoI* fragment with the correct sequence was then excised from this vector and subcloned into the luciferase reporter plasmid pGL3-basic (Promega), creating pGL3(-1576). A shorter 1577-bp *XmaI-XhoI* DNA fragment spanning -1502 to 75 was obtained from pGL3(-1576) and cloned into pGL3-basic, creating pGL3(-1502). This construct was subsequently used as a template for generation of a series of 5'-end unidirectional truncation mutants. In brief, pGL3(-1502) was first digested with a combination of *KpnI* and *XmaI*, and then its 5' end was unidirectionally truncated by using the Erase-A-Base Kit (Promega), creating pGL3(-1325), pGL3(-937), and pGL3(-539). An internal deletion mutant was constructed by PCR-based 2-step mutagenesis from pGL3(-1502) such that this mutant, pGL3-del(-191), bears a deletion of the segment between -191 and -39.

To more precisely locate the core promoter of *GP6*, 4 overlapping DNA fragments were amplified by PCR from the genomic DNA of each of 1 high-content and 1 low-content donor<sup>4</sup> by using 1 of the 4 forward primers: -322 (5'-CCC GGAGCCATTCAGAAAGGTGTGA-3'), -191 (5'-CCC GGAGGCTGGCCAGAGGAGATAA-3'), -159 (5'-CCC GGAGGCTGTGTGCAAGT-CAC-3'), and -38 (5'-CCC GGAGCCTGTGCCCCAGGCTCCCA-3') together with the common reverse primer, GP6PR. After nucleotide sequencing of the DNA fragments subcloned in pGEM-T easy, the 4 *GP6* promoter constructs were each subcloned into pGL3-basic, and the resultant constructs were designated pGL3(-322), pGL3(-191), pGL3(-159), and pGL3(-38). The substitution mutant pGL3( $\Delta$ GATA) was made by PCR-based mutagenesis from pGL3(-322), whereby the consensus GATA motif at -176 was mutated to GAGC. Similarly, in the mutant pGL3( $\Delta$ Ets), the Ets-1 site at -45 was mutated from ACATCCTG to ACACTTCG, whereas in the combined mutant pGL3( $\Delta$ GATA/Ets), the GATA-1 and Ets-1 sites were mutated. The sequences of all mutants were confirmed before use.

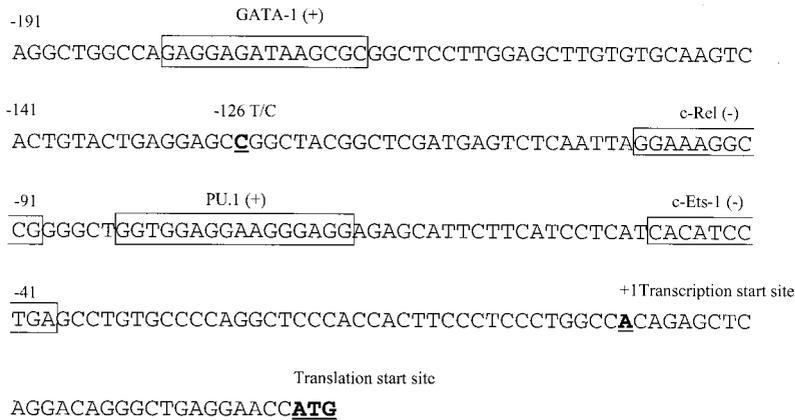
The murine counterparts of the core promoter regions -193 to 75 and -159 to 75 were amplified from mouse genomic DNA by PCR with the use of corresponding primers designed from the sequence of *Mus musculus* clone RP23-106A10 (GenBank accession No. AC087129). After verifying sequences, these were also subcloned into pGL3-basic to generate pGL3(m-193) and pGL3(m-159).

### Transient Expression Analysis of *GP6* Promoter Activities by Luciferase Assay

Dami cells were transfected with 1 of the pGL3-*GP6* promoter constructs together with the Renilla luciferase reporter vector pRL-TK (Promega) by using the Effectene transfection reagent (Qiagen). In brief,  $5 \times 10^5$  Dami cells were transfected with a combination of 0.1  $\mu$ g of pRL-TK and 0.2  $\mu$ g of 1 of the pGL3-*GP6* promoter constructs. Additional cells were transfected with the pGL3-control vector (Promega) and pRL-TK in the same medium. The cells were cultured for 48 hours in the presence of the DNA mixture and then harvested. Firefly and Renilla luciferase activities were measured on a luminometer by using the Dual-Luciferase Reporter Assay System (Promega) according to the manufacturer's instructions. To normalize for transfection efficiency, the promoter activity of each construct was expressed as the ratio of firefly luciferase activity to Renilla luciferase activity. When experiments in different cell lines were compared, the results were expressed as the ratio of *GP6* activity to pGL3-control activity.

