Complete localization of the intrachain disulphide bonds and the *N*-glycosylation points in the α -subunit of human platelet glycoprotein IIb

Juan J. CALVETE,* Agnes HENSCHEN* and José GONZÁLEZ-RODRÍGUEZ†‡

*Max-Planck-Institut für Biochemie, D-8033, Martinsried/München, Federal Republic of Germany, and †Instituto de Química Física, C.S.I.C., Serrano, 119, 28006 Madrid, Spain

Glycoprotein IIb (GPIIb), one of the two molecular components of the inducible receptor for fibrinogen on the platelet surface, is formed from two subunits, GPIIb α (114 kDa) and GPIIb β (22.5 kDa), joined by a single disulphide bond. CNBr cleavage of GPIIb, together with tryptic or endoproteinase Lys-C digestion of some of the isolated CNBr peptides, followed by amino acid and *N*-terminal sequence analysis of the isolated fragments, allowed us to locate unambiguously all the unknown disulphide bonds and the *N*-glycosylation points in platelet GPIIb. It could be established that each cysteine residue in GPIIb, beginning at α -Cys-56, is disulphide-bonded to its nearest neighbour in the amino acid sequence. Given the extensive structural similarity among the two-chain α -subunits of Arg-Gly-Asp adhesion receptors and the conservative positions of cysteine residues in their amino acid sequences, the intrachain and interchain disulphide-bond pattern found here in GPIIb will most probably be conserved in all two-chain α -subunits of these receptors. The *N*-linked glycosylation points found here in platelet GPIIb are the same as the five *N*-glycosylated asparagine residues suggested after cDNA sequencing of human erythroleukaemic-cell GPIIb [Poncz, Eisman, Heindenreich, Silver, Vilaire, Surrey, Schwartz & Bennett (1987) *J. Biol. Chem.* **262**, 8476–8482]. Some of the general features of the structure of GPIIb, such as the existence of distinct domains in the α - and β -subunits, as well as the identification of well-defined points in its external topography, are discussed.

INTRODUCTION

Platelet glycoprotein IIb (GPIIb), together with glycoprotein IIIa (GPIIIa), forms the Ca²⁺-dependent heterodimer GPIIb–IIIa (Kunicki *et al.*, 1981), which serves, after platelet activation, as the receptor for plasma fibrinogen and other adhesive proteins (Marguerie *et al.*, 1987). Recently, the amino acid sequence for a GPIIb-like glycoprotein from HEL cells was predicted from the cDNA sequence (Poncz *et al.*, 1987). GPIIb (136.5 kDa) is formed by two subunits, α (GPIIb α , 114 kDa) and β (GPIIb β , 22.5 kDa), joined by a single disulphide bridge between GPIIb β Cys-9 and GPIIb α Cys-826 [Nurden *et al.*, 1986; Usobiaga *et al.*, 1987; the preceding paper (Calvete *et al.*, 1989)]. The other two cysteine residues of GPIIb β (β Cys-14 and β Cys-19) form the single intrachain disulphide bond in this subunit (Calvete *et al.*, 1989). The rest of the 14 cysteine residues in GPIIb α are distributed in three main peptide stretches: the *N*-terminal stretch (α 1–285), containing six cysteine residues, and two central regions, containing four cysteine residues each and separated by the main chymotryptic cleavage point of GPIIb in whole platelets. The *C*-terminal stretch (from residue α 704 onwards up to the *C*-terminus, α 871) only carries α -Cys-826 (Calvete *et al.*, 1989).

In the present work we have determined the localization of all the intrachain disulphide bonds whose

positions remained unknown in GPIIb. We have found that each cysteine in this glycoprotein, beginning by α Cys-56, is disulphide-bonded to its nearest neighbour in the amino acid sequence.

MATERIALS AND METHODS

Endoproteinase Lys-C was from Boehringer-Mannheim, and tosylphenylalanylchloromethane ('TPCK')-treated trypsin from Sigma. The other chemicals and biochemicals were of analytical or chromatographic grade. The isolation of platelet GPIIb was performed as described previously (Eirín *et al.*, 1986).

