Quaternary organization of GPIb-IX complex and insights into Bernard-Soulier syndrome revealed by the structures of GPIbβ and a GPIbβ/GPIX chimera

Running title: Glycoprotein GPIbβ-GPIX structure

Paul A. McEwan†, Wenjun Yang‡, Katherine H. Carr†, Xi Mo‡, Xiaofeng Zheng‡, Renhao Li‡*, Jonas Emsley†*

†Centre for Biomolecular Sciences, School of Pharmacy, University of Nottingham, Nottingham, NG7 2RD, U.K.; ‡Center for Membrane Biology, Department of Biochemistry and Molecular Biology, The University of Texas Health Science Center at Houston, Houston, Texas 77030, USA.

†These authors contributed equally to this work.
*Corresponding Authors:
Jonas Emsley.
Centre for Biomolecular Sciences, School of Pharmacy, University of Nottingham, Nottingham, NG7 2RD, U.K.
Phone: 44-1158467092; Fax: 44-1158468002; Email: jonas.emsley@nottingham.ac.uk

and

Renhao Li.
current address; Division of Hem/Onc/BMT, Department of Pediatrics, Emory University School of Medicine, 2015 Uppergate Drive, Room 464, Atlanta, GA 30322.
Phone: 404-727-8217; Fax: 404-727-4859; Email: Li; renhao.li@emory.edu
Abstract

Platelet glycoprotein (GP)Ib-IX receptor complex has three subunits GPIbα, GPIbβ and GPIX which assemble with a ratio of 1:2:1. Dysfunction in surface expression of the complex leads to Bernard-Soulier syndrome (BSS). We have crystallised the GPIbβ ectodomain (GPIbβE) and determined the structure to reveal a single leucine rich repeat with N- and C-terminal disulphide bonded capping regions. The structure of a chimera of GPIbβE and three loops (a,b,c) taken from the GPIX ectodomain sequence was also determined. The chimera (GPIbβEabc), but not GPIbβE, forms a tetramer in the crystal, revealing a quaternary interface between GPIbβ and GPIX. Central to this interface is residue Y106 from GPIbβ which inserts into a pocket generated by two loops (b,c) from GPIX. Mutagenesis studies confirmed this interface as a valid representation of interactions between GPIbβ and GPIX in the full-length complex. Eight GPIbβ missense mutations identified from BSS patients were examined for changes to GPIb-IX complex surface expression. Two mutations, A108P and P74R, were found to maintain normal secretion/folding of GPIbβE but were unable to support GPIX surface expression. The close structural proximity of these mutations to Y106 and the GPIbβE interface with GPIX indicates they disrupt the quaternary organization of the GPIb-IX complex.
**Introduction**

Glycoprotein (GP)Ib-IX-V complex is an abundant membrane receptor complex on the platelet surface that plays a critical role in mediating platelet adhesion to the damaged vessel wall under conditions of high shear stress\(^1\). Platelets adhere, and integrins are subsequently activated by interactions of GPIb-IX-V with von Willebrand factor (VWF) bound to the subendothelium. How GPIb-IX-V transmits the VWF-binding signal across the membrane is not clear, partly because the structure and organization of this complex receptor remains to be elucidated. As GPV is only weakly associated with the receptor complex and is not essential for complex expression, assembly, VWF-binding or signal transduction\(^2,3\), we focus on the GPIb-IX complex here.

The GPIb-IX complex contains 3 subunits, GPIb\(_{\alpha}\), GPIb\(_{\beta}\) and GPIX, with a 1:2:1 stoichiometry\(^4\). Each subunit is a type I transmembrane protein, containing an ectodomain with leucine-rich repeats (LRR)\(^5\), a single transmembrane helix and a relatively short cytoplasmic tail. The GPIb\(_{\alpha}\) ectodomain contains binding sites for a growing list of hemostatically important ligands including VWF and thrombin\(^6^-8\). Covalent and non-covalent interactions are important to the quaternary stabilization of the receptor. GPIb\(_{\alpha}\) links to two GPIb\(_{\beta}\) subunits through membrane-proximal disulfide bonds to constitute the GPIb complex\(^4\). GPIX tightly associates with GPIb through noncovalent interactions\(^9\). Assembly of these subunits into a tightly integrated complex is also supported by genetic evidence. Bernard-Soulier syndrome (BSS) is a hereditary bleeding disorder that is characterized in most cases by giant platelets, low platelet counts and little or no expression of GPIb-IX on the platelet surface\(^10,11\). More than 30 mutations have been identified from BSS patients and mapped to GPIb\(_{\alpha}\), GPIb\(_{\beta}\) and GPIX\(^12,13\), indicating that all 3 subunits are required for proper surface expression of the complex.
Consistent with genetic evidence, efficient surface expression of GPIb-IX also requires all 3 subunits in transfected mammalian cells\textsuperscript{14}. For instance, GPIX alone cannot be expressed on the surface of transfected Chinese hamster ovary (CHO) cells. Only when co-expressed with GPIbβ, can it be detected on the cell surface, indicating that GPIbβ interacts with and stabilizes GPIX\textsuperscript{15,16}. Co-expression with GPIbβ and GPIbα produces even higher surface expression levels of GPIX\textsuperscript{14}. Thus, the surface expression level of individual subunits can be used as an indicator for the assembly of GPIb-IX and, implicitly, quaternary interactions among the subunits. Using this approach, we had previously shown that transmembrane domains of GPIb-IX are essential for complex assembly\textsuperscript{17,18}. Biophysical characterization of recombinant GPIb-IX derived transmembrane peptides in detergent micelles indicated that they form a parallel 4-helical bundle, and that their association leads to formation of membrane-proximal disulfide bonds between GPIbα and GPIbβ, further stabilizing the complex\textsuperscript{4,19}.

