Biochemical and Biophysical Research Communications **277,** 27–36 (2000)
doi:10.1006/bbrc.2000.3624, available online at http://www.idealibrary.com on IDEAL®

Molecular Cloning, Genomic Structure, Chromosomal Localization, and Alternative Splice Forms of the Platelet Collagen Receptor Glycoprotein VI

Yasuharu Ezumi, Takashi Uchiyama, and Hiroshi Takayama[1]
Department of Hematology and Oncology, Clinical Sciences for Pathological Organs, Graduate School of Medicine, Kyoto University, 54 Shogoin-Kawaracho, Sakyo-ku, Kyoto 606-8507, Japan

Received August 10, 2000

Glycoprotein VI (GPVI) is the major collagen receptor underlying platelet activation. We cloned the full-length cDNA for GPVI (GPVI-1) and its two isoforms (GPVI-2 and -3) from phorbol-ester-stimulated CMK cells. The GPVI-1 cDNA was identical in the coding region with the cDNA that has recently been reported to belong to the immunoglobulin superfamily. The GPVI gene consisted of 8 exons spanning over 23 kbp and was mapped on the chromosome 19q13.4. The promoter of GPVI gene lacked TATA and CAAT boxes and had multiple transcription start sites like other megakaryocytic genes. When COS-7 cells were cotransfected with the GPVI isoforms and Fc receptor gamma chain, Fc receptor gamma chain was associated with GPVI-1 and -2 but did not affect the GPVI expression levels. GPVI-1 and -2 could bind the collagen-related peptide, which exhibits triple-helical and polymeric structure of collagen to activate platelets via GPVI. © 2000 Academic Press

Key Words: **glycoprotein VI; collagen receptor; human platelets; isoforms.**

Platelets adhere to the extracellular matrix protein, collagen, at the site of vascular damage and become activated through specific membrane receptors, resulting in shape change, granule release, and aggregation. These processes contribute to the formation of a hemostatic plug. Although many platelet surface glycoproteins (GPs) have been proposed as receptors for collagen, it has generally been accepted that the integrin $\alpha_2\beta_1$ and GPVI play a critical role in adhesion and activation, respectively (1).

The first report that GPVI may be involved in collagen-platelet interactions came from the clinical study of a patient with mild bleeding disorders whose platelets lack GPVI but not $\alpha_2\beta_1$ and show selective deficiency in collagen-induced platelet aggregation and release reaction (2). Cross-linking GPVI by the $F(ab')_2$ of anti-GPVI IgG stimulates platelet activation and tyrosine phosphorylation of Syk tyrosine kinase and phospholipase Cγ2 similar to collagen stimulation, whereas GPVI-deficient platelets specifically lack activation of Syk and phospholipase Cγ2 in response to collagen (3, 4). Since then the character of GPVI has been studied mainly from the point of view of intracellular signaling (1). Furthermore, two specific agonists for GPVI have been discovered. One is the snake venom convulxin from *Crotalus durissus terrificus,* which has been known to be a powerful platelet activator (5, 6). The other is collagen-related peptide (CRP), which is made up from a backbone of Gly-Cys-Hyp-$(Gly\text{-}Pro\text{-}Hyp)_{10}$-Gly-Cys-Hyp-Gly and crosslinked via the cysteine residues (7). CRP exhibits triple-helical and polymeric structure to activate platelets via GPVI independent of $\alpha_2\beta_1$ (8, 9). Now has been evident that GPVI is non-covalently and constitutively associated with Fc receptor γ chain (FcRγ) as a signal transducing subunit containing an immunoreceptor tyrosine-based activation motif (10). Analogy to the antigen and Fc receptor signaling, the GPVI-FcRγ complex is associated with the Src family kinases Fyn and Lyn and mediates tyrosine phosphorylation of multiple proteins, including FcRγ, Syk, phospholipase Cγ2, and the adaptor molecules LAT and SLP-76, which are essential for platelet activation upon collagen stimulation (10–14). In contrast to characterization of the GPVI signaling, molecular nature of GPVI itself has not been known, until Clemetson *et al.* had recently cloned the cDNA for GPVI from a bone marrow cDNA

Abbreviations used: GP, glycoprotein; CRP, collagen-related peptide; FcRγ, Fc receptor γ chain; RT, reverse transcription; PMA, phorbol 12-myristate 13-acetate; RACE, rapid amplification of cDNA ends; PBS, phosphate-buffered saline.

The nucleotide sequences reported have been submitted to the DDBJ/EMBL/GENBANK data bank with Accession Nos. AB043819, AB043820, AB043821, and AB043943.

[1] To whom correspondence should be addressed. Fax: +81-75-751-3201. E-mail: hiro@kuhp.kyoto-u.ac.jp.

0006-291X/00 $35.00

library (15). They revealed that GPVI is a member of the immunoglobulin superfamily closely related to Fc receptor for IgA and the natural killer cell receptors.

In this report, we cloned the full-length cDNA for GPVI that was identical with the cDNA reported by Clemetson *et al.* in the coding region but different in the 3′ untranslated region. In addition, we report two other alternative splice forms of GPVI cDNA, the genomic structure, the transcription start sites, the chromosomal localization, and transfection experiments in COS-7 cells.

MATERIALS AND METHODS

Reagents. Convulxin was purified as described by Polgár *et al.* (5) from lyophilized *Crotalus durissus terrificus* venom. Convulxin-coupled Sepharose 4B beads were prepared as described previously (11). CRP was synthesized as described previously (7). Human anti-GPVI IgG was prepared from serum of a patient with GPVI deficiency (2, 3), who has been followed up as an outpatient in our department. Anti-FcRγ IgG was produced by immunizing rabbits as described previously (16). [α-^{32}P]dCTP and [γ-^{32}P]ATP (3000 Ci/mmol, 10 mCi/ml) were supplied by DuPont NEN (Boston, MA). Restriction enzymes, S1 nuclease, and terminal deoxynucleotidyl transferase were obtained from Takara Shuzo Co., Ltd. (Kyoto, Japan). All other reagents were obtained as previously reported (10, 11).

Cell culture. The CMK cells were provided by Dr. T. Sato (Chiba University, Japan) and maintained in RPMI Medium 1640 (GIBCO BRL, Tokyo, Japan) with 10% fetal calf serum (GIBCO BRL), 100 IU/ml penicillin, and 100 μg/ml streptomycin (17). This cell line was derived from a child with megakaryoblastic leukemia and shows some characteristics of mature megakaryocytes after phorbol 12-myristate 13-acetate (PMA) treatment (18). COS-7 cells (American Type Culture Collection, Rockville, MD) were grown in Dulbecco's modified Eagle medium (GIBCO BRL) supplemented with 10% fetal calf serum, 2 mM 1-glutamine, 100 IU/ml penicillin, and 100 μg/ml streptomycin.