Analytical methods

Amino acid and amino sugar analyses were performed with a Biotronik amino acid analyser, after sample hydrolysis at 110 °C in 6 M-HCl for 24 h and in 4 M-HCl for 4 h, respectively. *N*-Terminal sequence analyses were effected in a prototype automated spinning-cup sequenator as described previously (Edman & Henschen, 1975). The amino acid phenylthiohydantoin derivatives were analysed by reverse-phase h.p.l.c. using a Spherisorb ODSII column (Henschen, 1986).

Abbreviations used: GPIIb, glycoprotein IIb; GPIIIa, glycoprotein IIIa; GPIIb α and GPIIb β , the α - and β -subunits respectively of GPIIb; HEL: human erythroleukaemic-cell line.

‡ To whom correspondence and reprint requests should be sent.

CNBr cleavage of GPIIb

Isolated GPIIb (10 mg/ml) in 70 % (v/v) formic acid was cleaved by incubation with CNBr (100 mg/ml final concn.) for 2 h at room temperature under an inert atmosphere and in the dark. The sample was then diluted 10-fold with water and evaporated to dryness in a Speed Vac concentrator (Savant). The cleavage products were resuspended in 70 % formic acid and separated by reverse-phase on a C4 [pore size 30 nm (300 Å), 5 µm particle size] Vydac column (25 cm × 0.46 cm) using an acetonitrile gradient elution, or, alternatively, by size-exclusion chromatography on a column (130 cm × 1 cm) of Sephacryl S-200 which was eluted with 10 % formic acid.

Enzymic digestions

Endoproteinase Lys-C digestion of CNBr peptides from GPIIb was performed in 100 mM-Tris/HCl, pH 8.6, at an enzyme/peptide ratio of 1:50 (w/w) for 2 h at 37 °C. Formic acid was then added to the samples to 30 % (v/v) final concn. The digestion products were separated by acetonitrile gradient elution on a C4 Vydac column.

Tryptic digestion of the largest CNBr peptide from GPIIb, previously denominated 35 kDa-CNBr (α 1–285) (Calvete *et al.*, 1989), was done as follows. To a 5 mg/ml solution of this peptide fragment in 50 mM- NH_4HCO_3 /0.1 % (v/v) *N*-ethylmorpholine, pH 8.0, tosylphenylalanylchloromethane-treated trypsin was added at an enzyme/peptide ratio of 1:50 (w/w), and the mixture was incubated at 37 °C for 3 h. Formic acid was added to the samples to 30 % (v/v) final concn., and the digestion products were separated by reverse-phase h.p.l.c. on a C4 Vydac column.

RESULTS

Localization of Cys-490–Cys-545, Cys-602–Cys-608 and Cys-674–Cys-687 disulphide bonds in GPIIb α

By using the amino acid sequence predicted from the cDNA sequence for HEL GPIIb (Poncz *et al.*, 1987) as a reference for the primary structure of platelet GPIIb, 12 CNBr cleavage points can be predicted. However, when the CNBr products of cleavage of GPIIb in which the disulphide bonds had not been reduced were separated by reverse-phase chromatography (Fig. 1), only ten fractions, from C1 to C10 according to their elution time, were obtained. After amino acid and *N*-terminal sequence analyses of each fraction (Tables 1 and 2), it was found that, of the thirteen expected CNBr peptide fragments, only nine were obtained. The peptides α 1–285, α 315–489, β 25–31 and β 32–116 were not found. Most probably the first two missing peptides were retained in the reverse-phase column because of their large size, whereas the last one (β 32–116), carrying the intrabilayer stretch and 65 % of hydrophobic amino acid residues, was retained because of its hydrophobicity.

Of the ten fractions isolated, only C6, C7, C9 and C10 had cysteine in their composition. Fraction C10 contained two *N*-terminal amino acid sequences which were present in equal molar amounts. These could be identified with peptide stretches α 704–871 and β 1–24 and contained a total of four cysteine residues per molecule of peptide. These cysteine residues can be accounted for by the interchain disulphide bond, α 826– β 9 and the intrachain disulphide bond β 14– β 19, as expected from our previous findings (Calvete *et al.*, 1989). Fraction C9 had a major *N*-terminal sequence, accounting for 90 % of the whole, which was identified as peptide α 581–635. The fraction contained two cysteine residues per molecule

Fig. 1. Isolation of CNBr cleavage products of GPIIb by reverse-phase h.p.l.c.