In addition to transmembrane domains, GPIbβ and GPIX ectodomains, termed GPIbβ\textsubscript{E} and GPIX\textsubscript{E} in this paper, respectively, are required for proper assembly and efficient surface expression of GPIb-IX, as numerous BSS-causing mutations have been mapped to these domains. However, GPIX\textsubscript{E} is intrinsically unstable, impeding structural and biochemical investigation. Taking advantage of the high sequence homology between GPIbβ\textsubscript{E} and GPIX\textsubscript{E} (Figure 1A), we had earlier identified a GPIbβ\textsubscript{E}/GPIX\textsubscript{E} chimera (abbreviated to GPIbβ\textsubscript{Eabc}) that interacts with GPIbβ in a manner that mimics the GPIX ectodomain\textsuperscript{16}. Here we report X-ray crystal structures of human GPIbβ\textsubscript{E} and GPIbβ\textsubscript{Eabc}. These structures, in combination with assays to probe inter-subunit interactions in the context of full-length subunits, give insight into GPIb-IX quaternary assembly and provide greater understanding of the molecular mechanism of BSS.
Methods

Expression and purification of recombinant GPIb$\beta_E$ and GPIb$\beta_{Eabc}$

To produce recombinant GPIb$\beta_E$ and GPIb$\beta_{Eabc}$, a gene fragment encoding human GPIb$\beta$ and GPIb$\beta_{Eabc}$ residues Cys1-Leu121, respectively, was cloned into a pBlueBac4.5-derived baculovirus expression vector (Invitrogen) that appended the signal sequence of baculovirus envelope glycoprotein gp67 (amplified from pAcGP67A vector, BD Biosciences) and a Ser-Ser-hexahistidine tag to the N-terminus and C-terminus of GPIb$\beta_E$, respectively. The chimeric GPIb$\beta_{Eabc}$ contains three stretches of GPIX sequences: Ala29-Arg36, Ser49-Gln60, and Ser76-Arg87 (Fig. 1A). The mature protein with the hexahistidine tag was secreted from the infected insect cells into the culture media. After 40-85% ammonium sulfate fractionation of the collected culture media, the protein was dissolved in 50 mM Tris·HCl, 20 mM imidazole, 4 mM EGTA, pH 7.6 at 4 °C overnight, loaded onto the Ni-sepharose column (Qiagen), and subsequently eluted with 50 mM Tris·HCl, 300 mM NaCl, 100 mM imidazole, 4 mM EGTA, pH 7.6. The eluent was concentrated and further purified by gel filtration chromatography (Superdex 75, GE Healthcare) in 20 mM Tris·HCl, 100 mM NaCl, pH 7.0 (Supplemental Figure 1). N-terminal sequencing analysis of the purified protein confirmed the proper removal of the signal sequence.

Crystallization and structural determination

Purified GPIb$\beta_E$ with the hexahistidine tag was dialyzed into 20 mM Tris·HCl, 100 mM NaCl, pH 7.0 and concentrated to 18 mg/mL for crystallization at 19°C. Sparse matrix screens (Qiagen, UK) obtained initial conditions and refinement resulted in the crystallization condition of 1.6 M (NH$_4$)$_2$SO$_4$, 0.4 M LiCl, 0.1 M MES, pH 6.5. For GPIb$\beta_{Eabc}$ the sample was
concentrated to 1.5 mg/ml and sparse matrix screens (PACT, Molecular Dimensions) identified initial conditions at 10°C from D2 (form 1) and A9 (form 2). The D2 condition is 0.1 M MMT buffer pH 5 (MMT buffer is a mixture of DL-malic acid, MES and Tris base in the molar ratios 1:2:2 – DL-malic acid:MES:Tris base.), 25% (w/v) PEG 1500 and A9 is 0.2M lithium chloride, 0.1M sodium acetate pH 5 and 20% PEG 6000. Single crystals were transferred to the same solution containing 25% and 10% glycerol respectively and flash cooled in liquid nitrogen. Diffraction data were collected using beam line ID23-2, ID29-1 and ID23-2 respectively for GPIbβE, GPIbβEabc (form 1) and GPIbβEabc (form 2) at the European Synchrotron Radiation Facility (ESRF). The GPIbβE structure was determined by molecular replacement using the structure for C-terminal 133 residues of the Nogo-66 ectodomain (PDB code: 1P8T)20 and programs MrBUMP21 and PHASER22. Initial electron density maps were improved using 2-fold non-crystallographic symmetry and solvent flattening using the CCP4 program suite. Model rebuilding was performed using COOT23 and crystallographic refinement carried out in REFMAC24. Crystallographic statistics are listed in Table 1. A Ramachandran plot shows 115 residues in preferred regions, 2 in allowed regions and none in outlier regions. The GPIbβEabc form 1 structure was determined by molecular replacement using the GPIbβE structure. The model was built using COOT and refined using REFMAC. A Ramachandran plot shows 116 of residues in favoured regions and 3 in allowed regions with none in outlier regions. The form 2 structure was determined by molecular replacement using the form 1 structure and is identical with the exception of side chains involved in crystal packing.
**GPIbβ and GPIX constructs**

Vectors expressing HA-GPIbβ (full-length GPIbβ with N-terminal HA epitope tag, YPYDVPDYA), HA-GPIbβE (GPIbβ extracellular residues Cys1-Leu121 with N-terminal HA tag), HA-GPIbβEabc-GPIXTC (HA-tagged GPIbβEabc fused to GPIX transmembrane and cytoplasmic domains), HA-GPIbβEabc (HA-tagged GPIbβEabc), GPIX and GPIbα had been described\(^{16,17,25}\). Site directed mutagenesis was carried out by the overlap extension PCR procedure using the above vectors as the template. Each PCR fragment was digested by appropriate restriction enzymes and subcloned into its target vector as described earlier\(^ {17} \). All constructs were confirmed by DNA sequencing.