### DNase I Footprinting

To confirm the occupancy of putative transcription factor binding elements in the *GP6* core promoter region, we performed DNase I-based footprinting with the use of nuclear extracts from Dami and HeLa cells. The nuclear extract of HeLa cell (HeLaScribe Nuclear



**Figure 1.** Nucleotide sequence of the 5' region of *GP6*. The major transcription start site for *GP6* was revealed by 5' RACE to be an adenine located 28 bp upstream from the ATG translation start site. Notable transcription factor binding motifs in the region from -191 to 75 were identified by using MatInspector<sup>25</sup> and are indicated.

Extract) was purchased from Promega, and that of Dami was prepared by the method of Dignam et al.<sup>17</sup> The footprinting reaction was performed by using the Core Footprinting System (Promega) according to the manufacturer's instructions with minor modifications. Briefly, 5  $\mu$ g pGL3(-322) was digested with *Xho*I and then dephosphorylated at its 5' ends with calf intestine alkaline phosphatase. After <sup>32</sup>P-labeling of both 5' ends with the use of T4 polynucleotide kinase, the DNA fragment was further digested with *Xma*I, and the resultant 397-bp DNA fragment (probe) was gel-purified. This probe was incubated either with 20  $\mu$ g or 50  $\mu$ g Dami nuclear extract or 40  $\mu$ g HeLa nuclear extract for 30 minutes at ambient temperature; a control was incubated in the absence of nuclear extract. The DNA-protein mixture was subjected to RQ1 RNase-free DNase digestion for 2 minutes at ambient temperature. Each reaction mixture was extracted with phenol-chloroform and subjected to electrophoresis on a 6% denaturing sequencing acrylamide gel together with sequence ladders of the same probe. The latter were obtained by using a femtomole DNA Cycle Sequencing System (Promega) and served to locate the occupied sites. The gel was dried and exposed to x-ray film with an intensifying screen at -70°C.

### EMSA for GATA-1 at -176 and Ets-1 at -45

Double-stranded DNA probes were obtained by annealing the following individually synthesized single-stranded oligonucleotides: for GATA -176 (nucleotides -185 to -166, 5'-GCCAGAGGAGATAAGCGCGG-3'); for Ets -45 (nucleotides -51 to -32, 5'-CATCACATCCTGAGCCTGTG-3'). The probes were 5' end-labeled with <sup>32</sup>P before electrophoretic mobility shift assay (EMSA). The following double-stranded competitors were similarly synthesized: GATA mutant (5'-GCCAGAGGAG-AGCAGCGCGG-3'), GATA consensus (5'-CACTTGATA-ACAGAAAGTGATAACTCT-3'); GATA Gel Shift Oligonucleotides, Santa Cruz Biotechnology), Ets mutant (5'-CATCACACTTCGAGCCTGTG-3'), Ets consensus (5'-GATCTCGAGCAGGAAGTTCGA-3'); Ets-1/PEA3 Gel Shift Oligonucleotides, Santa Cruz Biotechnology), and PU.1 consensus (5'-GGGCTGCTTGAGGAAGTATAAGAAT-3'; Ets Family Gel Shift Oligonucleotides, Santa Cruz Biotechnology). The same HeLa and Dami nuclear extracts prepared for DNase I footprinting were used in this assay.

Each <sup>32</sup>P-labeled probe (50 fmol) was mixed with 10  $\mu$ g nuclear extract, 2  $\mu$ g poly(dI-dC) (Amersham Pharmacia Biotech), and 10-fold or 100-fold molar excess of competitor in the binding buffer (25 mmol/L HEPES, pH 7.9, containing 10% [vol/vol] glycerol, 50 mmol/L KCl, 0.5 mmol/L EDTA, 0.5 mmol/L phenylmethylsulfonyl fluoride, and 0.5 mmol/L dithiothreitol) and incubated at ambient temperature for 30 minutes. Samples were then analyzed in a 4% or 5% polyacrylamide gel with the use of a running buffer of 50 mmol/L Tris and 0.38 mol/L glycine, pH 8.5, containing 5 mmol/L EDTA. After electrophoresis, the gel was dried and exposed to x-ray film.

### Nucleotide Sequence Variation in the 5' Regulatory Region of *GP6*

Genomic DNA was isolated from peripheral blood leukocytes of each of 1 high-content and 1 low-content GPVI donor<sup>4</sup> by using the Puregene DNA isolation kit (Blood kit, Genra Systems). A DNA fragment of 1.5 kb corresponding to the *GP6* 5' regulatory region was amplified by PCR by using primers *Kpn*I-GP6P1F (5'-CGGGTACCTGGAGACAGAATCTCGCTCTG-3') and *Bg*/III-GP6P1R (5'-CGAGATCTCACAGGCTCAGGATGTGATG-3') and then directly sequenced.