Purification of GPVI. Human platelet concentrates outdated for clinical use were obtained from Japanese Red Cross Blood Centers. Platelets (4.0×10^{13} cells) were isolated from pooled concentrates as previously described (19) and centrifuged to form pellets. The pellets were lysed overnight in 2 l of buffer A (20 mM Tris–HCl, pH 7.4, 50 mM NaCl, 5 mM EDTA, 1% Triton X-100, 1 mM benzamidine, 1 mM phenylmethylsulfonyl fluoride, and 50 μg/ml leupeptine) with a mild shake. This and all following steps were performed at 4°C. The lysate was cleared by centrifugation at 10,000*g* for 1 h and determined a protein concentration by BCA Protein Assay Kit (Pierce, Rockford, IL). The supernatant (2 l, 2.5 mg/ml) was loaded on a DEAE Sepharose FF column (2×10 cm, Amersham Pharmacia Biotech UK Ltd., Buckinghamshire, UK) equilibrated in buffer A. The column was washed with 4 column volumes of buffer A containing 150 mM NaCl and eluted with buffer A containing 350 mM NaCl, since our preliminary experiment showed that a major portion of GPVI bound to DEAE Sepharose beads in the buffer containing 150 mM NaCl but unbound to them in the buffer containing 350 mM NaCl. Fractions containing GPVI (180 ml) were determined by immunoblotting with anti-GPVI IgG as described below and were twice passed over a column of 1 ml of convulxin-coupled beads. The column was washed with 300 ml of washing buffer (20 mM Tris–HCl, pH 7.4, 500 mM NaCl, 5 mM EDTA, 1% Triton X-100, 0.1% SDS, 1 mM benzamidine, 1 mM phenylmethylsulfonyl fluoride, and 50 μg/ml leupeptine). Absorbed protein was eluted with 800 μl SDS sample buffer (2% SDS, 5% glycerol, 5% 2-mercaptoethanol, and 62.5 mM Tris–HCl, pH 6.8). Twenty microliters of the sample was subjected to 10% SDS–PAGE under reducing conditions, followed by staining with Coomassie brilliant blue R-250 and immunoblotting with anti-GPVI IgG. We confirmed that a 62-kDa band was GPVI and that the sample contained GPVI at a concentration of 5 ng/μl, comparing with bovine serum albumin as a standard. Thus, we could isolate about 4 μg of GPVI.

Amino acid sequencing. The GPVI sample (750 μl) was subjected to 10% SDS–PAGE on two gels ($160 \times 100 \times 1$ mm), followed by Coomassie brilliant blue staining. GPVI was subjected to in-gel digestion with lysyl endopeptidase (Wako, Tokyo, Japan) and to reversed-phase HPLC on a TSK gel ODS-80Ts QA (2.0×250 mm, Tosoh, Tokyo, Japan). Three peptides among four isolated peptides were sequenced with the Hewlett-Packard G1005A Protein Sequencing System (Hewlett-Packard, Co., Palo Alto, CA).

Construction of a cDNA library. CMK cells were cultured in RPMI 1640 with 20 nM PMA for 3 days. Poly(A)+ RNA was isolated from the cells with the FastTrack 2.0 Kit (Invitrogen, Carlsbad, CA) according to the manufacturer's instructions. A cDNA library was constructed from 5 μg of the poly(A)+ RNA primed with oligo(dT) using the ZAP Express cDNA Synthesis Kit (Stratagene, La Jolla, CA) and ZAP Express cDNA Gigapack III Gold Cloning Kit (Stratagene) according to the manufacturer's instructions. The primary cDNA library in the ZAP Express lambda vector (5×10^5 plaque-forming units) was once plaque-amplified for screening by infection into *Escherichia coli* XL1-Blue MRF'.

cDNA cloning of the GPVI isoforms. The degenerate primers were designed based on the amino acid sequences obtained above. Degenerate PCR was performed on the first strand cDNA synthesized from the poly(A)+ RNA of CMK cells using the Omniscript Reverse Transcriptase (Qiagen Inc., Valencia, CA) and oligo(dT) primers. A 209-bp DNA product was amplified by PCR using the two primers, 5′-CARGCNGTNYTNTTYATHCC (a sense primer based on QAVLFIP) and 5′-CCRTANCKNGTYTGRCAYTG (an antisense primer based on QCQTRYG). The PCR product was labeled with [α-^{32}P]dCTP using Random Primer DNA Labeling Kit (Takara Shuzo Co., Ltd.). The CMK cDNA library was screened by the labeled DNA according to the standard method described by manufacturer. Positive clones were *in vivo* excised into insert-containing pBK-CMV phagemid vector by infection with ExAssist helper phage (Stratagene) followed by transduction of filamentous phage particles into *Escherichia coli* XLOLR according to the manufacturer's protocols.

Rapid amplification of cDNA ends. The rapid amplification of 5′ cDNA ends (5′-RACE) of GPVI was performed as described previously (20). The poly(A)+ RNA from PMA-treated CMK cells was reversed-transcribed into cDNA using the gene-specific primer 5′-GTAGGGTTACGTCCCCTCCT (positions +406 to +425 of the cDNA). The first strand cDNA was purified with QIAquick PCR Purification Kit (Qiagen) and added oligo(dA) at the 5′ ends by terminal transferase. 5′-ends of the GPVI cDNA was PCR-amplified using the gene-specific primer SP5 5′-CTTCCGTTCTGGTAGGAGCA (positions +290 to +309 of the cDNA) and oligo(dT) primers and sequenced.

DNA sequence analysis. PCR products subcloned into pBluescript vector or cDNAs inserted into pBK-CMV vector were sequenced both strands by the BigDye Terminator Cycle Sequencing Kit using the ABI PRISM 310 Genetic Analyzer (Applied Biosystems, Inc., Foster City, CA). Analysis of the sequence data was performed using the BLAST and Electronic PCR programs on the World Wide Web server of the National Center for Biotechnology Information and the Genetyx-Mac ver. 7.3 program (Software Development Co., Ltd., Tokyo, Japan).

S1 mapping. S1 nuclease protection analysis was performed on the poly(A)+ RNA from PMA-treated CMK cells as described (21). A single-stranded DNA probe (positions −213 to +72) was synthesized by genomic PCR using a ^{32}P-labeled primer at a 5′ end and was isolated by electrophoresis on a denaturing 5% polyacrylamide gel. The probe and 2 μg of the RNA was hybridized over-