**Transient transfection of CHO cells**

CHO K1 cells were grown in Dulbecco’s Modified Eagle’s medium supplemented with 10% fetal calf serum at 37 °C and 5% CO\(_2\). Transient transfection of vectors encoding GPIbα, GPIbβ and GPIX-derived constructs, in desired combinations, into CHO K1 cells was performed using lipofectamine 2000 (Invitrogen, Carlsbad, CA) as described earlier\(^ {17} \). Key parameters of the transfection, such as cell density and DNA amount, were kept the same as previously described in order to allow proper comparison among various constructs\(^ {17} \). After transfection, the cells were grown for an additional 48 hours before being analyzed.

**Characterization of effects of GPIbβ or GPIX mutations**

The secretion and folding of mutant HA-GPIbβE and HA-GPIbβEabc proteins were characterized as previously described\(^ {16} \). Briefly, the N-terminally HA-tagged protein secreted into the culture medium was collected by co-immunoprecipitation using the anti-HA antibody,
resolved in 12% Bis-Tris SDS gel in the presence or absence of reducing agents, and immunoblotted by HRP-conjugated anti-HA antibody. The cellular expression and assembly of the GPIb-IX complex was assessed by western blot of each subunit as previously described\textsuperscript{17,18}. The surface expression levels of GPIb\(\alpha\), HA-tagged GPIb\(\beta\) and GPIX in transiently transfected cells was measured using WM23, anti-HA and FMC25 antibodies, respectively, on a Beckman-Coulter Gallios flow cytometer as described\textsuperscript{17,25}. To assess the mutational effect, the measured mean fluorescence value of the entire cell population (10,000 cells) is normalized with the value of CHO cells expressing wild-type GPIb-IX complex (GPIb\(\alpha\)/HA-GPIb\(\beta\)/GPIX) being 100\% and that of empty vector-transfected cells 0\%\textsuperscript{17}. Groups were compared using the two-tailed Student’s \(t\)-test.

## Results

### Crystal structure of GPIb\(\beta\textsubscript{E})

To explore the architecture of the GPIb-IX receptor complex we crystallized recombinant GPIb\(\beta\textsubscript{E}) and solved the crystal structure to 1.25 Å resolution (Table 1). The topology of GPIb\(\beta\textsubscript{E}) spanning residues 1 to 118 is shown in Figure 1B, revealing the first structure with only a single LRR repeat. A typical example of the refined electron density is illustrated in stereo in Figure 1C and Supplemental Video 1. As seen in other LRR proteins, GPIb\(\beta\textsubscript{E}) assumes a right-handed coiled structure with a parallel \(\beta\)-sheet on one side (the concave face) and connecting loops containing \(\beta\)-turns and short 3\textsubscript{10} helices on the opposite (convex) face. The central LRR is covered on both ends by N- and C-terminal capping regions that contain several short \(\alpha\)- and 3\textsubscript{10}-helices. Four disulfide bonds are observed in the GPIb\(\beta\textsubscript{E}) structure: two (Cys1-Cys7, Cys5-Cys14) are located in the N-terminal cap and the other two (Cys68-Cys93, Cys70-Cys116) in the
C-terminal cap, which are topologically equivalent to those in the Nogo-66 and SLIT receptor\textsuperscript{20,26}. With only a single LRR it assumes a compact rectangular shape with a relatively flat, rather than a curved concave face commonly observed in multi-LRR structures. Moreover, the single LRR accommodates a unique feature in GPIb\textsubscript{β} that has not been observed in previously reported LRR structures; interactions of side chains bridging N- and C-terminal capping regions. Extending over the convex face, the aromatic ring in Trp21 is locked between the amino group of Pro46 and the guanidinium group of Arg71 by cation-π interactions (Figure 1B). These interactions exemplify numerous inter-loop interactions on the convex face, which likely add stability to the structure in a manner similar to the buried GPIb\textsubscript{β} Asn residues, Asn40 and Asn64, on the concave face\textsuperscript{27}.

**Different pathways in the pathogenesis of GPIb\textsubscript{β} BSS mutations**

Eight missense mutations in GPIb\textsubscript{β}E — C5Y\textsuperscript{28}, R17C\textsuperscript{29}, P29L\textsuperscript{30}, N64T\textsuperscript{31}, P74R\textsuperscript{32}, Y88C\textsuperscript{33,34}, P96S\textsuperscript{35} and A108P\textsuperscript{34} — have been identified in BSS patients. We have examined the context of these eight mutations using the GPIb\textsubscript{β}E structure. Residues affected are illustrated in a ribbon diagram of the structure in Figure 2A and also in Supplemental Video 2. Information on surface localization and conservation of each affected residue is summarized in Supplemental Table 1. Cys5 and Cys14 form a disulfide bond. Substitution of Cys5 with a tyrosine would result in loss of the disulfide bond and leave an unpaired Cys14. Asn64 is fully buried in the structure and bridges the C-terminal cap and the LRR with two hydrogen bonds, both of which would be lost if substituted by a smaller Thr residue. The other GPb\textsubscript{β}E residues — Arg17, Pro29, Pro74, Tyr88, Pro96, Ala108 — are present on the protein surface, although Arg17, Tyr88 and Pro96 are partially buried by surrounding side chains (Figure 2B). Both R17C and Y88C would introduce
an additional Cys residue to the domain, which is likely to interfere with formation of the four native disulfide bonds. The other 4 mutations involve either removal or addition of a Pro residue, which can impact on local conformation or global stability.

To elucidate the molecular pathogenesis of these BSS mutations, we have characterized systematically the mutational effects on GPIb-IX expression and assembly in transiently transfected CHO cells. Key parameters of transient transfection were kept constant to ensure proper comparison of protein expression levels among different experiments. CHO cells largely recapitulated the reported clinical observations that all 8 mutations led to a significant decrease in surface expression of the GPIb-IX complex, albeit to various degrees (Supplemental Figure 2).

