Volume 274, number 1,2, 189-192 FEBS 09100 November 1990

Human platelet glycoprotein IX

Characterization of cDNA and localization of the gene to chromosome 3

Mark J. Hickey[1], Larry L. Deaven[2] and Gerald J. Roth[1]

[1]*Seattle Veterans Administration Medical Center (111), 1660 S. Columbian Way, Seattle, WA 98108, USA* and [2]*Life Sciences Division, Los Alamos National Laboratory, L.S. -4; Mail Stop M8888, Los Alamos, NM 87545, USA*

Received 30 May 1990; revised version received 8 September 1990

Overlapping cDNAs encoding human platelet glycoprotein (Gp)IX were cloned from a human erythroleukemia cell λgt11 library. The possibly 'full-length' cDNA of 896 base pairs (bp) includes an open reading frame (528 bp), both 5′ (222 bp) and 3′ (146 bp) noncoding regions, and a poly(A) tail. Translation predicts a signal peptide of 16 amino acids and a mature protein of 160 amino acids that includes a 24 amino acid leucine-rich glycoprotein (LRG) segment. Southern blot analysis suggests the presence of a single copy of the Gp IX gene, and hybridization of Gp IX cDNA to sorted human chromosomes localizes the Gp IX gene to chromosome 3.

cDNA sequence; Human platelet glycoprotein IX; Leucine-rich glycoprotein family; Chromosomal localization

1. INTRODUCTION

Glycoprotein IX (Gp IX) is a small membrane glycoprotein found on the human platelet surface. Originally described by size (M_r 17 000) and p*I* (pH 5.8–6.3) [1,2], Gp IX forms a 1:1 non-covalent complex with Gp Ib [3], a heterodimeric transmembrane protein consisting of a disulfide linked 140 kDa α chain and 22 kDa β chain. Platelet adhesion to arterial blood vessels is mediated by the Gp Ib/Gp IX complex, which serves as the platelet receptor for von Willebrand factor (vWf), with the Gp Ib α chain providing the vWf binding site [4,5]. The individual contribution of Gp IX to adherence is unknown.

Patients with a congenital bleeding disorder, Bernard-Soulier syndrome, have platelets lacking all the components of the Gp Ib/Gp IX complex and also a third Gp, GpV (M_r 82 000) [1]. The molecular defect responsible for these deficiencies is unknown, but a recent report indicates that in one case, a truncated Gp Ib α chain is correlated with the disease state [6].

The primary structure for the Gp Ib α and β chains has been determined [5,7,8]. We have also recently described a partial cDNA for Gp IX with an open reading frame coding for part of the signal sequence and all of the mature Gp IX [9]. Gp Ib α, Gp Ib β, and Gp IX all contain sequences similar to the tandem repeats in leucine-rich α_2-glycoprotein (LRG) of human serum [10] and have one or more N-linked glycosylation sites. We now report the characterization of clones establishing the complete primary sequence of Gp IX and identify the chromosome containing the Gp IX gene.

Correspondence address: G. Roth, Seattle VAMC (111), 1660 S. Columbian Way, Seattle, WA 98108, USA

Abbreviations: Gp, glycoprotein; bp, base pairs; LRG, leucine-rich glycoprotein; BSA, bovine serum albumin; PVP, polyvinylpyrolidone; HEL, human erythroleukemia; SDS, sodium dodecyl sulfate; kDa, kilodaltons; vWf, von Willebrand factor

The nucleotide sequence reported here has been submitted to the EMBL/Genbank database under the accession number X52997

2. MATERIALS AND METHODS

2.1. *Materials*

[α-^{32}P]dCTP (3000 Ci/mmol) was purchased from New England Nuclear (Boston, MA). Nylon membranes and random priming reagents were obtained from Amersham (Arlington Heights, IL). 1× SSC is 0.15 M NaCl, 15 mM Na_3-citrate. 1× Denhardt's solution is 0.1% BSA, 0.1% PVP, 0.1% Ficoll.

2.2. *Screening and sequencing*

The previously isolated Gp IX cDNA (λH7, Fig. 1) was labelled with [α-^{32}P]dCTP to a specific activity of 1×10^8 cpm/μg by random priming [11], and used to screen 500 000 recombinants of a λgt11 cDNA library from HEL cells [7]. Nitrocellulose filters were hybridized at 42°C for 16 h in 50% formamide, 5× SSC, 1× Denhardt's solution, 25 mM $NaHPO_4$, pH 7.0, 250 μg/ml yeast tRNA, 10% dextran sulfate, and washed successively at 25°C in 2× SSC/0.1% SDS, and 52°C in 0.1× SSC/0.1% SDS [12]. Hybridization positive phage were plaque purified, and cDNA inserts were digested, subcloned, and sequenced as previously described [9].