We identified a T/C dimorphism at -126 and 2 nearby dimorphisms of G/A at -1536 and T/C at -1530 between the high- and low-content *GP6* donor sequences. To obtain genotypic frequencies of these dimorphisms, we similarly extracted genomic DNA from 15 more subjects. A genomic DNA fragment encompassing the 2 nearby dimorphisms was amplified by seminested PCR with the use of primers GP6P1F and GP6P1R in a primary PCR reaction and then GP6P1F and GP6P4R (5'-AGCCTGGGCAACACAGTAAA-3') in a secondary, seminested PCR reaction. The 196-bp DNA product was directly sequenced to identify the -1536 G/A and -1530 T/C dimorphisms. Another T/C dimorphism at -126 was genotyped by *Nae*I digestion of a 444-bp PCR product obtained with primers GP6P2F (5'-TGGGCAAAAAGAACGAGACT-3') and GP6PR. *Nae*I cuts the DNA fragment bearing -126C (GCCGGC), yielding 200-bp and 244-bp fragments in individuals homozygous for -126C. In individuals homozygous for -126T, a single 444-bp fragment is obtained. In individuals heterozygous for this dimorphism, 3 fragments are obtained: at 444, 244, and 200 bp.

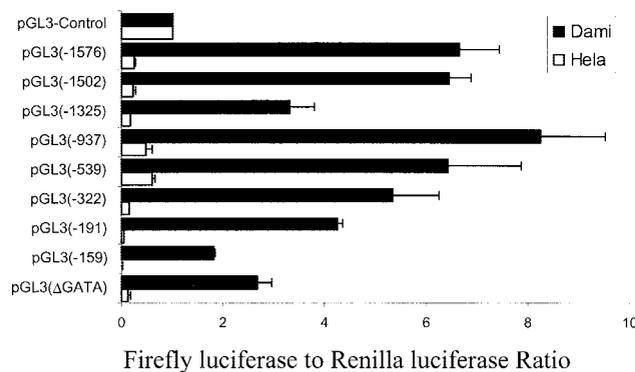
### DNA Sequencing

DNA sequencing was performed by the dideoxy method in the DNA Core Sequencing Laboratory of the Department of Experimental Medicine (The Scripps Research Institute).

## Results

### 5' RACE

A 330-bp DNA fragment was obtained after the final round of PCR. Three of the 5 sequenced clones began at an adenine located 28 bases upstream from the translation start codon ATG (Figure 1); the remaining 2 clones began at a cytosine located adjacent to and 1 base downstream from the aforementioned adenine. Consequently, we assigned the major transcription start site for *GP6* to the adenine located 28 bases upstream from the ATG. This experimental finding confirms the transcription start site predicted in the report of Ezumi.<sup>18</sup> This transcription start site is only 3 bases upstream from the original *GP6* cDNA sequence obtained by primer extension.<sup>2</sup> Throughout the present article, all of the nucleotides in the



**Figure 2.** Transient expression analysis of the *GP6* promoter in Dami and HeLa cells. A series of *GP6*-luciferase constructs were generated by 5' end unidirectional truncation. Each construct is identified regarding the base pair number at the 5' end relative to the transcription start site. Each bar represents normalized relative luciferase activity in Dami cells (solid bar) or HeLa cells (open bar) and is the mean of 3 independent experiments. Error bars represent 1 SD. The activity of each *GP6* promoter construct is expressed as a ratio of that to SV40 promoter-driven pGL3-control.

*GP6* 5' regulatory region are numbered relative to the adenine base transcription start site, now designated 1.

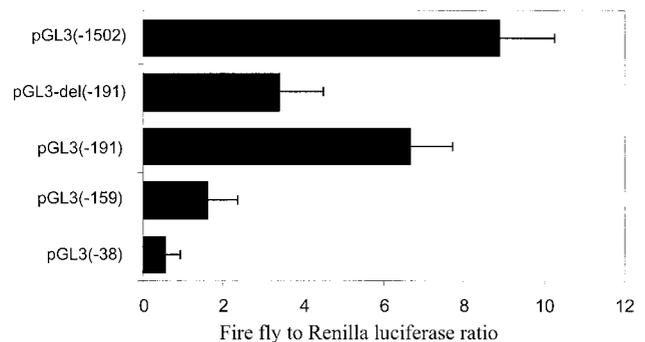
### Vista Analysis of Homologous Regions Between Human and Murine *GP6*

Both the *GP6* sequences of the *Mus musculus* clone RP23-106A10 (GenBank accession No. AC087129) and *Homo sapiens* chromosome 19 clone RP11-700B5 (GenBank accession No. AC019238) were reordered regarding the corresponding cDNA sequences. By comparing the murine and human *GP6* sequences with use of the Vista software package, we were able to identify highly homologous sequences in the region immediately 5' to exon 1. Within the proximal portion of the *GP6* 5' regulatory region, there is 1 conserved noncoding sequence (CNS) in which the human and murine sequences exhibit 77% identity over a 257-bp segment that spans from -182 to 75 (human sequence numbering; online Figure I, available at <http://www.ahajournals.org>). The high degree of homology within this sequence between these evolutionarily disparate species implies an important function for this region of the *GP6* gene.