A 100 µl portion of CNBr-cleaved GPIIb (10 mg/ml) in 70 % formic acid was injected into a C4 [30 nm (300 Å), 5 µm] Vydac column (25 cm × 0.46 cm), equilibrated in 0.1 % trifluoroacetic acid in water (A) and 0.1 % trifluoroacetic acid in acetonitrile (B) (9:1, v/v), and eluted at 1 ml/min, first isocratically (90 % A, 10 % B) for 5 min, followed by an acetonitrile gradient applied in successive linear steps: 10–22 % of B for 5 min; 22–27 % B for 15 min; 27–29 % B for 20 min; 29–35 % B for 5 min; 35–42 % B for 25 min; 42–70 % B for 10 min; and 70–10 % B for 5 min. The amino acid composition and the *N*-terminal sequence of the separated fractions are given in Tables 1 and 2(a).

Table 1. Amino acid and amino sugar compositions of CNBr fragments of GPIIb

The fragments are denominated according to their order of elution from the C4 reverse-phase column (see Fig. 1). The theoretical amino acid composition of the corresponding peptide stretches assigned for each of these fragments in the cDNA-deduced GPIIb amino acid sequence (Poncz *et al.*, 1987) is given alongside (the stretches are preceded by α or β according to the subunit of GPIIb to which they belong). The experimental results are an average of two determinations with two different preparations; n.d. = not determined.

Residue	Composition (residues/molecule of peptide)																α 704-871 β 1-25	
	C1	α 694-703	C2	β 117-137	C3	α 636-660	C4(C5)	α 286-314	C6	α 490-549	C7	α 661-693	C8	α 550-580	C9	α 581-635	C10	
Asp + Asn	1.24	1	2.97	3	2.11	2	3.78	4	4.87	5	3.92	4	3.72	4	5.30	6	14.50	19
Thr	-	-	-	-	-	-	1.03	1	3.67	5	1.21	1	1.36	1	2.54	4	4.94	3
Ser	0.21	-	-	-	-	-	1.68	2	2.36	3	1.41	1	2.44	3	2.45	2	16.72	18
Hse	1.00	1	-	-	1.00	1	1.00	1	1.00	1	1.00	1	1.00	1	1.00	1	1.00	1
Glu + Gln	1.30	1	4.84	5	4.66	5	0.31	-	5.77	7	5.43	6	2.94	2	6.93	8	29.36	30
Pro	-	-	1.93	2	1.01	1	0.56	1	2.11	3	0.51	1	2.66	3	2.13	3	11.01	22
Gly	1.11	1	1.95	2	2.72	3	3.54	4	5.47	7	2.90	2	1.47	1	3.87	4	10.36	12
Ala	2.05	2	-	-	5.31	6	2.56	3	2.66	3	1.42	1	2.56	3	3.36	4	7.50	8
Cys	-	-	-	-	-	1	-	-	1.75	2	1.66	2	-	-	2.07	2	3.89	4
Val	-	-	1.00	1	0.71	-	3.49	4	2.36	2	2.43	3	1.70	2	6.21	9	13.56	17
Met	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Ile	1.89	2	-	-	-	-	-	-	1.01	2	0.68	1	1.10	1	1.22	1	5.52	8
Leu	0.11	-	0.96	1	1.87	2	3.03	3	8.31	10	4.30	4	4.48	5	6.15	8	16.66	22
Tyr	-	-	-	-	1.31	2	0.88	2	0.21	-	-	-	0.42	-	0.34	-	2.12	2
Phe	-	-	1.83	2	-	-	0.71	1	0.24	-	0.80	1	1.52	2	1.01	-	3.68	4
Lys	2.01	2	1.93	2	-	-	-	2	2.38	3	2.17	2	1.27	1	0.84	-	3.96	5
His	-	-	-	-	1.56	2	1.46	2	2.08	3	0.43	-	-	-	1.11	1	4.82	7
Arg	-	-	1.98	2	-	-	0.71	1	3.18	4	2.16	3	1.79	2	2.27	2	5.48	8
Trp	-	-	n.d.	1	-	-	-	-	-	-	-	-	-	-	-	-	n.d.	2
GlcN	-	-	-	-	-	-	-	-	-	-	3.32	-	3.88	-	-	-	-	-
GalN	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1.78	-

