JCI The Journal of Clinical Investigation

# The wanderings of a platelet gene: what is "neo" Ib beta telling us?

G J Roth

*J Clin Invest.* 1994;93(6):2301-2302. https://doi.org/10.1172/JCI117231.

Research Article



In this issue of *The Journal,* a new form of platelet GPIbβ is described in heart and brain endothelia (1). This GPIbβ-related protein, referred to here as "neo" Ibβ, is both larger and more complex than its previously described relative in platelets (2). What is this unique product, and what are its implications?

### *The observation*

Platelet GPIbβ is the smaller chain of the surface GPIb heterodimer comprised of GPIbα (molecular mass 143 kD) disulfide-linked to GPIbβ (molecular mass 22 kD). Some nonmegakaryocytic cells (endothelia) express the GPIbα gene in response to certain exogenous stimuli (3). Platelets use GPIbα + β as the receptor for vWf and adhere to injured vascular surfaces through the highly specific GPIb–vWf interaction (4). GPs Ibα and Ibβ along with their associated GPs, V and IX, make up a highly interrelated adhesive subgroup (GPIb-V-IX) of the leucine-rich (LR) family of proteins (4, 5).

In the current study, a larger than expected transcript (3.5-kb neo Ibβ) was detected in endothelial cell mRNA by a platelet GPIbβ cDNA probe, and screening of an endothelial cell cDNA library with the same probe led to a 2.8-kb neo Ibβ cDNA sequence, obviously much larger than the 1-kb sequence encoding platelet GPIbβ. The neo Ibβ transcript appears to start from an upstream, currently unidentified site (perhaps 1 kb 5′ of the GPIbβ site) and to contain a "complex" 1,233-base open reading frame (ORF) although the new ATG start codon is, again, unidentified (Fig. 1). In expanding the platelet GPIbβ cDNA, the complex neo Ibβ ORF adds a 5′ extension to it, reads through its 5′ untranslated region (UTR), shifts the reading frame of the first three codons of its ORF, and retains a genomic intron as an exon. At the 3′ end, the neo Ibβ ORF is "shifted back" to be "in frame" with the platelet ORF, making the COOH terminus of the neo Ibβ protein identical to that of the platelet protein. The elongated 3′ UTR of the neo Ibβ cDNA extends beyond that for platelet GPIbβ, implying that endothelia do not recognize the platelet GPIbβ polyadenylation site. The neo Ibβ transcript may lack a poly A tail, or an additional 3′ UTR sequence may have gone undetected. Further analysis of cDNAs can clarify this point.

The predicted 411–amino acid neo Ibβ translation product is encoded by different regions of cDNA as shown in Fig. 1. Using an antibody to platelet GPIbβ, an appropriately sized 45-kD protein was detected in an endothelial cell lysate by Western blotting, indicating that the neo Ibβ transcript produces protein. A useful confirmatory study would entail probing the same Western blot with an antibody to the nonplatelet portion of the protein. The predicted protein retains the platelet GPIbβ transmembrane domain and most of its signal peptide, which now becomes an internal hydrophobic region that escapes cleavage by signal peptidase. However, neo Ibβ protein lacks an $NH_2$-terminal signal peptide. Because of these unusual features, work to define both the degree of glycosylation and the cellular location of the neo Ibβ protein (intra- versus extracellular) will be informative, along with a definition of the transcriptional start site, the start codon, and the intron/exon structure of the neo Ibβ gene.

### *Implications*

*A novel form of genetic variation.* The elongated and enlarged neo Ibβ version of platelet GPIbβ shows a novel approach by which a cell can obtain a new protein product from an existing gene locus. Instead of shuffling exons or duplicating/diverging from a progenitor gene, nonmegakaryocytic cells read the platelet GPIbβ gene locus by starting transcription well upstream and by including a "former" intron in the transcript. Such use of an alternative promoter and alternative splicing of a transcript are well-known mechanisms for generating new but related products. However, these mechanisms in combination with a shift in reading frame and translation from a promoter, 5′ UTR, and intron are probably familiar only to virologists. As a result, a new and apparently useful protein is produced by extending and modifying a genomic template. This may be the first such example of a nonviral gene product being formed by such means. Perhaps, this will prove to be an efficient and even a general mechanism of genetic variation, and additional examples of it may come to light in other eukaryotic genes.