### Characterization of the 5' Regulatory Region of *GP6* by Luciferase Reporter Assay

The plasmid pGL3(-1576) contains a segment of the *GP6* 5' regulatory region DNA from nucleotide -1576 to 75 that includes the 5' regulatory region and exon I, which encodes the first 11 amino acids of the signal peptide, as well as a part of intron I. A second plasmid, pGL3(-1502), contains the same *GP6* DNA but spans from -1502 to 75. A series of 5' unidirectional deletion mutants (online Table II) were derived from pGL3(-1502), as described in Methods, and tested for their ability to drive the luciferase reporter gene in transiently transfected Dami cells. Figure 2 summarizes the results of these transfection studies.

Relative to pGL3(-1576), pGL3(-1325) showed a 50% decrease in promoter activity, whereas pGL3(-937) exhib-



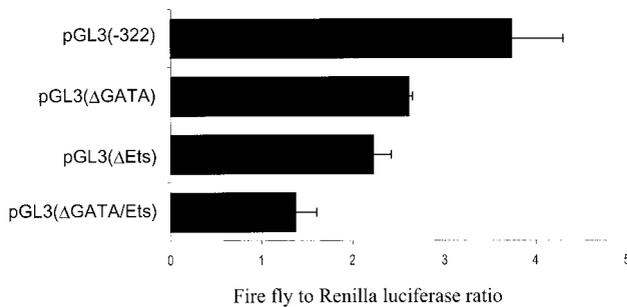
**Figure 3.** Determination of *GP6* core promoter sequence. An internal deletion mutant, pGL3-del(-191 to -39), and 5' truncation mutants were generated from a parent *GP6*-luciferase construct, pGL3(-1502), as described in Methods. Dami cells were transiently transfected with each of the promoter-luciferase reporter gene constructs together with pRL-TK, and firefly and Renilla luciferase activities were determined. Mean (solid bars) and 1 SD are shown.

ited a 24% increase in activity. These findings suggest that negative regulatory elements are located in a region between -1325 and -937. Relative to pGL3(-937), pGL3(-191) exhibited an  $\approx 50\%$  decrease in activity, and pGL3(-159) showed a  $>75\%$  decrease in activity. These results indicate that the *GP6* core promoter elements are located downstream from -191 and that there is an essential positive control element within the sequence spanning from -191 to -159. Qualitatively similar findings were obtained in K562 cells transfected with the same reporter constructs (not shown).

To evaluate cell-type specificity of the *GP6* promoter, the same promoter constructs were transfected into HeLa cells, and luciferase activities were similarly measured. Although pGL3(-937), pGL3(-539), and pGL3(-322) showed very weak promoter activities in HeLa cells, pGL3(-191) and pGL3(-159) showed little if any activity (Figure 2). Mutation of GATA at -176 decreased the promoter activity of pGL3( $\Delta$ GATA) in Dami cells compared with its parent construct, pGL3(-322), but showed no effect on the promoter activity in HeLa cells.

### Contribution of GATA and Ets Elements to *GP6* Core Promoter Activity

The deletion of the sequence -191 to -39 from pGL3(-1502) resulted in a 62% decrease in the activity of pGL3-del(-191) in transfected Dami cells, confirming that a major positive regulatory element resides within this segment (Figure 3). Within the -191/-39 region, there are 1 GATA-1 binding site at -176 and 1 Ets-1 binding site at -45 (Figure 1). To measure the contribution of these GATA-1 and Ets sites to promoter activity, we generated and tested 3 shorter promoter constructs, pGL3(-191), pGL3(-159), and pGL3(-38), in Dami cells. Although pGL3(-191) retained  $\approx 80\%$  of the activity of pGL3(-1502), deletion of a segment from -191 to -160, which contains the GATA-1 site at -176, decreased the promoter activity of pGL3(-191) by 4-fold. A further removal of the sequence from -159 to -39 decreased the promoter activity of pGL3(-38) to background



**Figure 4.** Effect of mutation of GATA-1 and Ets-1 site in *GP6* core promoter. Substitution mutants for either GATA-1 at  $-176$  [pGL3( $\Delta$ GATA)] or Ets-1 at  $-45$  [pGL3( $\Delta$ Ets)] and a combined mutant for both [pGL3( $\Delta$ GATA/Ets)] were made by PCR-based site-directed mutagenesis from a parent *GP6*-luciferase construct, pGL3( $-322$ ). Dami cells were transiently transfected with each *GP6*-luciferase construct, and the firefly and Renilla luciferase activities were determined. Mean (solid bars) and 1 SD of the normalized promoter activity are shown.