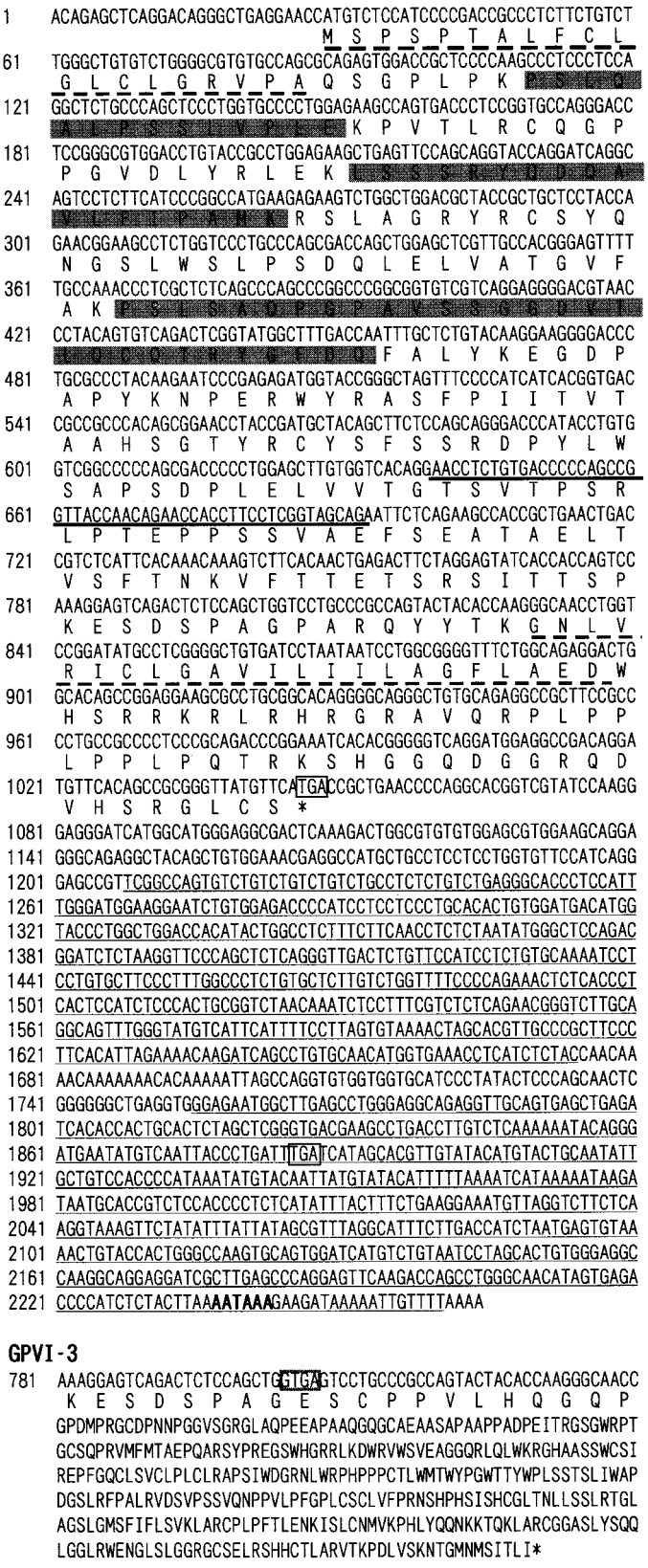


FIG. 1. Nucleotide sequence and the predicted amino acid sequence of the GPVI isoform cDNAs. The three peptides derived from purified GPVI are shaded. The putative signal sequence and transmembrane domain are indicated by dotted underlines. The expressed sequence tags AA308708 and AA494446 are thin underlined. The stop codons (TGA) are boxed. The polyadenylation motif (AATAAA) is in boldface type. The bold underlined nucleotide sequence is deleted in GPVI-2. The boxed nucleotides (GTGA) are inserted in GPVI-3, resulting in a frame shift.

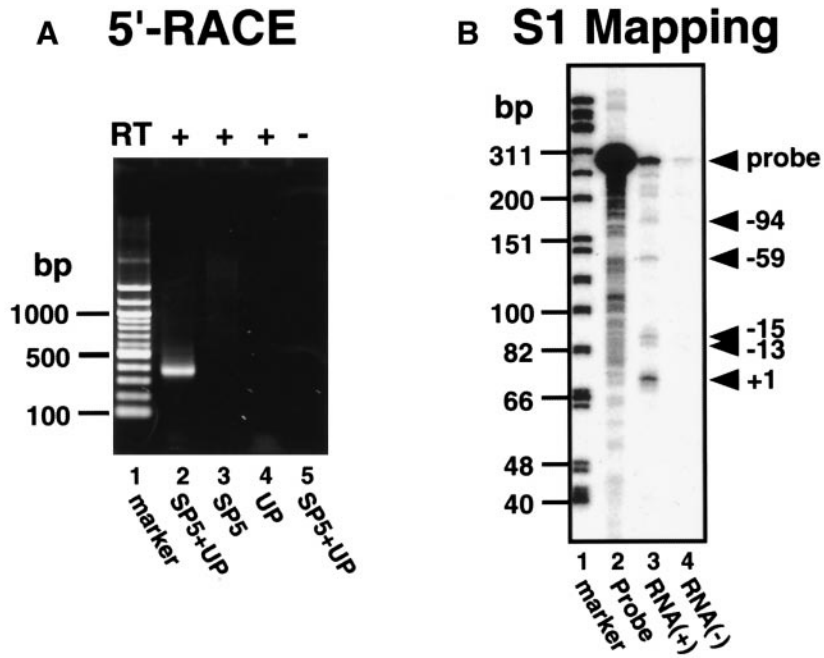
night at 55°C, digested with S1 nuclease for 30 min at 37°C, and analyzed on a denaturing 8% polyacrylamide gel, followed by autoradiography.

Transient expression in COS-7 cells. The cDNA of GPVI was amplified by reverse transcription (RT)-PCR from the CMK RNA using the primers 5′-CAGAGCTCAGGACAGGGCTG (positions +2 to +21) and 5′-GCCTCCCATGCCATGATCCC (positions +1083 to +1102), subcloned into pBluescript vector and sequenced. The insert was excised from the vector with *Eco*RI, since the GPVI cDNA had a restriction site for *Eco*RI at the positions +692 to +697. The insert was ligated into *Eco*RI-digested pBK-CMV containing the GPVI isoform cDNAs that was screened and cloned above. To enhance expression in eukaryotic cells, the vector was digested with *Nhe*I and *Spe*I and religated to remove the prokaryotic 5′-untranslated sequence as recommended by the supplier. Thus, we constructed the expression vectors pBK-CMV-GPVI-1, -2, and -3 that contained the entire cDNAs of GPVI isoforms. The cDNA of the entire coding region of FcRγ was also amplified by RT-PCR from the CMK RNA, subcloned into pBluescript vector, and sequenced. The insert was subcloned into pBK-CMV (pBK-CMV-FcRγ) at *Eco*RI and *Xho*I sites. The pBK-CMV-GPVI-1, -2, or -3 (1 μg/dish) was transfected with or without pBK-CMV-FcRγ (1 μg/dish) into subconfluent COS-7 cells in 60 mm dishes using the Effectene Transfection Reagent (Qiagen) following to the manufacturer's protocol. The cells were studied 72 h after transfection.

Affinity precipitation of convulxin-binding proteins. GPVI and FcRγ associated with it were isolated from COS-7 cells with convulxin-coupled Sepharose 4B beads as described previously (11). Briefly, transfected cells were washed twice with phosphate-buffered saline (PBS) and lysed in buffer containing 1% Triton X-100 (1 ml/dish). The lysates were incubated with 40 μl convulxin-coupled beads. After washing the beads, proteins were eluted in 90 μl SDS sample buffer and boiled for 5 min. For preparation of whole cell lysates, the Triton X-100 lysates were diluted with 2 × SDS sample buffer and boiled for 5 min.

Immunoblotting. Immunoblot analysis of GPVI and FcRγ was performed as described previously (10, 11). The whole cell lysates of COS-7 cells or convulxin-binding proteins were subjected to SDS–PAGE and transferred to nitrocellulose membranes. Immunoblotting was carried out with protein detection by enhanced chemiluminescence.

Flow cytometry. Convulxin and CRP were biotinylated with 5-(*N*-succinimidyloxycarbonyl)pentyl D-biotinamide (Dojindo, Kumamoto, Japan) as described previously (22). Transfected cells were washed twice with PBS and detached by treatment with PBS containing 5 mM EDTA. The cells suspended in 100 μl PBS containing 2 mM EDTA and 0.5% bovine serum albumin were stained with biotinylated convulxin or CRP for 30 min at 4°C, followed by staining for 15 min at 4°C with fluorescein-conjugated streptavidin (Caltag Laboratories, Burlingame, CA). After washing, the cells were analyzed on a FACScan (Becton Dickinson Immunocytometry Systems, San Jose, CA). Washed human platelets were prepared from normal donors or a GPVI deficient patient as described previously (4). The platelets (1×10^7 cells/100 μl) were analyzed by flow cytometry as described above.