2.3. *Northern blotting*

Isolation of total RNA from human platelets and macrophages, separation on denaturing agarose gels, and transfer to nitrocellulose have been described [13]. Probe labelling, hybridization, and wash conditions were as described in section 2.2.

2.4. *Southern blotting*

Genomic DNA was isolated from normal human lymphocytes,

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/90/$3.50

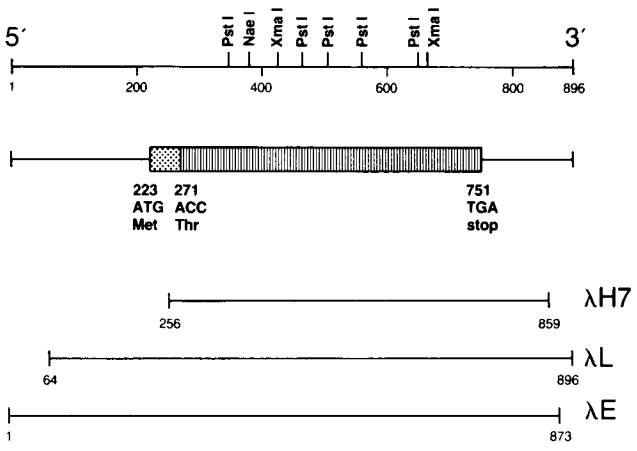

Fig. 1. Derivation of the Gp IX cDNA sequence from three inserts. A restriction map is shown (top line, numbers refer to nucleotide sequence) above a diagram depicting the open reading frame as boxes (dots = signal peptide; vertical lines = mature protein), noting the first base of the start (ATG) and stop (TGA) codons and the first codon of mature Gp IX (ACC). Three individual cloned inserts (λH7-previously reported [9], λL, and λE), were sequenced.

digested overnight with restriction enzymes, separated on a 1% agarose gel and transferred to nylon membranes [12]. Probe labelling, hybridization and wash conditions were identical to those described in section 2.2 except that 0.1% SDS was included in the hybridization solution.

2.5. *Chromosomal localization*

Human chromosomes were isolated and sorted directly on nitrocellulose discs as previously described [14,15] and hybridized to labelled Gp IX cDNA.

3. RESULTS AND DISCUSSION

3.1. *Isolation and characterization of cDNA inserts*

Screening 5×10^5 recombinants from the HEL cell library yielded 16 positive clones of which 8 were plaque purified. All contained inserts between 500 and 850 bp in size. The two longest clones, designated λE and λL (Fig. 1) were subcloned, and their sequences were determined on both strands, resulting in the nucleotide and deduced amino acid sequences shown (Fig. 2). The 5′ non-coding region is unusually long (222 bp); however, the putative initiator Met codon (nt 223–235) is 30 bp downstream from an in-frame stop codon (nt 193–195), precluding the existence of a longer translation product. Sequence alignment of Gp IX and Gp Ib α chain shows striking similarity in the nucleotides immediately preceding their initiator Met codons (12/14 identical). This is potentially significant since this region contains the only intron in the entire gpIbα gene [16,17].