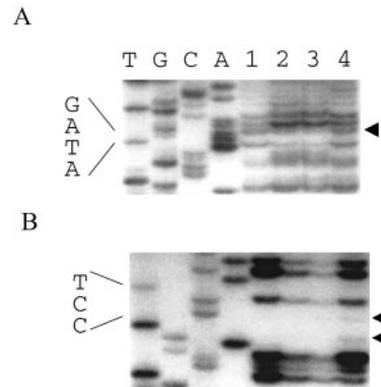
level. The  $-159/-39$  segment contains predicted binding sites for c-Rel, PU.1, and c-Ets-1.

The murine *GP6* sequence corresponding to the human sequence  $-191$  to  $-160$  is also a positive regulatory element (online Figure II, available at <http://www.ahajournals.org>). This segment also contains a T/C dimorphism at  $-126$  that is in linkage disequilibrium with 2 dimorphisms,  $-1536$  G/A and  $-1530$  T/C. The  $-126$ T and the  $-126$ C haplotypes were transfected into Dami cells in the context of pGL3( $-322$ ) and pGL3( $-159$ ). In either case, no significant differences were observed between promoter activities of the 2 haplotypes (data not shown).

Site-directed mutagenesis confirmed the positive regulatory role of the GATA-1 site at  $-176$  and the Ets-1 site at  $-45$ . As illustrated in Figure 4, substitution mutation of the GATA-1 site decreased the promoter activity by 30%, whereas that of the c-Ets-1 site decreased the promoter activity by 40%. The combined mutation of the GATA-1 and c-Ets-1 sites decreased the promoter activity by 3-fold, suggesting an additive effect of these transcription factors.

### Transcription Factor Binding Sites Within the 191 Nucleotides Preceding the Transcription Start Site

A number of transcription factor binding motifs are present in the putative core promoter sequence of *GP6* (Figure 1). To investigate whether any of these motifs can be bound by nuclear proteins from megakaryocytic lineage cells, we performed DNase I footprinting (Figure 5). The GATA-1 site at  $-176$  was clearly occupied by a Dami nuclear protein(s) but not by HeLa nuclear proteins, suggesting *in vivo* involvement of the GATA-1 site for megakaryocyte lineage-specific expression of *GPVI*. However, other predicted transcription factor binding motifs, including c-Rel and PU.1, were apparently not occupied by Dami nuclear proteins in the DNase I footprinting. The c-Ets-1 motif at  $-45$  might be protected, inasmuch as 2 bands indicated by arrowheads in Figure 5B were present in lane 4 but absent in lanes 2 and 3. However, absence of the bands in lane 1 (control) makes it difficult to conclude that the c-Ets-1 site is actually protected by Dami nuclear protein.



**Figure 5.** Occupation of GATA-1 and Ets-1 sites by Dami nuclear extract. The 397-bp fragment of the human *GP6* promoter region ( $-322$  to  $75$ ) was labeled with  $^{32}$ P at the 5' end of antisense strand and incubated alone (lane 1), in the presence of  $20 \mu\text{g}$  (lane 2) or  $50 \mu\text{g}$  (lane 3) of Dami nuclear protein, or in the presence of  $40 \mu\text{g}$  HeLa nuclear protein. The resulting mixture was digested with DNase I as described in Methods. The sequencing ladders of the same probe are shown on the left. Panel A shows a segment corresponding to the GATA-1 site at  $-176$ , and panel B shows a segment corresponding to the c-Ets-1 site at  $-45$ . Arrowheads to the right of the gel represent major protected bands.

### Binding of Nuclear Proteins to Transcription Factor Binding Motifs

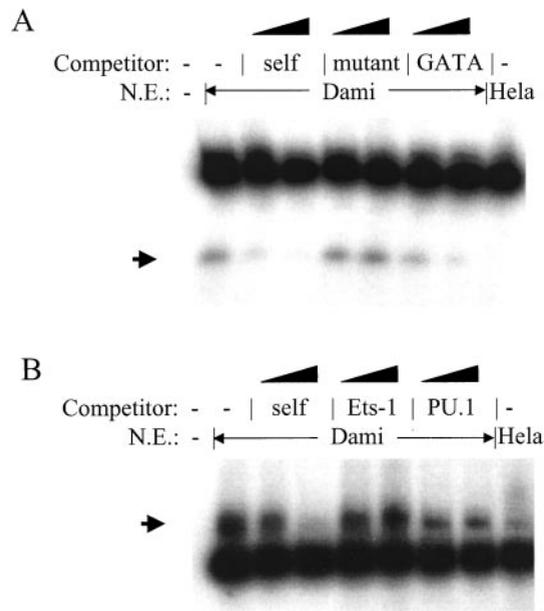
Using electrophoretic mobility shift analysis, we found that nuclear proteins from Dami but not from HeLa cells bind to a 20-bp core promoter fragment (nucleotides  $-185$  to  $-166$ ) that contains the GATA-1 binding motif (Figure 6A). Unlabeled probe (self DNA) and consensus GATA DNA competed away this binding in a dose-dependent manner. However, the mutated self-competitor, in which GATA sequence was mutated to GAGC, did not affect the binding of Dami nuclear protein to the probe. In agreement with the DNase I footprinting results, these observations clearly indicate that the Dami nuclear protein selectively binds to the GATA-1 motif on the probe.