C

```
-300  TCATATCTCACTGTGGTTTTGATGCGTATTTTCCTAGTGACATCAGGGAGTTTATGGGAG
-240  CACGGGAACACAGACCAGGCCCCAGCAGGCGGACAAACGGTGCAACGCCAGGCTGGCCAG
-180  AGGAGATAAGCGCGGCTCCTTGGAGCTTGTGTGCAAGTCACTGTACTGAGGAGCCGGCTA
-120  CGGCTCGATGAGTCTCAATTAGGAAAGGCCGGGGCTGGTGGAGGAAGGGAGGAGAGCATT
 -60  CTTCATCCTCATCACATCCTGAGCCTGTGCCCCAGGCTCCCACCACTTCCCTCCCTGGCC
   1  ACAGAGCTCAGGACAGGGCTGAGGAACCATGTCTCCATCCCCGACCGCCCTCTTCTGTCT
  61  TGGGCTGTGTCTGGGGCGTGTGCCAGCGCAGAGTGGACCGCTCCCCAAGCCCTCCCTCCA
 121  GGCTCTGCCCAGCTCCCTGGTGCCCCTGGAGAAGCCAGTGACCCTCCGGTGCCAGGGACC
 181  TCCGGGCGTGGACCTGTACCGCCTGGAGAAGCTGAGTTCCAGCAGGTACCAGGATCAGGC
 241  AGTCCTCTTCATCCCGGCCATGAAGAGAAGTCTGGCTGGACGCTACCGCTGCTCCTACCA
 301  GAACGGAAGCCTCTGGTCCCTGCCCAGCGACCAGCTGGAGCTCGTTGCCACGGGAGTTTT
      SP5
```

D

	Py	Py	Py	Py	-1	+1	Py	Py	Py	Py
Tdt	**C**	**C**	**C**	**T**	**C**	**A**	**T**	**T**	**C**	**T**
GPVI	**T**	G	G	**C**	**C**	**A**	**C**	A	G	A
α_{IIb}	G	G	**C**	**C**	**C**	**A**	**T**	**T**	**C**	**C**
α_2	**C**	**T**	**C**	**T**	**C**	**A**	**C**	**C**	G	G
GPIX	A	**T**	**C**	**C**	**C**	**A**	**T**	A	G	A
PF4	G	A	G	**T**	**C**	**A**	**T**	**T**	G	G
GPV	**C**	**T**	**T**	**T**	**C**	**A**	G	**T**	**T**	A

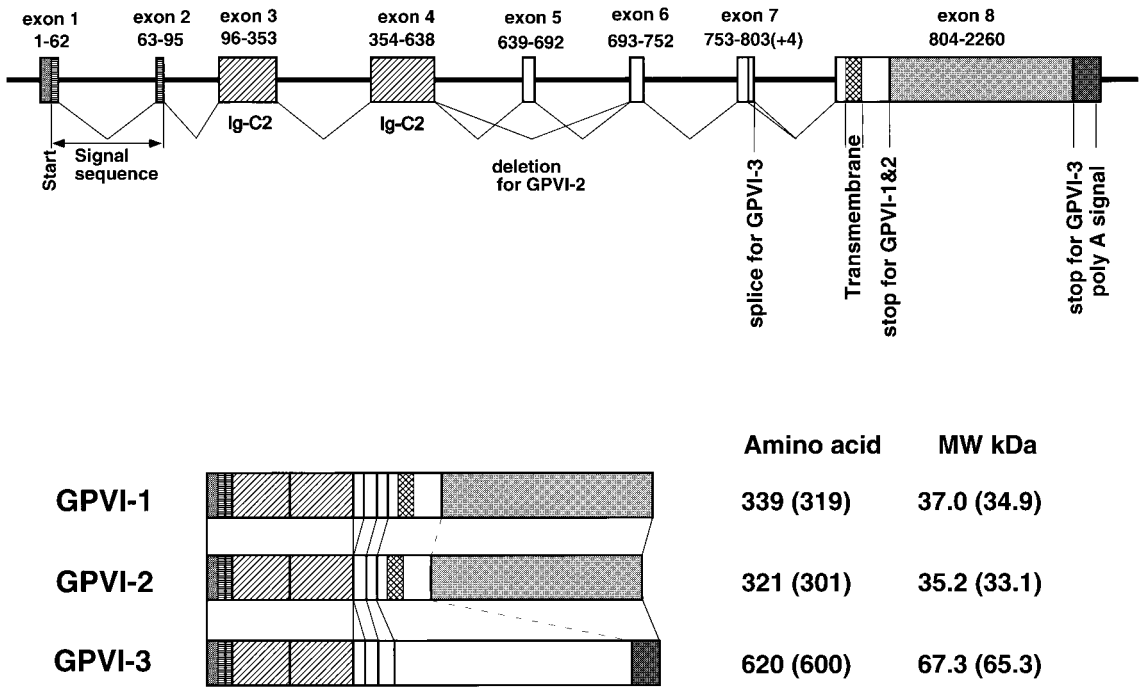


FIG. 3. Genomic structure of the GPVI gene. The eight exons are shown as boxes. The numbers over the boxes are nucleotides referred in Fig. 1. The positions of the start and stop codons are indicated. The signal sequence, Ig-C2 domains, and transmembrane domains are hatched. The untranslated regions are shaded. The numbers of amino acids and theoretical molecular masses of the GPVI isoforms are indicated. The values excluding the signal sequence are also denoted in the parentheses.

RESULTS AND DISCUSSION

cDNA cloning of the GVI isoforms. Because convulxin specifically and strongly binds GPVI (5, 6, 11), we used affinity chromatography on convulxin-coupled agarose beads to concentrate GPVI from the Triton-X 100 soluble phase of platelet lysate and directly subjected the GPVI-rich extract to SDS–PAGE. A 62-kDa band under reducing conditions was identified as GPVI with immunoblotting using anti-GPVI antibody. Using this two-step purification method, we could successfully isolate about 4 μg of GPVI from human platelet concentrates. Four peptides were separated by reversed-phase HPLC after the protein was digested with lysyl endopeptidase. We obtained the following amino acid sequences from three peptides; peptide-1, PSLQALPSSLVPLEK; peptide-2, PSLSAQPGPAVSSGGDVTLQCQTRYGFDQ; peptide-3, LSSSRYQDQAVLFIPAMK. Because we have found that treatment of CMK cells with PMA induced surface expression of GPVI accompanied by expression of FcRγ (manuscript in preparation), we chose PMA-treated CMK cells as a source of mRNA for RT-PCR and to construct a cDNA library for cloning of the GPVI cDNA. PCR with two degenerate primers, CARGCNGTNYTNTTYATHCC (a sense primer based on QAVLFIP of peptide-3) and CCRTANCKNGTYTGRCAYTG (an antisense primer based on QCQTRYG of peptide-2), generated a 209-bp DNA product. We then screened the CMK cDNA library with the PCR product as a hybridizing probe and obtained several clones matching the sequences of