The 16 amino acid signal peptide is typically hydrophobic and ends with a sequence consistent with a cleavage site (A-X-A) [18]. The polyadenylation signal AATAAA (nt 853–858) is 27 nt upstream from a poly(A_{11}) tail. Hydropathy analysis of the complete sequence indicates that Gp IX is a transmembrane protein (Fig. 3).

```
                CAGCTGTATCCCATAGAGTTGCCACCCAGGCCTCAGCCAGGACCTTTCA      49

GGCCAGACAGGAGCACCTGACCAAAGGCTTCACAGCCGCCCTCACCGCCCGGCCTTCTA     108

CGGTGTCCAGAGACAGTTAGCCAGGCCTGGGCTGGGCACACTCCACCTTCCCTAGTCAC     167

CAGCTGGTTTCCCAGAGGAGAAGGCTGAGACCCGAGAAGGGAGCCAGCCTGTCCC ATG     225
                                                          M      -16

CCT GCC TGG GGA GCC CTG TTC CTG CTC TGG GCC ACA GCA GAG GCC      270
 P   A   W   G   A   L   F   L   L   W   A   T   A   E   A       -1

ACC AAG GAC TGC CCC AGC CCA TGT ACC TGC CGC GCC CTG GAA ACC      315
 T   K   D   C   P   S   P   C   T   C   R   A   L   E   T       15

ATG GGG CTG TGG GTG GAC TGC AGG GGC CAC GGA CTC ACG GCC CTG      360
 M   G   L   W   V   D   C   R   G   H   G   L   T   A   L       30

CCT GCC CTG CCG GCC CGC ACC CGC CAC CTT CTG CTG GCC AAC AAC      405
 P   A   L   P   A   R   T   R   H   L   L   L   A   N   N       45
                                                      ◆

AGC CTT CAG TCC GTG CCC CCG GGA GCC TTT GAC CAC CTG CCC CAG      450
 S   L   Q   S   V   P   P   G   A   F   D   H   L   P   Q       60

CTG CAG ACC CTC GAT GTG ACG CAG AAC CCC TGG CAC TGT GAC TGC      495
 L   Q   T   L   D   V   T   Q   N   P   W   H   C   D   C       75

AGC CTC ACC TAT CTG CGC CTC TGG CTG GAG GAC CGC ACG CCC GAG      540
 S   L   T   Y   L   R   L   W   L   E   D   R   T   P   E       90

GCC CTG CTG CAG GTC CGC TGT GCC AGC CCC AGC CTC GCT GCC CAT      585
 A   L   L   Q   V   R   C   A   S   P   S   L   A   A   H      105

GGC CCG CTG CGG CTG ACA GGC TAC CAG CTG GGC AGC TGT GGC TGG      630
 G   P   L   R   L   T   G   Y   Q   L   G   S   C   G   W      120

CAG CTG CAG GCG TCC TGG GTG CGC CCG GGG GTC TTG TGG GAC GTG      675
 Q   L   Q   A   S   W   V   R   P   G   V   L   W   D   V      135

GCG CTG GTC GCC GTG GCC GCG CTG GGC CTG GCT CTT CTG GCT GGC      720
 A   L   V   A   V   A   A   L   G   L   A   L   L   A   G      150

CTG CTG TGT GCC ACC ACA GAG GCC CTG GAT TGA GCCAGGCCCCCAGAA      768
 L   L   C   A   T   T   E   A   L   D stop                    160

CCTGGCTCCAGCCAGGGGGCCAGTCCCTGAGGCAGGTCCCCAGACTCCACCAAGCCTGG     827

TCAGCCCAAACCACCAGAAGCCCAGAATAAACTGGCAGCTCAGCTGTTTTATATAAGCA     886

AAAAAAAAAA                                                      896
```

Fig. 2. Nucleotide and deduced amino acid sequence of human Gp IX. The amino terminus of the mature protein starts with +1; signal peptide is indicated by negative numbers. Amino acids confirmed by Edman degradation [20] are underlined, and the putative transmembrane domain is double underlined. LRG sequence is shown as a dashed box and a potential polyadenylation signal is overlined. An N-linked glycosylation site is marked (◆).

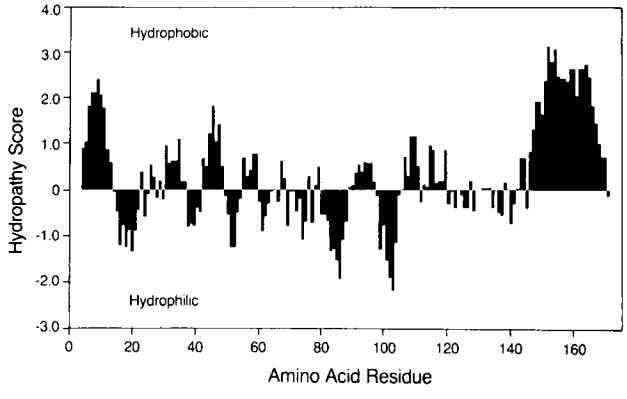

Fig. 3. Hydropathy plot of Gp IX. The complete amino acid sequence of Gp IX is analyzed by the method of Kyte and Doolittle [21] with a 7 residue window. The two major hydrophobic peaks correspond to the signal peptide (1–16) and a putative transmembrane domain (146–170).