Similarly, Dami nuclear proteins bound to the c-Ets-1 probe (Figure 6B). However, this binding was blocked only by self-competitor and the consensus PU.1 DNA but not by a consensus Ets competitor DNA. The thick band observed in Figure 6A and 6B was considered to be a nonspecific band because it was observed with Dami and HeLa nuclear extract and because it was not affected by any competitor.

### Discussion

*GPVI* is a platelet-specific collagen receptor that belongs to the killer cell receptor subgroup of the immunoglobulin gene family<sup>2,3</sup> and cooperates with integrin  $\alpha_2\beta_1$  *in vivo* to effect optimal collagen-induced platelet adhesion and activation.

In many respects, the *GP6* 5' regulatory and promoter regions resemble those of other genes encoding megakaryocyte lineage-specific proteins.<sup>5-9</sup> Most notably, the *GP6* promoter does not have a TATA box<sup>18</sup> but contains functional GATA-1 and Ets-1 binding motifs. In luciferase reporter assays using 5' unidirectional deletion mutants of the *GP6* 5' regulatory region, we determined that the core promoter of *GP6* lies within the sequence from  $-191$  to  $-39$ . This region



**Figure 6.** Binding of nuclear proteins to transcription factor binding motifs. Nuclear extracts (NEs) were prepared from Damı and HeLa cells and incubated with radiolabeled oligonucleotides either corresponding to GATA-1 at  $-176$  (nucleotides  $-185$  to  $-165$ , A) or Ets-1 at  $-45$  (nucleotides  $-51$  to  $-32$ , B) in the presence of 10-fold or 100-fold molar excess of cold competitor oligonucleotides as indicated. A thick nonspecific band was observed with both oligonucleotides incubated with either Damı or HeLa NE. A specific band that resulted from the binding of Damı nuclear protein to each oligonucleotide is indicated with an arrowhead.

is highly conserved between human and mouse *GP6*, as revealed by Vista analysis, and there is roughly 77% identity within a 259-bp segment from position  $-182$  to  $75$  (human sequence numbering).

One GATA-1 site at  $-176$  and 1 c-Ets-1 site at  $-45$  were identified in the core promoter. Deletion of the sequence  $-191$  to  $-160$  significantly decreased activity in transfected Damı cells, indicating that an essential positive regulatory element exists within this sequence. Several lines of evidence confirmed that GATA-1 at  $-176$  is the essential element. First, mutation of the GATA-1 site decreased the promoter activity in Damı cells by 30%. Second, this GATA-1 motif is fully conserved between the human and murine *GP6* sequences. Third, binding of Damı cell nuclear protein to this sequence was confirmed by DNase I footprinting and EMSA, where this binding was effectively blocked by consensus GATA oligonucleotides.

The *GP6* promoter is also regulated by  $\geq 1$  Ets binding protein, and mutation of the Ets binding element at  $-45$  resulted in a 40% decrease in activity in transfected Damı cells. Concomitant mutation of the GATA-1 and Ets-1 binding elements decreased the promoter activity by 3-fold, indicating an additive effect of GATA-1 and Ets-1. This conclusion is confirmed by the finding that deletion of  $-159$  to  $-39$  decreases promoter activity to almost the background level in Damı cells. However, it was not clear whether the Ets-1 site at  $-45$  was protected in DNase I footprinting assays. The binding of an Ets family transcription factor to

this site may be too weak to be detected by DNase I footprinting, or DNase I may not cut the segment of the probe corresponding to the Ets-1 site to yield a band to be protected.

Ets family transcription factors have been implicated in megakaryocyte lineage-specific expression of a number of genes.<sup>5-9,19-21</sup> For example, Ets-1 is involved in the regulation of platelet factor 4,<sup>7</sup> PU.1 is implicated in the expression of platelet basic protein,<sup>19</sup> and Fli-1 contributes to the regulation of glycoprotein IX expression.<sup>20</sup> The contribution of Ets family proteins to the regulation of  $\alpha_{IIb}$  expression is less clear. Era et al<sup>6</sup> demonstrated that c-Ets-1 is not involved by showing that *c-ets-1*-null murine embryonic stem cells differentiate normally into megakaryocytes in vitro and that the level of *LacZ* expression driven by the  $\alpha_{IIb}$  core promoter in *c-ets-1*-null murine megakaryocytes is similar to that observed in wild-type megakaryocytes.