Table 2. *N*-Terminal sequence of the fragments of cleavage of GPIIb

(a) *N*-Terminal sequence of the CNBr fragments (C1–C10) isolated as shown in Fig. 1 and labelled according to their order of elution from the C4 reverse-phase column. (b) *N*-Terminal sequence of the fragments (L1–L5) obtained by digestion of fractions C6 + C7 (see Fig. 1) with endoproteinase Lys-C and separated by reverse-phase chromatography. (c) *N*-Terminal sequence of the CNBr fragments isolated by size-exclusion chromatography in the major fractions SI and SII (see Fig. 2). (d) *N*-Terminal sequence of the tryptic fragments (R1–R9) of digestion of fraction SI (α 1–285) (see Fig. 2), isolated by reverse-phase chromatography (see Fig. 3). In those fractions where more than one *N*-terminal sequence was obtained, the relative proportions (R.P.) of each sequence are shown. On the right-hand side of the Table is the numbering in the cDNA sequence of the terminal residues of each peptide ('cDNA sequence').

Fragment	Sequence	R.P.	cDNA sequence
(a)			
C1	KKNAQI		α 694–703
C2	WKVGFF		β 117–137
C3	DAANE GEGAYEAELAVHLPQGA (H) Y		α 636–660
C4:C5	ASYFGHSAV		α 286–314
C6	XVGAT(+RALSN)	4:1	α 490–549
C7	RALSNVEGFE(+XVGAT)	5:1	α 661–693
C8	AFLRDEADFR		α 550–580
C9	APAVVLHGDT(+PLKVD)	10:1	α 581–635
C10	LVS VGNLEEA	} 1:1	α 704–871
	LQDPV LV SXD		β 1–24
(b)			
L1	XVGATGHNIP	} 1:1	α 490–501
	H(S)PIXHTT		α 541–549
L2	LSLNAELQL		α 502–514
L3	PRQGRR		α 515–540
L4	RALSNVEGFE	} 1:1	α 661–667
	KEXETRVVLX		α 678–693
L5	ASYFGHSV		α 286–314
(c)			
SI	LNLDPV	} 1:1:1	α 1–285
SII	ESRADRK		α 315–489
	LVS VGNL		α 704–871
	LQDPV LV		β 1–24
(d)			
R1	V A I V V		α 33–41
R2	N V G S Q T L		α 78–88
R3	I Y V E N D F S W D		α 154–165
R4	I Y V E N		α 154–164
R5	A S V Q L		α 456–471
R6	T L G P S Q E	} 1:1	α 42–59
	A E G G Q X P		α 60–73
R7	L N L D P V Q L T F	} 1:1	α 1–27
R8	Q G L G A S V V		α 91–118
	T P V S X F L A Q		α 125–139
R9	A E Y S P	} 1:1	α 141–147
	Y X E A G		α 166–208

and thus contains a single disulphide bond. The *N*-terminal sequence of the minor fraction could not be identified in the sequence of GPIIb and appears to arise from a contaminant. Fractions C6 and C7 contain the major *N*-terminal sequences corresponding to α 490–549 and α 661–693 respectively, but the two fractions are clearly cross-contaminated with each other. These two fractions also contain two cysteine residues per molecule and thus each contain a single disulphide bond.

To remove any uncertainty about the individuality and identification of the major CNBr fragments in fractions C6 and C7 and, therefore, about the localization of the disulphide bonds they contain, these two chromatographic fractions were mixed together and

digested by endoproteinase Lys-C. The digestion products were separated by reverse-phase chromatography (results not shown) and their amino acid composition (also not shown) and *N*-terminal sequence were analysed (Table 2b). Of the five major fractions obtained (L5 is a contamination of fraction C6 by fraction C5), only L1 and L4 contained cysteine and both gave two, equally contributing, *N*-terminal sequences. These *N*-terminal sequences could be identified as peptides α 490–501 and α 541–549 in L1, and α 661–677 and α 678–693 in L4. Given that these two fractions contained two cysteine residues per molecule, the Cys-490–Cys-545 and Cys-674–Cys-687 disulphide bonds could be confirmed.

Fig. 2. Isolation of the large CNBr products of cleavage of GPIIb by size-exclusion chromatography

A 200 μ l (5 mg) portion of the CNBr products of cleavage of GPIIb in 70 % formic acid was applied to a Sephacryl S-200 column (130 cm $\times$ 1 cm) equilibrated in 10 % formic acid, and two main fractions, SI and SII, were obtained. The *N*-terminal sequence and amino acid analyses of these fractions are given in Tables 2(c) and 3. Abbreviations: BSA, bovine serum albumin; Ovo, ovalbumin; TrI, trypsin inhibitor.