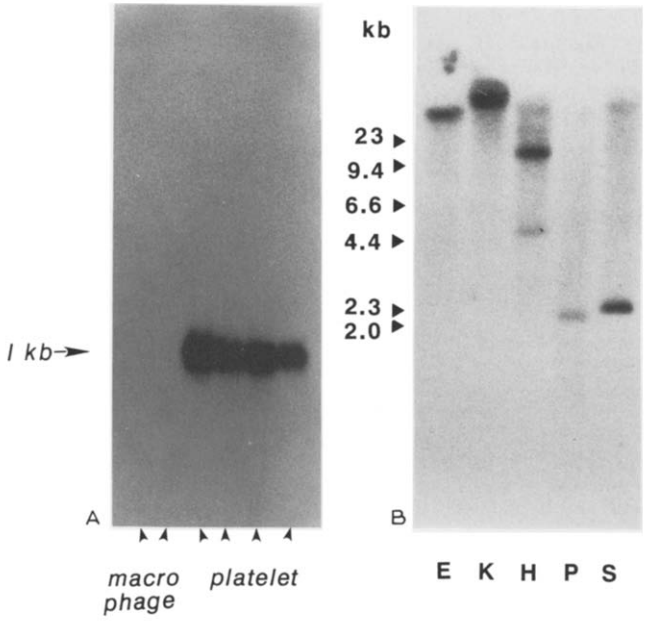

Fig. 4. (A) Northern blot analysis of the Gp IX transcript. Total RNA (12 μg) from human alveolar macrophages (lanes 1 and 2) and human platelets (lanes 3–6) was hybridized with ^{32}P-labelled Gp IX cDNA (λH7). (B) Southern blot analysis of Gp IX DNA. Human genomic DNA (15 μg) was digested with the indicated restriction enzymes and hybridized with ^{32}P-labelled Gp IX cDNA (λE and λL). E = *Eco*RI; K = *Kpn*I; H = *Hin*dIII; P = *Pst*I; S = *Sac*I. The high molecular weight band in the *Hin*dIII lane is a partial digestion product (data not shown). Sizes of hybridizing bands were estimated by comparison to *Hin*dIII digested λ phage DNA.

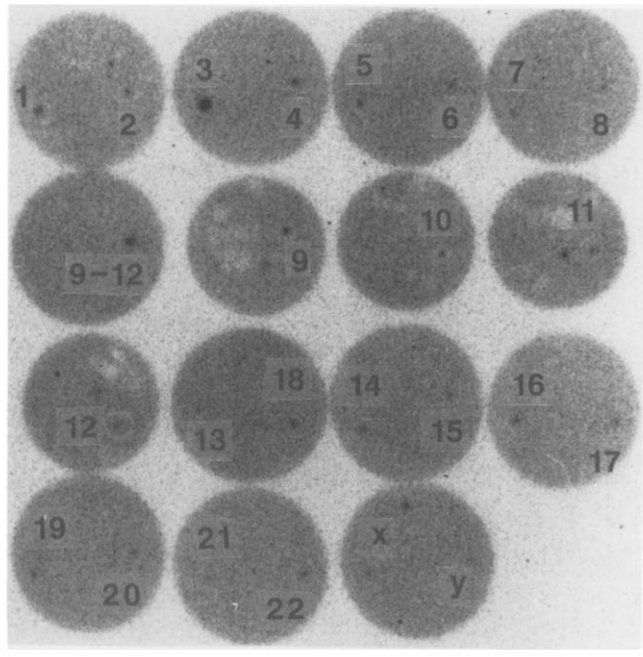

Fig. 5. Spot blot analysis of flow-sorted human chromosomes hybridized with ^{32}P-labelled Gp IX cDNA (λH7). DNA from chromosome 3 gives the only hybridization signal above background level.

Acknowledgements: This work was supported by a Merit Review Grant from the Veteran's Administration (G.J.R.) and Grant HL 39947 from the National Institutes of Health (G.J.R.)

3.2. *Blot analysis of Gp IX*

The cDNA sequence for Gp IX appears to be nearly 'full-length', comparing its size (896 bp) with that of the Gp IX transcript detected by Northern blot analysis of platelet mRNA (Fig. 4A). However, the transcriptional start site was not identified. This transcript is not found in human macrophages (Fig. 4A), or neutrophils [13] but is present in HEL cells, which have megakaryocytic and platelet features. This evidence supports the idea that Gp IX is a platelet-specific protein.