To test whether these BSS mutations are detrimental to the structural integrity of GPIbβE, each mutation was introduced to the HA-tagged GPIbβE (with the native signal sequence) and the resulting gene transfected transiently into CHO cells. Western blot of the cell lysate indicated that translation of the HA-GPIbβE gene was not affected by any of the mutations (Figure 2C). However, C5Y, P29L and P96S mutant proteins failed to secrete from the cells since no HA-tagged protein was detected in the culture media. The secretion of R17C was significantly decreased. In N64T-expressing cells, only the ectodomain with a higher molecular mass was detected in the culture media. The secretion of R17C was significantly decreased. In N64T-expressing cells, only the ectodomain with a higher molecular mass was detected in the culture media. The other mutants, P74R, Y88C and A108P, were secreted similar to the wild-type. Of the mutants that were secreted, R17C, N64T and Y88C exhibited significant formation of inter-molecular disulfide bonds as detected in SDS-PAGE under non-reducing conditions. Since GPIbβE contains 4 intra-molecular disulfide bonds and no inter-molecular ones (Figure 1B), the existence of the latter indicated misfolding of these mutant proteins. In contrast, P74R and A108P GPIbβE mutant proteins, like the wild-type, contained only intra-molecular disulfide bonds and should therefore be well-folded.
The effects of mutations on the interaction between GPIbβ and GPIX were analyzed next. GPIX does not express on the cell surface in isolation but becomes detectable when coexpressed with GPIbβ\textsuperscript{14}. GPIbβ is thought to interact with and stabilize GPIX, which involves the ectodomain and transmembrane domain of both proteins\textsuperscript{15,16}. As shown in Figure 2D, when HA-tagged GPIbβ and GPIX were co-transfected into CHO cells, both HA-GPIbβ and GPIX were detected on the cell surface. Mutations that disrupt secretion or folding of GPIbβ\textsubscript{E}, such as P29L, resulted in little cell surface expression of HA-GPIbβ and, as a consequence, GPIX as detected by flow cytometry. On the other hand, a distinctive feature of P74R and A108P was that HA-GPIbβ was present on the cell surface but GPIX was not. Overall, these results revealed that BSS missense mutations disrupt assembly and surface expression of the GPIb-IX complex via different mechanisms.

**Crystal structure of GPIbβ\textsubscript{Eabc} and conformational change in the Y88 loop**

Despite high sequence similarity between GPIbβ\textsubscript{E} and GPIX\textsubscript{E}, attempts to express recombinant GPIX\textsubscript{E} have not been successful\textsuperscript{16}. But sequence analysis did allow the engineering of GPIbβ\textsubscript{Eabc}, a chimera of GPIbβ\textsubscript{E} and GPIX\textsubscript{E}, as a stable protein that is readily secreted\textsuperscript{16}. Built on the GPIbβ\textsubscript{E} scaffold, GPIbβ\textsubscript{Eabc} incorporates three discontinuous stretches of GPIX\textsubscript{E} that correspond to the 3 convex loops including the α1 helix (termed loops a, b, c; spanning GPIX residues Ala29-Arg36, Ser49-Gln60, Ser76-Arg87 as shown in Figure 1A). GPIbβ\textsubscript{Eabc} was purified from insect cells, crystallized and the structure solved to 2.35Å resolution, using the GPIbβ\textsubscript{E} crystal structure for a molecular replacement calculation identifying four molecules in the asymmetric unit (Figure 3A, Table 1). To avoid confusion, residues of GPIbβ\textsubscript{Eabc} are labeled
in this paper by their residue numbers in respective source domains (e.g. GPIX-Asp56 in loop b) rather than by GPIb\textsubscript{Eabc}’s own.

The GPIb\textsubscript{Eabc} structure shares the same topology as GPIb\textsubscript{E}, but has three changes in the structure (Figure 3B). The first, as expected, is in the convex loops. Due to the substantial sequence difference, the 3\textsubscript{10} helix in loop a and the cation-π inter-loop interaction between Trp21 and Arg71 in GPIb\textsubscript{E} are not present in the GPIb\textsubscript{Eabc} structure (Figure 3A). Second, in helix α1 of loop c GPIX-Tyr79 replaces GPIb\textsubscript{E}-Pro74, which surprisingly does not affect the kinked main chain conformation but instead forms a substantial new interaction with loop b (Figure 3A). Here the GPIX-Tyr79 side chain is buried under the main chain of GPIX-Gly53 and the Tyr OH group forms a hydrogen bond to the main chain nitrogen of adjacent GPIX-Phe55 (Figure 3A). The extensive interactions observed in these convex loops substantiate our earlier observation that only when all three convex loops are grafted together onto the GPIb\textsubscript{E} scaffold did the GPIb\textsubscript{Eabc} chimeric protein become stable and well-folded\textsuperscript{16}. A third difference occurs in the C-terminal cap, despite GPIb\textsubscript{E} and GPIb\textsubscript{Eabc} sharing the same amino acid sequence in this region. Residues 86-89 that form the 3\textsubscript{10} helix in loop c of GPIb\textsubscript{E} unravel and residues 80-86 coil up to form a novel turn of helix in GPIb\textsubscript{Eabc} with GPIb\textsubscript{E}-Glu84 forming a new salt bridge to GPIb\textsubscript{E}-Arg57. In addition, the C-terminal helix α2 makes a rigid body movement of 3 Å away from helix α1. Compared to GPIb\textsubscript{E}, the angle between helices α1 and α2 in GPIb\textsubscript{Eabc} is widened (Figure 3C).

The re-coiling conformational change in the loop containing residue Y88 is in a topologically equivalent region of the LRR fold to the β-switch loop in GPIbα known to alter conformation when binding to vWF-A1 domain or peptide inhibitor\textsuperscript{6,36}. In GPIbα the conformational change involves unraveling of a short α-helix to form an extended β-hairpin whereas in GPIbβ the α-
helix unravels and a second stretch is formed, displaced towards the C-terminus in a twisting helical motion (Figure 3D).