In Fli-1 knockout mice, among the megakaryocyte lineage-specific transcripts that might be regulated by Ets proteins, including platelet glycoprotein IX, platelet glycoprotein IIb, *c-mpl*, and *vwf*, only platelet glycoprotein IX transcription is decreased.<sup>22</sup> It is unlikely that the expression of *GP6* is regulated by Fli-1 because the *GP6* promoter constructs that we tested showed nearly identical activities in Damı and K562 cells, and K562 is known to not express Fli-1.<sup>20</sup>

Another member of the Ets family, PU.1, is essential in the development of cells of the monocytic, granulocytic, and lymphoid lineage and is known to be a potential cross-antagonist of GATA-1.<sup>23</sup> Because in the case of *GP6*, the Ets-1 family protein in question cooperates additively with GATA-1, it is unlikely that PU.1 is the Ets family protein that controls the *GP6* promoter. Further studies are required to identify and characterize the Ets family protein that is responsible for the control of *GP6* promoter.

In normal subjects, we identified 3 *GP6* sequence dimorphisms at  $-1536$ ,  $-1530$ , and  $-126$ , which are in linkage disequilibrium and define 2 haplotypes: G-T-T and A-C-C. The T/C dimorphism at  $-126$  lies within the *GP6* core promoter region but has no effect on promoter activity in transfected Damı cells. The other 2 dimorphisms are located in a region of the 5' gene sequence that does not contribute significantly to promoter activity, inasmuch as deletion of these sequences has no effect. To further evaluate a possible contribution of these haplotypes to the control of platelet GPVI levels, we compared haplotype frequencies in a group of 17 randomly selected donors. There is no statistically significant correlation between these haplotypes and platelet GPVI content (data not shown). Because platelet GPVI levels are correlated with integrin  $\alpha_2\beta_1$  densities<sup>4</sup> and because the appearance of GPVI during human megakaryocyte maturation coincides temporally with the appearance of integrin  $\alpha_2\beta_1$ ,<sup>24</sup> it is possible that there is a common mechanism involved in the coordinated expression of these important collagen receptors.

The characterization of the *GP6* 5' regulatory region, as described in the present study, represents an important preliminary step in the elucidation of the precise molecular mechanisms responsible for the control of platelet GPVI levels.