*Elaborating the LR-LRG motif.* The leucine-rich/leucine-rich glycoprotein (LR/LRG) family of proteins is distinguished by a common, conserved 24–amino acid segment that is often present in a series of tandem repeats. Although the segment's function is uncertain, it appears to participate in protein/protein interactions. Members of the family are found in many species, tissues, and organs, indicating that the LR/

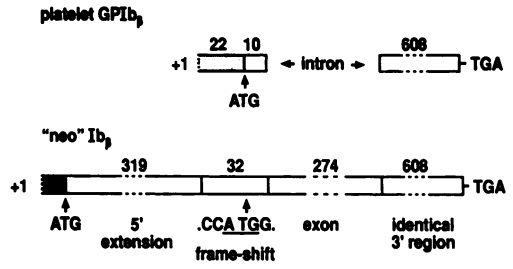


*Figure 1.* Diagram of cDNA encoding platelet GPIbβ and neo Ibβ. The 5′ UTR (*filled boxes*) and ORF (*open boxes*) of the cDNA sequences for platelet GPIbβ (*top*) and neo Ibβ (*bottom*) are shown. Putative transcriptional start sites (+1), ATG/TGA start/stop codons, and lengths of different regions (base:b number, not to scale, illustrated by *broken lines*) are noted. The platelet GPIbβ transcript reflects two exons with a 274-b intron spliced out, capable of translating a 206–amino acid (aa) polypeptide (10b + 608b ORF). The neo Ibβ transcript starts from an upstream site and can translate 411 aa from a 1,233-b ORF (319-b 5′ extension, 32-b platelet exon, 274-b "platelet intron"-retained as exon, 608-b 3′ region), reading three different amino acids from the ORF of the first platelet exon (i.e., frame shift: platelet ATG start codon → .CCA TGG.) but the same 203 COOH-terminal aa as the ORF of the second platelet exon. The neo translation product also includes 106 aa encoded by the 5′ extension, 11 aa by the first platelet exon, and 91 aa by the intron.

The Journal of Clinical Investigation, Inc.
Volume 93, June 1994, 2301–2302

LRG motif is a versatile functional element that has been adapted and retained by a variety of genes (4, 5). This exact process of "adaption/retention" appears to have occurred in the case of neo Ibβ and platelet GPIbβ. Apparently, the platelet GPIbβ gene, as part of both the GPIb-V-IX and the LR/LRG systems, was an attractive template for such variation, and the neo Ibβ gene was the result. A fascinating aspect of the neo Ibβ observation will be to pursue the location and function of the neo protein in comparison with its adhesive cousin, platelet GPIbβ. Undoubtedly, such an analysis will reveal new aspects of LRG function, complementary to recent progress made with another LR/LRG protein, ribonuclease inhibitor, by both deletion mutation and crystal structure studies (5, 6).

*Megakaryocytic versus endothelial gene expression.* Certain genes are known to be specifically expressed in either megakaryocytes (i.e., GPIIb) or endothelia (E-selectin) or in both (GPIIIa, vWf) (7). However, the pairing of neo and platelet Ibβ on a single locus of genomic DNA presents a new aspect of such "tissue specificity" in which a distinct promoter for each cell type could be studied with a single construct. Perhaps, the unusual pairing of these Ibβ genes will provide a uniquely informative model for the study of gene regulation in endothelial cells as compared with megakaryocytes.

Valuable research opens up areas of inquiry, and neo Ibβ is telling us to explore new aspects of eukaryotic genetic variation, the LR/LRG proteins, and megakaryocytic/endothelial genes.

Gerald J. Roth
Hematology Section, Medical and Research Services
Seattle Veterans Administration Medical Center
and Department of Medicine
University of Washington

## References

1. Kelly, M. D., D. W. Essex, S. S. Shapiro, F. J. Meloni, T. Druck, K. Huebner, and B. A. Konkle. 1994. Complementary DNA cloning of the alternatively expressed endothelial cell glycoprotein Ibβ (GPIbβ) and localization of the GPIbβ gene to chromosome 22. *J. Clin. Invest.* 93:2417–2424.

2. Lopez, J. A., D. W. Chung, K. Fujikawa, F. S. Hagen, E. W. Davie, and G. J. Roth. 1988. The α and β chains of human platelet glycoprotein Ib are both transmembrane proteins containing a leucine-rich amino acid sequence. *Proc. Natl. Acad. Sci. USA.* 85:2135–2139.

3. Konkle, B. A., S. S. Shapiro, A. S. Asch, and R. L. Nachman. 1990. Cytokine enhanced expression of GPIbα in human endothelium. *J. Biol. Chem.* 265:19833–19838.

4. Roth, G. J. 1991. Developing relationships: arterial platelet adhesion, glycoprotein Ib and leucine-rich glycoproteins. *Blood.* 77:5–19.

5. Lee, F. S., and B. L. Vallee. 1993. Structure and action of mammalian ribonuclease (angiogenin) inhibitor. *Prog. Nucleic Acid Res. Mol. Biol.* 44:1–30.

6. Kobe, B., and J. Deisenhofer. 1993. Crystal structure of porcine ribonuclease inhibitor, a protein with leucine-rich repeats. *Nature (Lond.).* 366:751–756.

7. Weller, A., S. Isenmann, and D. Vestweber. 1992. Cloning of the mouse endothelial selectins. Expression of both E- and P-selectin is inducible by tumor necrosis factor. *J. Biol. Chem.* 267:15176–15183.