**Tetrameric structure of GPIbβ_{Eabc} reveals a GPIbβ-GPIX interface**

Previous studies with the GPIbβ_{Eabc} construct demonstrated that the three loops (a,b,c) from GPIX form ectodomain interactions with GPIbβ and this is required for surface expression of the GpIb-IX complex\(^{16}\). The side chains from GPIbβ involved in this interaction with GPIX were unknown. In the crystal structure we observe four GPIbβ_{Eabc} molecules in the asymmetric unit form a tetrameric ring structure. Here the C-terminal cap region of one subunit packs against the convex loops (b,c) of another placing the α-helices towards the center of the ring where they are partially buried (Figure 4A). At this interface the side chain of GPIbβ-Tyr106 from one subunit lies at the center of a shallow pocket created between loop b and helix α1 (loop c) from a second subunit that is rotated by 90 degrees. Each interface buries a surface area of approximately 1,100Å\(^2\), which is significantly greater than values typical for crystal contacts\(^{37}\) or the contacts found in the GPIbβ_{E} structure. A second monoclinic crystal form (space group P2\(_1\)) of GPIbβ_{Eabc} grown under different conditions has an identical structure and tetrameric arrangement (labeled form 2 in Table 1).

Figure 4B illustrates hydrophobic contacts contribute to the GPIX pocket including GPIX-Tyr79, GPIX-Leu82, GPIX-Trp83 and more peripheral contacts to the interface come from GPIbβ-Ala108 and GPIX-His57. The OH group of GPIbβ-Tyr106 forms a hydrogen bond to the guanidinium of GPIX-Arg87. GPIbβ-Tyr106 is further surrounded by three salt bridges between GPIbβ residues Arg102, Asp90, Arg89 and GPIX residues Asp56, Arg87, Asp86 respectively.
GPIbβ-Arg92 and GPIX-Arg87 side chains form hydrogen bonds to the main chain carbonyls GPIX-Asp56 and GPIbβ-Ala86 respectively.

The interface between GPIbβEabc molecules is comprised of the GPIbβE-derived C-terminal region and GPIXE-derived convex loops. To test whether this interface is a valid representation of full-length GPIbβ and GPIX in the whole complex, three residues were selected for mutagenesis. GPIbβ-Tyr106 lies at the center of the interface and GPIX-Leu82, GPIX-Asp86 are examples of residues which contribute hydrophobic and electrostatic contacts to the interface respectively (illustrated as underlined in Figure 4B). In the GPIbβE crystal structure, the side chain of Tyr106 is exposed to the solvent and flanked by Arg85 and Arg89. Mutating Tyr106 to Phe, Asn, Asp, Ala or Val largely preserved proper secretion and folding of HA-GPIbβE that were expressed transiently from CHO cells (Figure 5A). However, when expressed as a full-length subunit on the surface of cells, all Tyr106 mutations resulted in the loss of the ability of HA-tagged GPIbβ (HA-GPIbβ) to enhance surface expression of GPIX in transfected CHO cells (Figure 5B). Thus, as predicted by the GPIbβEabc crystal structure, these results pinpoint GPIbβ-Tyr106 as a critical part of a GPIXE binding site.

Similarly HA-GPIbβEabc protein containing a GPIX-L82A or GPIX-D86E mutation was able to secrete to the culture media and fold well (Figure 6A). However, both mutations failed to retain the GPIbβ binding ability of GPIbβEabc in the context of the full-length subunit (Figure 6B). Since GPIXE alone cannot express as a well-folded form, it would be difficult to assess the effect of either mutation on its folding and secretion. Nonetheless, full-length GPIX bearing either mutation could not be expressed on the cell surface even in the presence of GPIbβ (Figure 6C), which is consistent with the GPIbβEabc crystal structure that both GPIX-Leu82 and GPIX-Asp86 participate directly in the interaction between GPIbβE and GPIXE.
Discussion

One of the fundamental but unanswered questions about the GPIb-IX complex is how the subunits organize and assemble. Earlier studies have shown that transmembrane (TM) helices of the GPIb-IX complex interact with one another to form a parallel tetrameric α-helical bundle\textsuperscript{4,18,19}. Here the GPIX TM helix bridges the two GPIb\(\beta\) TM helices and GPIb\(\alpha\) performs a similar function providing non-covalent and covalent quaternary interactions and stability by formation of the two interchain disulfide bonds bridging two GPIb\(\beta\). Previously it has also been shown that the LRR containing ectodomains form an interface involving the convex loops of GPIX interacting with GPIb\(\beta\)\textsuperscript{16}. Overall protein-protein interactions mediated by LRR domains utilize a diverse range of scaffolds with different numbers of LRRs (ranging from 1 to >15) mediating homo and heterotypic interactions\textsuperscript{5}. These interactions are principally mediated via the LRR concave face which typically has a curved or horseshoe shape as observed in GPIb\(\alpha\)\textsuperscript{6}. GPIb\(\beta\) and GPIX are unusual in that they have only a single LRR. The GPIb\(\beta_E\) structure is the first description of this fold and reveals a compact shape with little LRR curvature.

A crystal structure for the SLIT receptor LRR ectodomain domain 4 (pdb code 2WFH) reveals a homodimer with a concave face to face interaction of the LRRs. We initially speculated that the two copies of GPIb\(\beta_E\) may also interact in a similar way to form a dimer, however, in the crystal structure we only observe a monomer. By contrast the structure of a chimeric GPIb\(\beta_{Eabc}\), where three loops from GPIX are added, reveals a heterodimeric GPIb\(\beta\)-GPIX interface. Due to the positioning of the two subunits in the dimer at right angles this structure is then able to cyclize with a second dimer to form a ring like tetramer where the same interface is observed four times. Rather than utilizing the LRR concave face this structure reveals interactions occur through the C-terminal cap of GPIb\(\beta_E\) and convex face loops of GPIX\(E\). Here a central side chain
Y106 from the GPIb\(\beta_E\) cap inserts into a pocket created by two loops from GPIX on the convex face. This is more familiar to the interfaces observed in Toll-like receptors where heterodimeric interactions form between the C-terminal cap and one side of the LRRs\(^{38}\).