FIG. 2. Determination of the transcription start sites for GPVI. (A) 5′-RACE was performed using the primers SP5 (lanes 2, 3, and 5) and UP (lanes 2, 4, and 5) with (lanes 2–4) or without reverse transcription (lane 5) on poly(A)+ RNA from PMA-treated CMK cells. The products were subjected to 1.2% agarose gel electrophoresis and ethidium bromide staining. Molecular size markers were run in lane 1. (B) The 5′-end-labeled probe (extending from −213 to +72) was annealed to poly(A)+ RNA from PMA-treated CMK cells (lane 3) or no RNA (lane 4). The product was digested by S1 nuclease and separated on a denaturing 8% polyacrylamide gel, followed by autoradiography. The 5′-end-labeled probe (lane 2) and molecular size markers (lane 1) were also run in parallel. The positions of the transcription start sites are shown on the right. (C) The 5′-end of GPVI cDNA and the 5′-flanking region of the gene are shown. The 5′ ends of the 5′-RACE products from A are denoted by asterisks. Five of the ten clones sequenced were positioned at +1. The primer SP5 used in A is indicated by an arrow. The probe used in B is dotted underlined. (D) Comparison of the transcription start sites of megakaryocyte- or platelet-specific genes, including GPVI, human α_{IIb} and α_2 integrins, GPIX, PF4, and GPV genes, with mouse terminal deoxynucleotidyl transferase gene (Tdt). The sequences flanking the +1 position of them are shown. Bold letters represent nucleotides corresponding to the consensus sequence PyPyPyPyCAPyPyPyPy of the Tdt gene.

TABLE I

Sequences at the Exon–Intron Junctions of the GPVI Gene

Exon #	(size)	Sequence at exon/intron junction: 5′ splice donor . . .	(intron size)	. . . 3′ splice acceptor	Exon #
1	(62 bp)	CTG TCT TG gtgagtcctga. . . C L G	(5665 bp)	. . .ctttcctcag G GCT GTG L C	2
2	(31 bp)	GCA GAG TG gtgagtcctt. . . Q S G	(109 bp)	. . .cctcttccag G ACC GCT P L	3
3	(257 bp)	TGC CAC GG gtaaaggaag. . . A T G	(4276 bp)	. . .tttctcccag G AGT TTT V F	4
4	(284 bp)	GGT CAC AG gtaggggtag. . . V T G	(2306 bp)	. . .gatttcccag G AAC CTC T S	5
5	(53 bp)	GGT AGC AG gtaggttctg. . . V A E	(6506 bp)	. . .gctttcttag A ATT CTC F S	6
6	(59 bp)	CAC AAC TG gtgagtaacc. . . T T E	(2001 bp)	. . .tggttcccag A GAC TTC T S	7
7	(50 bp)	TCC AGC TG gtgagtaagt. . . P A G	(517 bp)	. . .cattctccag G TCC TGC P A	8
7′	(54 bp)	TCC AGC TGG TGA gtaagta. . . P A G E	(513 bp)	. . .cattctccag GTC CTG S C	8

Note. Exon 7′ represents a splice variant for GPVI-3.

peptide-2 and -3. Because none of them contained the sequence of peptide-1 and an expected start ATG codon, we performed 5′-RACE on RNA from PMA-stimulated CMK cells with the primer SP5 (Fig. 2C). This yielded a product of about 380 bp (Fig. 2A). We did subcloning the DNA fragment and sequenced ten clones. 5′-end of a half of the sequenced clones started at the same position that we defined as +1, while the other clones started at upstream positions between −1 and −113 (Fig. 2C). These clones encoded peptide-1 and -3, and a putative signal peptide with a starting methionine. Thus, we cloned a full-length cDNA for GPVI of 2260 bp with an open-reading frame and a poly-A signal at the 3′ end (designated as GPVI-1 in Fig. 1). We also isolated two other variant cDNAs for GPVI with the same screening. Some positive clones with the screening described above had 54-base deletion from +657 to +710. Another clone had additional 4 bases (GTGA) from +822. We named the deletion variant as GPVI-2 and the insertion variant as GPVI-3. The coding region of the GPVI-1 cDNA was identical to the cDNA for GPVI that Clemetson *et al.* have recently reported (15), while the 3′-untranslated region (from +1109 to the last) differed from their cDNA. Although the reason why it differs from ours is unknown, our sequence of the 3′-untranslated region was confirmed by the genomic sequence of GPVI described below. We isolated and sequenced the cDNA for GPVI-1 (+1 to +1756) with RT-PCR on poly(A)+ RNA from human platelets and conformed that it was the same as the GPVI-1 cDNA from CMK cells. As previously reported (15), the cDNA for GPVI-1 encoded 339 amino acids including a leader signal sequence, two Ig-C2-like domains, a putative transmembrane domain, and a unique cytoplasmic tail (Fig. 1). The deletion of 18 amino acids in GPVI-2 was between the Ig-C2 domains and the transmembrane domain without a frameshift (Fig. 1). The 4-base insertion in GPVI-3 caused a frame shift resulting in elongation of 361 amino acids with no apparent transmembrane domain following two Ig-C2-like domains. The 3′-untranslated region contained two expressed sequence tags, AA308708 and AA494446, derived from colon cancer cells. This sug-

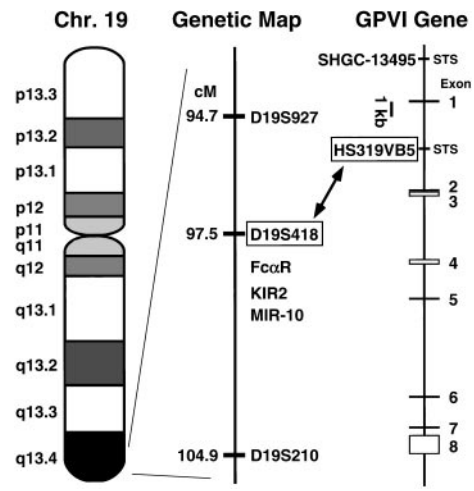


FIG. 4. Chromosomal localization of the GPVI gene. On the middle are shown relative positions of framework markers on the Généthon human genetic linkage map and FcαR, KIR2, and MIR-10 genes on the chromosome 19q13.4. The sequence-tagged sites SHGC-13495 and HS319VB5 are positioned relative to the exons of GPVI on the right. Note that HS319VB5 is derived from D19S418.