Localization of Cys-473–Cys-484 disulphide bond

We have demonstrated previously that none of the six cysteine residues in the *N*-terminal part of the molecule (CNBr product α 1–285, formerly referred to as ‘35 kDa CNBr’) is disulphide-bonded to any cysteine residue in the rest of the molecule (Calvete *et al.*, 1989). Thus the experiments so far described not only locate the Cys-490–Cys-545, Cys-602–Cys-608 and Cys-674–Cys-687 intrachain bonds, but also strongly support the existence of another disulphide bond between Cys-473 and Cys-484. Further evidence for the existence of this bond is provided by analysis of the CNBr peptides obtained by separation using size-exclusion chromatography on Sephacryl S-200 rather than reverse-phase chromatography. Two main fractions, SI and SII, were obtained (Fig. 2). The CNBr fragments present in each fraction were identified by *N*-terminal sequence and amino acid analyses (Tables 2 and 3). These show that SI contains the CNBr peptide α 1–285, whereas SII contains a mixture of the two fragments α 315–489 and (α 704–871 + β 1–24). Thus the two cysteine residues in α 315–489 must form a single disulphide bond.

Tryptic digestion of CNBr fragment α 1–285 and localization of Cys-56–Cys-65, Cys-107–Cys-130 and Cys-146–Cys-167 disulphide bonds

To be able to locate the remaining disulphide bonds in GPIIb α , the CNBr fragment α 1–285 isolated by size-exclusion chromatography (Fig. 2), was digested with trypsin and the tryptic peptides separated by reverse-phase chromatography on a C4 column (Fig. 3). Nine fractions (from R1 to R9 in order of elution from the

Table 3. Amino acid and amino sugar compositions of the two major fractions of CNBr fragments of cleavage of GPIIb separated by size-exclusion chromatography (see Fig. 2)

The theoretical amino acid composition of the corresponding peptide stretches assigned to each of these fractions in the cDNA-deduced GPIIb amino acid sequence (Poncz *et al.*, 1987) is given alongside (α and β denote the subunit of GPIIb to which the fragments belong). The experimental results are an average of two determinations with two different preparations; n.d. = not determined.

Residues	Composition (residues/homoserine residue)				
	SI	α 1–285	SII	α 315–489	α 704–871 + β 1–24
Asp + Asn	24.56	22	16.15	16	19
Thr	11.73	14	3.74	5	3
Ser	22.17	27	14.15	14	18
Glu + Gln	36.30	32	18.08	14	30
Pro	17.83	15	17.14	14	22
Gly	22.17	30	17.86	20	12
Ala	15.65	20	11.13	17	8
Cys	5.95	6	2.68	1	4
Val	19.78	23	17.15	17	17
Met	–	–	–	–	–
Ile	4.56	6	6.01	5	8
Leu	19.78	27	18.14	20	22
Tyr	8.04	12	4.11	7	2
Phe	9.13	15	5.01	7	4
Lys	5.86	5	3.78	4	5
His	4.34	5	4.15	1	7
Arg	11.52	15	10.33	11	8
Hse	1.00	1	1.00	1	1
Trp	n.d.	9	n.d.	–	2
GlcN	11.02	–	–	–	–
GalN	2.48	–	1.64	–	–

Fig. 3. Isolation of the tryptic products of digestion of the *N*-terminal fragment of CNBr cleavage of GPIIb

A 50 μ l (4 mg/ml) portion of the *N*-terminal CNBr product of cleavage of GPIIb, isolated in fraction SI (α 1–285) as described in Fig. 2, was further digested with tosylphenylalanylchloromethane-treated trypsin and the digestion products separated by reverse-phase chromatography as in Fig. 1. Column elution was at 1 ml/min, first isocratically (90% A, 10% B) for 10 min, followed by a linear acetonitrile gradient from 10 to 70% B for 90 min, and back to 10% in 5 min. The *N*-terminal sequences of the isolated fractions are given in Table 2(d).

column) were obtained and identified by amino acid (results not shown) and *N*-terminal sequence analyses (Table 2d). It was found that fraction R5 was a peptide, α 456–471, derived from fraction SII. Of the others, only fractions R6, R8 and R9 had cysteine residues in their amino acid composition, and each of these three fractions contained two equally contributing *N*-terminal sequences. From these *N*-terminal sequences it was possible to identify the following peptide stretches: α 42–59 + α 60–73 in R6; α 91–118 + α 125–139 in R8; and α 141–147 + α 166–208 in R9. Thus the pair of cysteine residues in each of these three peptides form the remaining three intrachain disulphide bonds in GPIIb α .