An important spatial constraint for the GPIb-IX complex lies in the uniform topology of its subunits. Since GPIb\(\alpha\), GPIb\(\beta\) and GPIX are all type I transmembrane proteins, the C-termini of all 4 ectodomains in the GPIb-IX complex should remain in close proximity to one another to enable formation of the adjacent parallel TM helix bundle. We noticed that in the GPIb\(\beta_{Eabc}\) crystal structure, the C-termini of all 4 GPIb\(\beta_{Eabc}\) molecules are located on the same side of the tetramer (Figure 4A, Supplemental Video 3). If the GPIb\(\beta_1\)-GPIX-GPIb\(\beta_2\) trimer has a similar cyclic arrangement positioned above the TM helices then a schematic model of this trimer is shown in Figure 7A with the GPIb\(\alpha\) TM helix added (shown in green) extending towards the N-terminal ligand binding domain. A second observation from the GPIb\(\beta_{Eabc}\) structure is that GPIX can only bind one GPIb\(\beta_E\) through its \(b,c\) loop pocket and thus further studies will be required to define what the context and conformation of a second GPIb\(\beta_E\) (GPIb\(\beta_{E2}\)) is. As we observe no interaction between two GPIb\(\beta_E\) molecules in the crystal structure the model in Figure 7A shows GPIb\(\beta_{E2}\) on the opposite side of GPIX to GPIb\(\beta_{E1}\).

**Different pathways to BSS**

The genetic basis for BSS was established 3 decades ago\(^{11}\). Over 30 mutations have since been identified from BSS patients and the number is likely to grow\(^{39}\). As we showed in this study several novel mutations could produce BSS-like phenotypes in transfected cells (Figure 5,6). Some BSS mutations are located in the promoter region and are presumed to decrease gene transcription\(^{40}\). Some are frame-shifting or nonsense mutations, resulting in a typically truncated
and nonfunctional protein product\textsuperscript{41-43}. In this study, we have for the first time systematically characterized all the reported missense BSS mutations in GPIbβ and observed distinct mechanisms that underlie the pathogenesis of the disease (Figure 2). Six out of the eight missense mutations are detrimental to proper tertiary folding or secretion of GPIbβ\textsubscript{E}. On the other hand, the other two mutations A108P and P74R maintain the native disulphide bonds and are likely to fold normally.

The clinical data on the A108P mutation describes a patient who is compound heterozygous with mutation Y88C\textsuperscript{34}. In these platelets GPIb-IX receptor complex is present on the surface at a reduced level and does bind VWF. Antibody SZ1 (binds GPIbβ/GPIX subunits in complex but not in isolation) recognizes the GPIb-IX present. As a compound heterozygote of GPIbβ the mutations could form different combinations of the chains i.e. A108P/A108P, A108P/Y88C, Y88C/Y88C. A separate family which are homozygous for Y88C have classical BSS giant platelets with no receptor present at the platelet surface\textsuperscript{33}. As illustrated in Figure 4B Ala108 is adjacent to Y106 which is central to the GPIbβ\textsubscript{Eabc} interface. We showed mutating Y106 abolishes GPIX expression in the same way as A108P. Mutating Ala108 to Pro at the periphery of the interface, may produce a less stable subunit association which, in turn, gives rise to the reduction in surface expression of the whole complex which is what we observe in CHO cells. The P74R mutation shows classic BBS platelets in homozygous patients with no receptor complex detected at the surface\textsuperscript{32}. This mutation is more severe than A108P in our cell based assays and does not support any GPIX expression at the cell surface even in the presence of GPIbα (Supplemental Figure 2). Pro74 does not contribute directly to the Y106 interface but is located close by in helix α1 which does contribute through side chains from residues Ala86, Arg89, Asp90 (Figure 4B). Proline residues are commonly found at the N-termini of helices
where the special main chain properties influence the secondary structure formation. A change here could disrupt the orientation of the helix and thus its contributions to the interface.

Finally, not all mutations characterized in GPIbβ result in loss or reduction of receptor complex at the platelet surface. The polymorphism G15E is associated with the human platelet-specific alloantigen (HPA) Iy<sup>a</sup> and does not adversely affect the expression level of GPIb-IX<sup>44</sup>. Gly15 is exposed on the protein surface and located in the N-terminal cap. Figure 7B shows the locality of these mutations in the same model as Figure 7A and illustrates how two copies of GPIbβ in this arrangement presents mutations in different contexts (Supplemental Video 4). Overall these studies provide valuable insights on the assembly of GPIb-IX complex and will prove a scaffold for further investigation of this important platelet receptor.

**Acknowledgments**

We acknowledge the European Synchrotron Radiation Facility for provision of synchrotron radiation source, and thank David Flot and Gordon Leonard for assistance in using the beamline ID23-2. This work was supported in part by NIH grant HL082808 (R.L.) and Cancer Research UK grant C21418/A8573, British Heart Foundation grant RG/07/002/23132 (J.E.). Protein Data Bank: Coordinates and structure factors deposited are 3RFE for GPIb<sub>E</sub>β and 3REZ for GPIb<sub>Eabc</sub>β.

**Authorship contributions**

PAM and KHC crystallized and determined the structures of GPIb<sub>E</sub>β and GPIb<sub>Eabc</sub>β respectively. WY produced recombinant proteins for crystallization, characterized and analyzed mutational effects on GPIb-IX expression and assembly, and wrote the paper. XM and XZ cloned many
constructs and established the baculovirus protein expression system. RL initiated and designed research, analyzed results, and wrote the paper. JE analyzed results, and wrote the paper.

Disclosure of conflict of interest

The authors declare that they have no conflict of interest.
References


10.3324/haematol.2010.032631.


Table 1. Crystallographic data collection and refinement statistics.