Southern blot analysis of human genomic DNA with a Gp IX probe suggests that Gp IX is a small, single copy gene similar to Gp Ib α, since the probe containing the entire Gp IX transcript hybridizes with a single $\cong$ 2.0 kb *Sac*I fragment (Fig. 4B). However, the small size and relative lack of intron sequences may not be characteristic of all LRG proteins, since the gene for chaoptin, another LRG protein found in *Drosophila*, contains 12 exons spanning over 6 kb [19].

Hybridization of labelled Gp IX cDNA to individual human chromosomes is shown in Fig. 5, and multiple hybridizations and exposures indicate that the Gp IX gene is located on chromosome 3. Since the gpIbα gene is found on chromosome 17 [17], the genes are not part of a complex but are dispersed on separate chromosomes.

REFERENCES

[1] Clemetson, K.J., McGregor, J.L., James, E., Dechavanne, M. and Luscher, E.F. (1982) J. Clin. Invest. 70, 304–311.
[2] Berndt, M.C., Gregory, C., Chong, B.H., Zola, M. and Castaldi, P.A. (1983) Blood 62, 800–807.
[3] Du, X., Beutler, L., Ruan, C., Castaldi, P.A. and Berndt, M.C. (1987) Blood 69, 1524–1527.
[4] Handa, M., Titani, K., Holland, L.Z., Roberts, J.R. and Ruggeri, Z.M. (1986) J. Biol. Chem. 261, 12579–12585.
[5] Titani, K., Takio, K., Handa, M. and Ruggeri, Z.M. (1987) Proc. Natl. Acad. Sci. USA 84, 5610–5614.
[6] Ware, J., Russell, S.R., Vicente, V., Scharf, R.E., Tomer, A., McMillan, R. and Ruggeri, Z.M. (1990) Proc. Natl. Acad. Sci. USA 87, 2026–2030.
[7] Lopez, J.A., Chung, D.W., Fujikawa, K., Hagen, F.S., Papayannopoulou, T. and Roth, G.J. (1987) Proc. Natl. Acad. Sci. USA 84, 5615–5619.
[8] Lopez, J.A., Chung, D.W., Fujikawa, K., Hagen, F.S., Davie, E.W. and Roth, G.J. (1988) Proc. Natl. Acad. Sci. USA 85, 2135–2139.
[9] Hickey, M.J., Williams, S.A. and Roth, G.J. (1989) Proc. Natl. Acad. Sci. USA 86, 6773–6777.
[10] Takahashi, N., Takahashi, Y. and Putnam, F.W. (1985) Proc. Natl. Acad. Sci. USA 82, 1906–1910.
[11] Feinberg, A.P. and Vogelstein, B. (1984) Anal. Biochem. 137, 266–267.
[12] Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
[13] Roth, G.J., Hickey, M.J., Chung, D.W. and Hickstein, D.D. (1989) Biochem. Biophys. Res. Commun. 160, 705–710.

[14] Lebo, R.V., Gorin, F., Fletterick, R.J., Kao, F.T., Cheung, M.C., Bruce, B.D. and Kan, Y.W. (1984) Science 225, 57-59.
[15] Deaven, L.L., Van Dilla, M.A., Bartholdi, M.F., Carrano, A.V., Cram, L.S., Fuscoe, J.C., Gray, J.W., Hildebrand, C.E., Moyzis, R.K. and Perlman, J. (1986) Cold Spring Harb. Symp. Quant. Biol. 51, 159-167.
[16] Wenger, R.H., Kieffer, N., Wicki, A.N. and Clemetson, K.J. (1988) Biochem. Biophys. Res. Commun. 156, 389-395.
[17] Wenger, R.H., Wicki, A.N., Kieffer, N., Adolph, S., Hamiester, H. and Clemetson, K.J. (1989) Gene 85, 517-524.
[18] von Heijne, G. (1985) J. Mol. Biol. 184, 99-105.
[19] Reinke, R., Krantz, D.E., Yen, D. and Zipursky, S.L. (1988) Cell 52, 291-301.
[20] Roth, G.J., Ozols, J., Nugent, D.J. and Williams, S.A. (1988) Biochem. Biophys. Res. Commun. 156, 931-939.
[21] Kyte, J. and Doolittle, R.F. (1982) J. Mol. Biol. 157, 105-132.