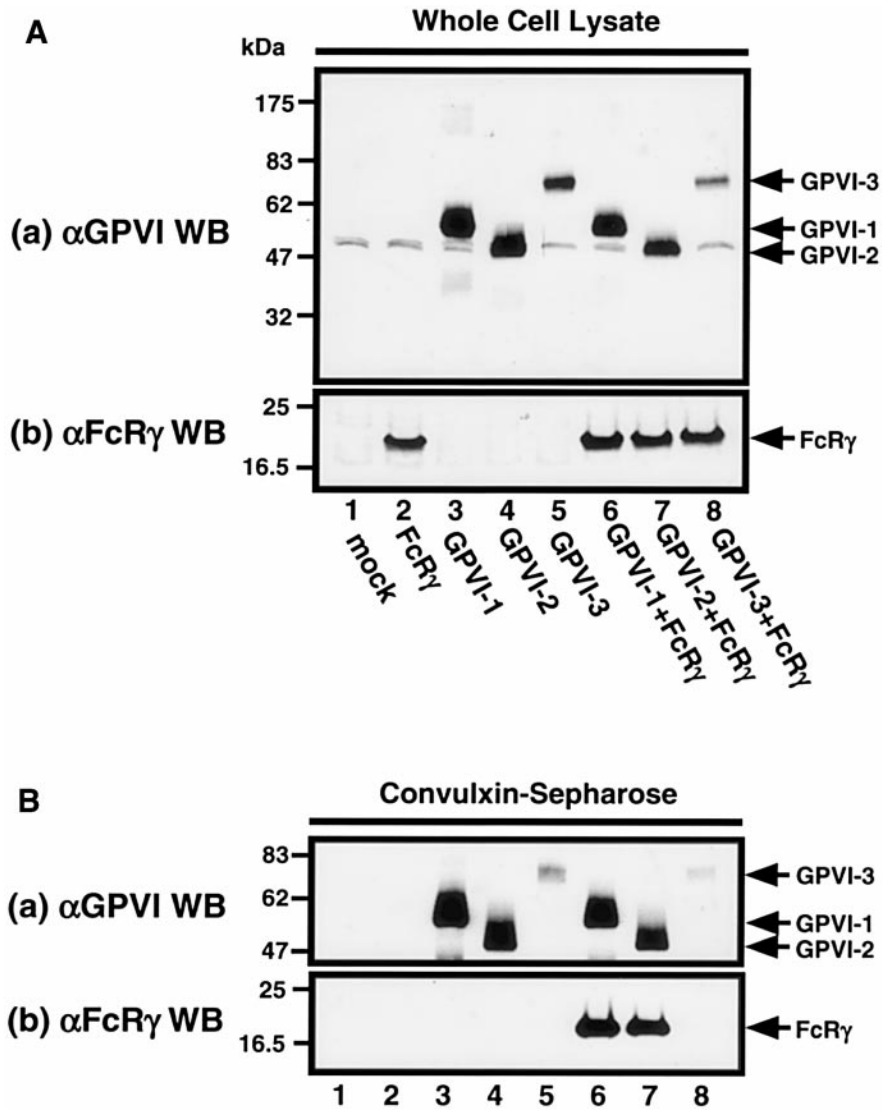


FIG. 5. Western blot analysis of COS-7 cells transfected with the GPVI isoforms and FcRγ. COS-7 cells were transfected with pBK-CMV alone, pBK-CMV-FcRγ, pBK-CMV-GPVI-1, -2, or -3 (lanes 1–5), or were cotransfected with pBK-CMV-FcRγ and pBK-CMV-GPVI-1, -2, or -3 (lanes 6–8). (A) Cells were lysed in SDS sample buffer, resolved on 10% (a) or 12.5% (b) SDS–PAGE, transferred to nitrocellulose membranes, and immunoblotted with anti-GPVI IgG (a) or anti-FcRγ IgG (b). (B) Cells were lysed in Triton X-100-containing lysis buffer. Proteins were precipitated by affinity with convulxin-coupled Sepharose 4B and subjected to immunoblot analysis as described in A. Molecular mass markers are indicated in kDa on the left. Positions of the GPVI isoforms and FcRγ are indicated on the right.

gests that GPVI might be expressed and plays some roles in cancer cells, although GPVI has been reported to be only expressed in platelets and megakaryocytes among normal tissues or cells (10, 23).

Genomic structure and chromosomal localization of GPVI. By searching the database with the GPVI cDNA sequence, we found the full span of GPVI gene in the chromosome 19 clones CTC-550B14 and RP11-700B5, which have been sequenced in the Human Genome Project. CTC-550B14 and RP11-700B5 consisted of unordered 11 and 23 contigs, respectively. A 58-kbp contig in length of CTC-550B14 contained the full GPVI gene consisted of 8 exons spanning over 23 kbp, whereas the GPVI gene divided into three contigs of RP11-700B5 (Fig. 3). Thus, we could determine the genomic structure of the GPVI gene and its 5′-flanking region. The sequences at the all exon–intron junctions followed to the GT/AG rules (Table I). The exon 1 and 2 encoded a signal sequence of 20 amino acids, whereas the exon 3 and 4 encoded Ig-C2-like domains. A mucin-like sequence was encoded in the exon 5–7. The exon 8 of 1460 bp in length encoded a transmembrane domain

and contained a 3′-untranslated region and a poly-A signal. The exon 5 was deleted in the cDNA for GPVI-2 without a frameshift, whereas the GPVI-3 cDNA had a 4-bp insertion at the end of the exon 7, producing a frameshift. It became evident that GPVI-2 and -3 were alternatively spliced forms (Fig. 3). To determine the localization of the GPVI gene on the genome, we searched the contig including the GPVI gene for sequence-tagged site markers with the electric PCR program of the National Center for Biotechnology Information. We detected two sequence-tagged sites, HS319VB5 and SHGC-13495, in the first intron and the 5′-flanking region of the GPVI gene, respectively. HS319VB5 is derived from the microsatellite frame marker D19S418, which has been mapped on the chromosome 19q13.4 on the Généthon genetic map (24), whereas SHGC-13495 is derived from an expressed sequence tag and falls at the same position by RH mapping (25). Thus, the GPVI gene was mapped on the chromosome 19q13.4 (Fig. 4). The genes for FcαR and a series of the NK cell receptors are located near the GPVI gene (25). They all belong to the immunoglobulin superfamily and are closely related.

Determination of transcription start sites for GPVI gene. To confirm the transcription start sites of GPVI, we conducted an S1 nuclease mapping experiment on the mRNA from PMA-treated CMK cells with a probe designed from the 5′ end of GPVI gene (positions −214 to +72) (Figs. 2B and 2C). The results obtained from the S1 nuclease mapping consisted with that of the 5′-RACE described above. The major start site in S1 nuclease mapping was +1 position, whereas the minor sites were positioned at between −13 and −94. The promoter of the GPVI gene lacked TATA and CAAT boxes and had multiple transcription start sites like other megakaryocytic genes, such as the integrin α_{IIb} and α_2, GPIX, PF4, and GPV genes (26–29). The major start sites of those genes including the GPVI gene matched the consensus sequence for initiation found in the mouse terminal transferase gene (Fig. 2D) (26–29).

Transient expression of the GPVI isoforms in COS-7 cells. To characterize the GPVI isoforms that we cloned, we performed a transient expression experiment in COS-7 cells. When COS-7 cells were transfected with GPVI-1, -2, and -3, the GPVI isoforms were expressed with molecular masses of 56, 50, and 70 kDa, respectively, under non-reduced conditions (Fig. 5A), and of 60, 54, and 74 kDa, respectively, under reduced conditions (data not shown) on immunoblot with anti-GPVI antibody. The theoretical molecular masses of GPVI-1, -2, and -3 excluding the signal sequence were 34.9, 33.1, and 65.3 kDa, respectively (Fig. 3). The increase of molecular masses seems to be due to glycosylation of GPVI in COS-7 cells. GPVI has a putative N-glycosylation site in the