## References

- Nieswandt B, Schulte V, Bergmeier W, Mokhtari-Nejad R, Rackebbrandt K, Cazenave JP, Ohlmann P, Gachet C, Zirngibl H. Long-term antithrombotic protection by in vivo depletion of platelet glycoprotein VI in mice. *J Exp Med*. 2001;193:459–470.
- Clemetson JM, Polgar J, Magnenat E, Wells TNC, Clemetson KJ. The platelet collagen receptor glycoprotein VI is a member of the immunoglobulin superfamily closely related to Fc $\alpha$ R and the natural killer receptors. *J Biol Chem*. 1999;274:29019–29024.
- Jandrot-Perrus M, Busfield S, Lagrue A-H, Xiong X, Debili N, Chickering T, Le Couedic J-P, Goodearl A, Dussault B, Fraser C, Vainchenker W, Villevall J-L. Cloning, characterization, and functional studies of human and mouse glycoprotein VI: a platelet-specific collagen receptor from the immunoglobulin superfamily. *Blood*. 2000;96:1798–1807.
- Furihata K, Clemetson KJ, Deguchi H, Kunicki TJ. Variation in human platelet glycoprotein VI content modulates glycoprotein VI-specific prothrombinase activity. *Arterioscler Thromb Vasc Biol*. 2001;21:1857–1863.
- Prandini MH, Martin F, Thevenon D, Uzan G. The tissue-specific transcriptional regulation of the megakaryocytic glycoprotein IIb gene is controlled by interactions between a repressor and positive cis-acting elements. *Blood*. 1996;88:2062–2070.
- Era T, Takagi T, Takahashi T, Bories JC, Nakano T. Characterization of hematopoietic lineage-specific gene expression by ES cell in vitro differentiation induction system. *Blood*. 2000;95:870–878.
- Minami T, Tachibana K, Imanishi T, Doi T. Both Ets-1 and GATA-1 are essential for positive regulation of platelet factor 4 gene expression. *Eur J Biochem*. 1998;258:879–889.
- Lepage A, Uzan G, Touche N, Morales M, Cazenave J-P, Lanza F, del la Salle C. Functional characterization of the human platelet glycoprotein V gene promoter: a specific marker of late megakaryocytic differentiation. *Blood*. 1999;93:3366–3380.
- Hashimoto Y, Ware J. Identification of essential GATA and Ets binding motifs within the promoter of the platelet glycoprotein Iba $\alpha$  gene. *J Biol Chem*. 1995;270:24532–24539.
- Kunicki TJ. The influence of platelet collagen receptor polymorphisms in hemostasis and thrombotic disease. *Arterioscler Thromb Vasc Biol*. 2002;22:14–20.
- Corbi AL, Miller LJ, O'Connor K, Larson RS, Springer TA. cDNA cloning and complete primary structure of the alpha subunit of a leukocyte adhesion glycoprotein, p150,95. *EMBO J*. 1987;6:4023–4028.
- Gumina RJ, Kirschbaum NE, Piotrowski K, Newman PJ. Characterization of the human platelet/endothelial cell adhesion molecule-1 promoter: identification of a GATA-2 binding element required for optimal transcriptional activity. *Blood*. 1997;89:1260–1269.
- Pan J, McEver RP. Characterization of the promoter for the human P-selectin gene. *J Biol Chem*. 1993;268:22600–22608.
- Fong AM, Santoro SA. Transcriptional regulation of  $\alpha_{IIb}$  integrin gene expression during megakaryocytic differentiation of K562 cells: role of a silencer element. *J Biol Chem*. 1994;269:18441–18447.
- Shou Y, Baron S, Poncz M. An Sp1-binding silencer element is a critical negative regulator of the megakaryocyte-specific alpha IIb gene. *J Biol Chem*. 1998;6:5716–5726.
- Albanese P, Leboeuf M, Rosa J-P, Uzan G. Identification of a GATA-overlapping sequence within the enhancer of the murine GPIIb promoter that induces transcriptional deregulation in human K562 cells. *Blood*. 2000;96:1348–1357.
- Dignam JD, Lebovitz RM, Roeder RG. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Res*. 1983;11:1475–1489.
- Ezumi Y, Uchiyama T, Takayama H. Molecular cloning, genomic structure, chromosomal localization, and alternative splice forms of the platelet collagen receptor glycoprotein VI. *Biochem Biophys Res Commun*. 2000;277:27–36.
- Zhang C, Gadue P, Scott E, Atchison M, Poncz M. Activation of the megakaryocyte-specific gene platelet basic protein (PBP) by the Ets family factor PU.1. *J Biol Chem*. 1997;272:26236–26246.
- Bastian LS, Kwiatkowski BA, Breininger J, Danner S, Roth G. Regulation of the megakaryocytic glycoprotein IX promoter by the oncogenic Ets transcription factor Fli-1. *Blood*. 1999;93:2637–2644.
- Mignotte V, Deveaux S, Filipe A. Transcriptional regulation in megakaryocytes: the thrombopoietin receptor gene as a model. *Stem Cells*. 1996;14(suppl 1):232–239.
- Hart A, Melet F, Grossfeld P, Chien K, Jones C, Tunnacliffe A, Favier R, Bernstein A. Fli-1 is required for murine vascular and megakaryocytic development and is hemizygotously deleted in patients with thrombocytopenia. *Immunity*. 2000;13:167–177.
- Cantor AB, Orkin SH. Hematopoietic development: a balancing act. *Curr Opin Genet Dev*. 2001;11:513–519.
- Lagrue-Lak-Hal AH, Debili N, Kingbury G, Lecut C, Le Couedic JP, Villevall JL, Jandrot-Perrus M, Vainchenker W. Expression and function of the collagen receptor GPVI during megakaryocyte maturation. *J Biol Chem*. 2001;276:15316–15325.
- Quandt K, Frech K, Karas H, Wingender E, Werner T. MatInD and MatInSpector: new fast and versatile tools for detection of consensus matches in nucleotide sequence data. *Nucleic Acids Res*. 1995;23:4878–4884.

# Arteriosclerosis, Thrombosis, and Vascular Biology



JOURNAL OF THE AMERICAN HEART ASSOCIATION

## Characterization of Human Glycoprotein VI Gene 5' Regulatory and Promoter Regions Kenichi Furihata and Thomas J. Kunicki

*Arterioscler Thromb Vasc Biol.* 2002;22:1733-1739; originally published online August 22, 2002;

doi: 10.1161/01.ATV.0000034493.76465.FF

*Arteriosclerosis, Thrombosis, and Vascular Biology* is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231

Copyright © 2002 American Heart Association, Inc. All rights reserved.

Print ISSN: 1079-5642. Online ISSN: 1524-4636

The online version of this article, along with updated information and services, is located on the World Wide Web at:

<http://atvb.ahajournals.org/content/22/10/1733>

Data Supplement (unedited) at:

<http://atvb.ahajournals.org/content/suppl/2002/10/09/22.10.1733.DC1>

**Permissions:** Requests for permissions to reproduce figures, tables, or portions of articles originally published in *Arteriosclerosis, Thrombosis, and Vascular Biology* can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the [Permissions and Rights Question and Answer](#) document.

**Reprints:** Information about reprints can be found online at:

<http://www.lww.com/reprints>

**Subscriptions:** Information about subscribing to *Arteriosclerosis, Thrombosis, and Vascular Biology* is online at:

<http://atvb.ahajournals.org/subscriptions/>