<table>
<thead>
<tr>
<th>Sample</th>
<th>GPIβε</th>
<th>GPIβε_{Eabc} (form 1)</th>
<th>GPIβε_{Eabc} (form 2)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data collection</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Space group</td>
<td>P2(1)</td>
<td>P3(1)21</td>
<td>C2(1)</td>
</tr>
<tr>
<td><strong>Cell dimensions</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a, b, c (Å)</td>
<td>61.60, 34.83, 71.77</td>
<td>72.01, 72.01, 171.73</td>
<td>124.27, 72.72, 72.57</td>
</tr>
<tr>
<td>α, β, γ (°)</td>
<td>90, 90.31, 90</td>
<td>90, 90, 120</td>
<td>90, 107.06, 90</td>
</tr>
<tr>
<td>Resolution (Å)*</td>
<td>31.54-1.25 (1.28-1.25)</td>
<td>42.17-2.35 (2.48-2.35)</td>
<td>36.36-3.2 (3.82-3.20)</td>
</tr>
<tr>
<td>Rsym * a</td>
<td>0.145 (0.400)</td>
<td>0.131 (0.517)</td>
<td>0.136 (0.630)</td>
</tr>
<tr>
<td>I / sigI*</td>
<td>9.8 (3.8)</td>
<td>14.4 (5.0)</td>
<td>7.6 (1.9)</td>
</tr>
<tr>
<td>Completeness (%)*</td>
<td>93.4 (92.2)</td>
<td>100.0 (100.0)</td>
<td>95.2 (95.0)</td>
</tr>
<tr>
<td>Redundancy*</td>
<td>4.6 (4.0)</td>
<td>10.9 (11.1)</td>
<td>2.4 (2.4)</td>
</tr>
<tr>
<td><strong>Refinement</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of reflections</td>
<td>78922</td>
<td>19565</td>
<td>9796</td>
</tr>
<tr>
<td>Rwork / Rfree b</td>
<td>0.203/ 0.223</td>
<td>0.211/ 0.259</td>
<td>0.246/ 0.284</td>
</tr>
<tr>
<td>B-factors (Å²)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein</td>
<td>12.5</td>
<td>8.8</td>
<td>25.2</td>
</tr>
<tr>
<td>Ligands</td>
<td>21.0</td>
<td>39.3</td>
<td>61.3</td>
</tr>
<tr>
<td>Solvent</td>
<td>28.1</td>
<td>13.1</td>
<td>18.7</td>
</tr>
<tr>
<td>R.m.s. deviations</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bond lengths (Å)</td>
<td>0.0188</td>
<td>0.189</td>
<td>0.008</td>
</tr>
<tr>
<td>Bond angles (°)</td>
<td>1.889</td>
<td>1.301</td>
<td>1.435</td>
</tr>
</tbody>
</table>

*Values in parentheses are for highest-resolution shell.

a \( R_{sym} = \frac{\text{Sum}(h) [\text{Sum}(j) [I(hj) - <Ih>])}{\text{Sum}(hj) <Ih>} \) where I is the observed intensity and <Ih> is the average intensity of multiple observations from symmetry-related reflections calculated with SCALA.

b \( R_{work} = \frac{\|Fo\| - |Fc|}{\text{Sum}(h)[Fo|h}, where Fo and Fc are the observed and calculated structure factors, respectively. \( R_{free} \) computed as in \( R_{work} \), but only for (5%) randomly selected reflections, which were omitted in refinement, calculated using REFMAC.
Figure Legends

Figure 1. The crystal structure of GPIbβ ectodomain (GPIbβE). (A) Sequence alignment of GPIbβE (blue), GPIXE (purple), and GPIbβEabc (blue and purple) with respective residue numbers for GPIbβE and GPIX marked on the top and bottom. Elements of secondary structure in GPIbβE are illustrated on top and colored as in Figure 1B. Residues affected by BSS missense mutations are in red. Three stretches of the GPIXE sequence that are included in GPIbβEabc as are shown in purple. (B) Two orientations are shown of a ribbon diagram of the GPIbβE structure viewed from the concave face (left), with β-strands labeled in blue, α-helices red, 310 helices purple and loop regions grey. Asn residues 40, 41 and 64 are shown as stick and a single residue from the N-linked oligosaccharide attached to Asn41 is also shown in green. The diagram on the right is rotated 180 degrees with side chains in the inter-LRR cap cation-π interaction shown as stick. (C) Refined 1.25 Å electron density, calculated using 2Fo-Fc co-efficients and contoured at 1.5 r.m.s.

Figure 2. BSS mutations in GPIbβE disrupt expression and assembly of GPIb-IX complex by different mechanisms. (A) Two views of the GPIbβE structure are shown as a ribbon diagram related by a 90 degree rotation. Residues affected by BSS missense mutations are highlighted as stick and colored green for being solvent accessible and red for buried. (B) Space-filling representation of the GPIbβE structure, showing the concave (left) and convex (right) faces. Main chain atoms are colored white, and residues affected by BSS mutations colored green. (C) SDS gels showing differential effects of BSS mutations on expression (top panel), secretion (middle) and folding (bottom) of GPIbβE expressed from transfected CHO cells. Each
mutation is identified by the residue number. The HA epitope tag was appended to the N-terminal end of GPIbβ for easy detection. Immunoprecipitation was carried out using anti-HA monoclonal antibody, and immunoblotting using HRP-conjugated anti-HA monoclonal antibody (HRP-HA). Molecular weight markers are marked on the left of each gel. (D) Sample flow cytometry histograms showing the effects of selected BSS mutations on surface expression levels of HA-tagged full-length GPIb (HA-GPIb) and GPIX (GPIX) that are co-expressed transiently in CHO cells. (E) Relative surface expression levels of HA-GPIb (grey column) and GPIX (white) in transfected CHO cells. The surface expression levels were measured by flow cytometry and quantified as mean fluorescence intensity, which were normalized with expression levels in cells transfected with wild-type GPIb-IX (GPIbα/HA-GPIbβ/GPIX) being 100% and those in cells transfected with sham vectors 0%\textsuperscript{17}. The data are presented as mean±s.d. (n=3). *, \( p < 0.001 \).