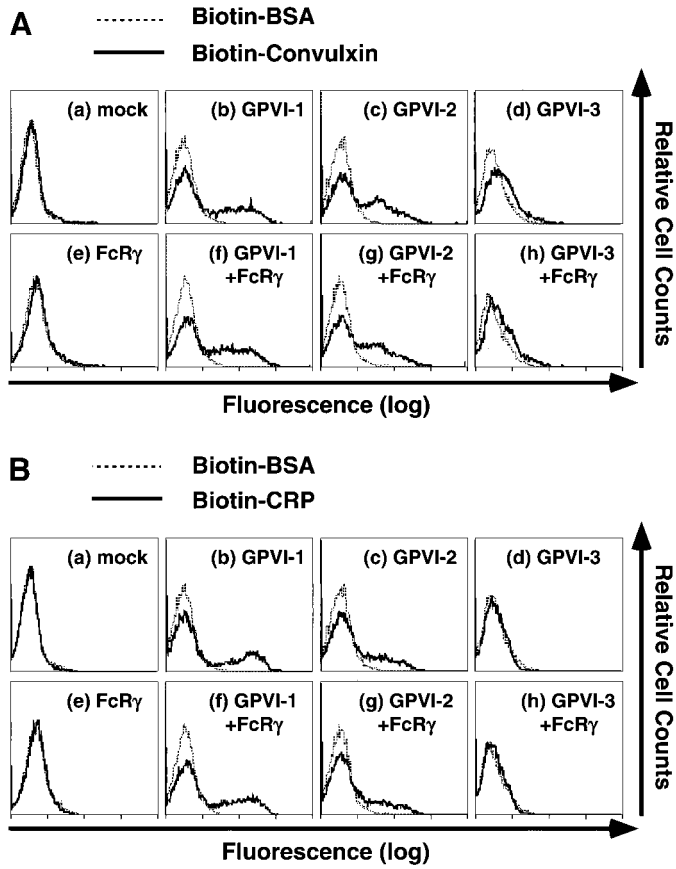


FIG. 6. Flow cytometric analysis of COS-7 cells transfected with the GPVI isoforms and FcRγ. COS-7 cells were transfected with pBK-CMV alone, pBK-CMV-GPVI-1, -2, -3, or pBK-CMV-FcRγ (panels a–e), or were cotransfected with pBK-CMV-FcRγ and pBK-CMV-GPVI-1, -2, or -3 (panels f–h). (A) Transfectants were stained with biotinylated convulxin or bovine serum albumin as a negative control, followed by fluorescein-labeled streptavidin and were analyzed by flow cytometry. (B) Transfectants were stained with biotinylated CRP or bovine serum albumin as a negative control, followed by fluorescein-labeled streptavidin and were analyzed by flow cytometry.

Ig-C2 like domains and the mucin-like region that might be modified by *O*-glycosylation (15). To detect surface expression level of the GPVI isoforms, we developed flow cytometry with biotin-labeled convulxin, which bound to normal platelets but not GPVI-deficient platelets (data not shown). GPVI-1 and -2 were strongly expressed on the surface of COS-7 cells, whereas GPVI-3 was weakly expressed (Fig. 6A). Since GPVI is non-covalently and constitutively associated with FcRγ in platelets (10), we cotransfected COS-7 cells with the GPVI isoforms and FcRγ. The cotransfection of FcRγ did not affect either the protein expression level or surface expression level of the GPVI isoforms in COS-7 cells, judging from immunoblot of whole cell lysates (Fig. 5A) and flow cytometry (Fig. 6A). Convulxin-coupled beads precipitated the GPVI isoforms from the trans-

fected COS-7 cells (Fig. 5B). FcRγ was coprecipitated with GPVI-1 and -2 but not with GPVI-3 (Fig. 5B). GPVI-1 and -2 have a positively charged arginine in the transmembrane domain that seems to play an important role in association with FcRγ, as reported on the Fc receptors associated with FcRγ (30), whereas GPVI-3 did not have a distinct transmembrane domain. Although FcRγ is known to enhance expression of FcϵRI and FcγRIIIA, but not FcγRI and FcαR among the Fc receptors (30), the expression levels of the GPVI isoforms were not affected by FcRγ in our transfection experiment. CRP exhibits triple-helical and polymeric structure of collagen and activates platelets via GPVI independent of $\alpha_2\beta_1$ (7–9). We determined the ability of the GPVI isoforms to bind CRP by flow cytometry with biotin-labeled CRP. We confirmed that biotin-labeled CRP bound to normal platelets but not GPVI-deficient platelets (data not shown). Biotin-labeled CRP bound to GPVI-1 and -2 expressed on COS-7 cells, irrespectively of FcRγ, whereas GPVI-3 did not bind CRP (Fig. 6B). Thus, GPVI-1 and -2 by themselves had ability to recognize and bind the triple-helical structure of collagen that CRP mimics.

We reported the full-length cDNA for the GPVI isoforms, the genomic structure, the transcription start sites, the chromosomal localization, and transfection experiments in COS-7 cells. Although physiological and pathological roles of the GPVI isoforms remain to be investigated, this study provides the molecular basis for further understanding the collagen receptor GPVI.

ACKNOWLEDGMENTS

We thank Drs. Etsuya Matsutani, Sei Yoshida, Yoshiyuki Kaneko, and Keisuke Shindoh for helpful advice, Dr. Takeyuki Sato for a kind gift of CMK cells, and Ms. Ikuyo Nakamura for secretarial assistance. This work was supported by a Grant-in-Aid for Scientific Research from the Ministry of Education, Science, and Culture of Japan and Sankyo Foundation of Life Science.