Figure 3. The crystal structure of GPIb\textsubscript{Eabc} and its comparison with GPIb\textsubscript{E}. (A) A ribbon diagram of the GPIb\textsubscript{Eabc} structure, showing the grafted GPIX convex loops (magenta) in front. Residues derived from GPIb\textsubscript{E} are colored in blue. Side chains of several residues are shown in stick and labeled. GPIX-A29 and GPIX-S76 correspond to GPIbβ-W21 and GPIbβ-R71, respectively. (B) Superposition of GPIb\textsubscript{E} (green) and GPIb\textsubscript{Eabc} (blue/magenta) structures. (C) A close-up view of the conformational difference in the C-terminal cap region between GPIb\textsubscript{E} (green) and GPIb\textsubscript{Eabc} (blue) structures. The locations of several side chains in both structures are marked. (D) Topology diagrams of ligand free (green) and ligand bound structures in orange for GPIbα and blue for GPIbβ. The observed conformational change occurs in a topologically
equivalent loop (boxed) and in each case involves unwinding of a helix for GPIbα (R-loop) and GPIbβ (loop containing residue Y88) upon engaging ligand; vWF-A1 or GPIX respectively.

**Figure 4. The tetramer structure of GPIbβ_{Eabc} reveals a binding interface between GPIbβ_{E}-derived C-terminal cap region and GPIX_{E}-derived convex loops.** (A) Cartoon diagram showing two views of the GPIbβ_{Eabc} tetramer structure related by a 90 degree rotation. The N-termini (N) of the molecules are at the peripheral of the tetramer, and the C-termini (C) are all located in proximity at the bottom of the lower panel. Sequences derived from GPIbβ_{E} are colored in blue, and those derived from GPIX_{E} in magenta. The side chain of GPIbβ-Y106 at each interface is shown in green as stick. (B) A close-up view of the interface between two GPIbβ_{Eabc} molecules (colored blue, GPIbβ_{E}-derived and magenta, GPIX_{E}-derived residues, respectively). Interacting side chains are labeled and colored accordingly and underlined residues were subject to mutagenesis.

**Figure 5. Tyr106 mutations do not disrupt secretion and folding of GPIbβ_{E}, but disrupt its interaction with GPIX_{E}** (A) SDS gels showing the lack of disruptive effects of Tyr106 mutations on expression, secretion and folding of GPIbβ_{E} expressed from transfected CHO cells. The identity of each Tyr106 mutation is marked on top of the gels. The other annotations follow those of Figure 2C. (B) Relative surface expression levels of HA-GPIbβ (grey column) and GPIX (white) in transfected CHO cells measured by flow cytometry. The annotations follow those of Figure 2E. Note that none of the Tyr106 mutations retain the ability of wild-type HA-GPIbβ to enhance surface expression of GPIX. The data are presented as mean±s.d. (n=3). *, p < 0.001.
Figure 6. GPIX<sub>E</sub> convex loops participate in direct interaction with GPIbβ<sub>E</sub> in the full-length complex. (A) SDS gels showing the lack of disruptive effects of the L82A or D86E mutation on expression, secretion and folding of HA-GPIbβ<sub>Eabc</sub>. Residue numbers of both mutations are in the context of GPIX. The other annotations follow those of Figure 2C. (B) Relative surface expression levels of HA-GPIbβ<sub>Eabc</sub>-GPIX<sub>TC</sub>, either wild-type or containing the indicated mutation, in the absence (grey column) or presence (white) of coexpressing GPIbβ in transfected CHO cells. HA-GPIbβ<sub>Eabc</sub>-GPIX<sub>TC</sub> is a protein that contains N-terminally HA-tagged GPIbβ<sub>Eabc</sub> and GPIX transmembrane and cytoplasmic domains<sup>16</sup>. The expression level was measured by flow cytometry using anti-HA monoclonal antibody and quantified as mean fluorescence intensity, which was normalized with that in cells transfected with wild-type GPIb-IX (GPIbα/GPIbβ/HA-GPIX) being 100% and those in cells transfected with sham vectors 0%<sup>16</sup>. Note that the enhancement of HA-GPIbβ<sub>Eabc</sub>-GPIX<sub>TC</sub> surface expression by GPIbβ is significantly diminished by both mutations. The data are presented as mean±s.d. (n=4). *, p < 0.01. (C) Relative surface expression levels of GPIX, either wild-type or containing the indicated mutation, in the absence (grey column) or presence (white) of coexpressing GPIbβ in transfected CHO cells. GPIX surface expression level was measured by flow cytometry using anti-GPIX monoclonal antibody FMC25 and analyzed as described above (n=3).

Figure 7. A schematic model for the membrane-proximal portion of the GPIb-IX complex. (A) Cartoon diagram showing two views related by a 90 degree rotation illustrating the GPIbα chain TM domain (in green) extending towards the N-terminus (top), GPIbβ ectodomain and TM domain (in sky blue and dark blue), and GPIX ectodomain and TM domain (in purple). Tyr106
from GPIbβ1 is shown as stick along with interchain disulphides. (B) Same view to A except without TM domains. GPIbβ is colored white showing residues affected by BSS and HPA mutations. Side chains of GPIbβ residues Ala108 (Blue), Pro74 (orange) and Gly15 (red) are shown in space-filling mode, Tyr106 is shown as stick (grey).
Figure 1
Figure 2
Figure 3
Figure 4
Figure 5
Figure 6
Quaternary organization of GPIb-IX complex and insights into Bernard-Soulier syndrome revealed by the structures of GPIb β and a GPIb β/GPIX chimera

Paul A. McEwan, Wenjun Yang, Katherine H. Carr, Xi Mo, Xiaofeng Zheng, Renhao Li and Jonas Emsley