REFERENCES

1. Alberio, L., and Dale, G. L. (1999) Platelet-collagen interactions: Membrane receptors and intracellular signalling pathways. *Eur. J. Clin. Invest.* **29,** 1066–1076.
2. Sugiyama, T., Okuma, M., Ushikubi, F., Sensaki, S., Kanaji, K., and Uchino, H. (1987) A novel platelet aggregating factor found in a patient with defective collagen-induced platelet aggregation and autoimmune thrombocytopenia. *Blood* **69,** 1712–1720.
3. Ichinohe, T., Takayama, H., Ezumi, Y., Yanagi, S., Yamamura, H., and Okuma, M. (1995) Cyclic AMP-insensitive activation of c-Src and Syk protein-tyrosine kinases through platelet membrane glycoprotein VI. *J. Biol. Chem.* **270,** 28029–28036.
4. Ichinohe, T., Takayama, H., Ezumi, Y., Arai, M., Yamamoto, N., Takahashi, H., and Okuma, M. (1997) Collagen-stimulated activation of Syk but not c-Src is severely compromised in human platelets lacking membrane glycoprotein VI. *J. Biol. Chem.* **272,** 63–68.
5. Polgár, J., Clemetson, J. M., Kehrel, B. E., Wiedemann, M., Magnenat, E. M., Wells, T. N. C., and Clemetson, K. J. (1997) Platelet activation and signal transduction by convulxin, a C-type lectin from Crotalus durissus terrificus (tropical rattlesnake) venom via the p62/GPVI collagen receptor. *J. Biol. Chem.* **272,** 13576–13583.
6. Jandrot-Perrus, M., Lagrue, A.-H., Okuma, M., and Bon, C. (1997) Adhesion and activation of human platelets induced by convulxin involve glycoprotein VI and integrin $\alpha_2\beta_1$. *J. Biol. Chem.* **272,** 27035–27041.
7. Morton, L. F., Hargreave, P. G., Farndale, R. W., Young, R. D., and Barnes, M. J. (1995) Integrin alpha 2 beta 1-independent activation of platelets by simple collagen-like peptides: Collagen tertiary (triple-helical) and quaternary (polymeric) structures are sufficient alone for alpha 2 beta 1-independent platelet reactivity. *Biochem. J.* **306,** 337–344.
8. Asselin, J., Gibbins, J., Achison, M., Lee, Y. H., Morton, L. F., Farndale, R. W., Barnes, M. J., and Watson, S. P. (1997) A collagen-like peptide stimulates tyrosine phosphorylation of syk and phospholipase Cγ2 in platelets independent of the integrin $\alpha_2\beta_1$. *Blood* **89,** 1235–1242.
9. Kehrel, B., Wierwille, S., Clemetson, K. J., Anders, O., Steiner, M., Knight, C. G., Farndale, R. W., Okuma, M., and Barnes, M. J. (1998) Glycoprotein VI is a major collagen receptor for platelet activation: It recognizes the platelet-activating quaternary structure of collagen, whereas CD36, glycoprotein IIb/IIIa, and von Willebrand factor do not. *Blood* **91,** 491–499.
10. Tsuji, M., Ezumi, Y., Arai, M., and Takayama, H. (1997) A novel association of Fc receptor gamma-chain with glycoprotein VI and their co-expression as a collagen receptor in human platelets. *J. Biol. Chem.* **272,** 23528–23531.
11. Ezumi, Y., Shindoh, K., Tsuji, M., and Takayama, H. (1998) Physical and functional association of the Src family kinases Fyn and Lyn with the collagen receptor glycoprotein VI-Fc receptor γ chain complex on human platelets. *J. Exp. Med.* **188,** 267–276.
12. Poole, A., Gibbins, J. M., Turner, M., van Vugt, M. J., van de Winkel, J. G. J., Saito, T., Tybulewicz, V. L. J., and Watson, S. P. (1997) The Fc receptor γ-chain and the tyrosine kinase Syk are essential for activation of mouse platelets by collagen. *EMBO J.* **16,** 2333–2341.
13. Gibbins, J. M., Briddon, S., Shutes, A., van Vugt, M. J., van de Winkel, J. G. J., Saito, T., and Watson, S. P. (1998) The p85 subunit of phosphatidylinositol 3-kinase associates with the Fc receptor γ-chain and linker for activitor of T cells (LAT) in platelets stimulated by collagen and convulxin. *J. Biol. Chem.* **273,** 34437–34443.
14. Gross, B. S., Lee, J. R., Clements, J. L., Turner, M., Tybulewcz, V. L. J., Findell, P. R., Koretzky, G. A., and Watson, S. P. (1999) Tyrosine phosphorylation of SLP-76 is downstream of Syk following stimulation of the collagen receptor in platelets. *J. Biol. Chem.* **274,** 5963–5971.
15. Clemetson, J. M., Polgar, J., Magnenat, E., Wells, T. N. C., and Clemetson, K. J. (1999) The platelet collagen receptor glycoprotein VI is a member of the immunoglobulin superfamily closely related to FcαR and the natural killer receptors. *J. Biol. Chem.* **274,** 29019–29024.
16. Letourneur, O., Kennedy, I. C. S., Brini, A. T., Ortaldo, J. R., O'Shea, J. J., and Kinet, J.-P. (1991) Characterization of the family of dimers associated with Fc receptors (FcϵRI and FcγRIII). *J. Immunol.* **147,** 2652–265.
17. Sato, T., Fuse, A., Eguchi, M., Hayashi, Y., Ryo, R., Adachi, M., Kishimoto, Y., Teramura, M., Mizoguchi, H., Shima, Y., Komori, I., Sunami, S., Okimoto, Y., and Nakajima, H. (1989) Establishment of a human leukaemic cell line (CMK) with megakaryocytic characteristics from a Down's syndrome patient with acute megakaryoblastic leukaemia. *Br. J. Haematol.* **72,** 184–190.

18. Komatsu, N., Suda, T., Moroi, M., Tokuyama, N., Sakata, Y., Okada, M., Nishida, T., Hirai, Y., Sato, T., Fuse, A., and Miura, Y. (1989) Growth and differentiation of a human megakaryoblastic cell line, CMK. *Blood* **74,** 42–48.
19. Tandon, N. N., Lipsky, R. H., Burgess, W. H., and Jamieson, G. A. (1989) Isolation and characterization of platelet glycoprotein IV (CD36). *J. Biol. Chem*. **264,** 7570–7575.
20. Frohman, M. A., Dush, M. K., and Martin, G. R. (1988) Rapid production of full-length cDNAs from rare transcripts: Amplification using a single gene-specific oligonucleotide primer. *Proc. Natl. Acad. Sci. USA* **85,** 8998–9002.
21. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
22. Harlow, E., and Lane, D. (1988) Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
23. Jandrot-Perrus, M., Busfield, S., Lagrue, A. H., Xiong, X., Debili, N., Fraser, C., Vainchenker, W., and Villeval, J. L. (1999) Platelet glycoprotein VI: cloning, structure, and expression of a functional protein. *Blood* **94** (Suppl. 1), 371a.
24. Dib, C., Faure, S., Fizames, C., Samson, D., Drouot, N., Vignal, A., Millasseau, P., Marc, S., Hazan, J., Seboun, E., Lathrop, M., Gyapay, G., Morissette, J., and Weissenbach, J. (1996) A comprehensive genetic map of the human genome based on 5,264 microsatellites. *Nature* **380,** 152–154.
25. Deloukas, P., Schuler, G. D., Gyapay, G., Beasley, E. M., Soderlund, C., Rodriguez-Tome, P., Hui, L., Matise, T. C., McKusick, K. B., Beckmann, J. S., Bentolila, S., Bihoreau, M., Birren, B. B., Browne, J., Butler, A., Castle, A. B., Chiannilkulchai, N., Clee, C., Day, P. J., Dehejia, A., Dibling, T., Drouot, N., Duprat, S., Fizames, C., Fox, S., Gelling, S., Green, L., Harrison, P., Hocking, R., Holloway, E., Hunt, S., Keil, S., Lijnzaad, P., Louis-Dit-Sully, C., Ma, J., Mendis, A., Miller, J., Morissette, J., Muselet, D., Nusbaum, H. C., Peck, A., Rozen, S., Simon, D., Slonim, D. K., Staples, R., Stein, L. D., Stewart, E. A., Suchard, M. A., Thangarajah, T., Vega-Czarny, N., Webber, C., Wu, X., Hudson, J., Auffray, C., Nomura, N., Sikela, J. M., Polymeropoulos, M. H., James, M. R., Lander, E. S., Hudson, T. J., Myers, R. M., Cox, D. R., Weissenbach, J., Boguski, M. S., and Bentley, D. R. (1998) A physical map of 30,000 human genes. *Science* **282,** 744–746.
26. Uzan, G., Prenant, M., Prandini, M. H., Martin, F., and Marguerie, G. (1991) Tissue-specific expression of the platelet GPIIb gene. *J. Biol. Chem*. **266,** 8932–8939.
27. Zutter, M. M., Santoro, S. A., Painter, A. S., Tsung, Y. L., and Gafford, A. (1994) The human alpha 2 integrin gene promoter. Identification of positive and negative regulatory elements important for cell-type and developmentally restricted gene expression. *J. Biol. Chem*. **269,** 463–469.
28. Hickey, M. J., and Roth, G. J. (1993) Characterization of the gene encoding human platelet glycoprotein IX. *J. Biol. Chem*. **268,** 3438–3443.
29. Ravanat, C., Morales, M., Azorsa, D. O., Moog, S., Schuhler, S., Grunert, P., Loew, D., Van Dorsselaer, A., Cazenave, J. P., and Lanza, F. (1997) Gene cloning of rat and mouse platelet glycoprotein V: Identification of megakaryocyte-specific promoters and demonstration of functional thrombin cleavage. *Blood* **89,** 3253–3262.
30. Daëron, M. (1997) Fc receptor biology. *Annu. Rev. Immunol.* **15,** 203–234.