Distribution of amino sugars and glycosylation points in GPIIb

The amino acid sequence for HEL-cell GPIIb-like glycoprotein (Poncz *et al.*, 1987), predicted by cDNA sequence, shows five potential *N*-linked glycosylation sites (asparagine residues: α 15, α 249, α 570, α 680 and β 70). Amino sugar analysis of the cleavage products of GPIIb referred to above showed glucosamine to be in the following fractions: 11 residues/molecule in SI (α 1–285) and four residues/molecule of peptide in C7 (α 661–693), C8 (α 550–580), L4 (α 661–677 + α 678–693) and R7 (α 1–27). Therefore the asparagine residues in positions α 15, α 570 and α 680 are *N*-glycosylated. Given that peptides SI (α 1–285) and R7 (α 1–27) contained 11 and four residues of glucosamine/molecule respectively, it is probable that α Asn-249 is glycosylated as well. GPIIb β does not contain galactosamine (Calvete & González-Rodríguez, 1986), and as it is not thought to be *O*-glycosylated, β Asn-60 is the only point available for *N*-glycosylation in this subunit.

Galactosamine was found in only the following isolated fractions: 2 residues/molecule in C10 (α 704–

871 + β 1–24); 3 residues/molecule in SI (α 1–285); and 2 residues/molecule of homoserine in SII (α 315–489 and α 704–871 + β 1–24). Therefore, given the galactosamine content of C10 and SII and that GPIIb β does not contain any galactosamine, each peptide stretch, α 315–489 and α 704–871, contains two residues of this amino sugar.

DISCUSSION

In Fig. 4, using the evidence provided by the amino acid sequence predicted from the cDNA sequence for HEL GPIIb (Poncz *et al.*, 1987), we summarize the information obtained in the present work on the localization of the intrachain disulphide bonds and *N*-glycosylation points in GPIIb α and GPIIb β , together with the localization of the single interchain disulphide bridge found previously (Calvete & González-Rodríguez, 1986; Calvete *et al.*, 1989).

By combining CNBr cleavage of GPIIb with tryptic or endoproteinase Lys-C digestion of some of the isolated CNBr products, followed by amino acid and *N*-terminal sequence analyses of the obtained fragments, it was possible to locate unambiguously all the unknown disulphide bonds in GPIIb α . In contrast with the situation in GPIIIa, where both nearest-neighbour and non-nearest-neighbour cysteine pairing is observed (J. J. Calvete, A. Henschen & J. González-Rodríguez, unpublished work), it could be established here that each cysteine residue in GPIIb, beginning at α Cys-56, is disulphide-bonded to its nearest neighbour in the amino acid sequence, the actual pairing being as follows: α Cys-56–Cys-65; α Cys-107–Cys-130; α Cys-146–Cys-167; α Cys-473–Cys-484; α Cys-490–Cys-545; α Cys-602–Cys-608; α Cys-674–Cys-687; α Cys-826– β Cys-9; and β Cys-14–Cys-19. Given the extensive structural similarity found so far among the α -subunits of Arg-Gly-Asp adhesion receptors (platelet fibrinogen receptor, and

Fig. 4. Summary of the covalent structure of human platelet GPIIb, showing the nearest-neighbour cysteine-residue-pairing pattern, the location of the *N*-glycosylation points and the peptide stretches where *O*-glycosylation points were found.

Key to symbols: ◆, *N*-glycosylation points; ●, *O*-glycosylation points; M_{314} etc., Met-314 etc.; N_{15} etc. Asn-15 etc.; [C], cysteine disulphide bonds.

vitronectin and fibronectin receptors) (Poncz *et al.*, 1987; Fitzgerald *et al.*, 1987; Uzan *et al.*, 1988), the intrachain and interchain disulphide-bond pattern, outlined in Fig. 4 for platelet GPIIb, will most probably be conserved in the α -subunits of other cytoadhesins, as is the case with the positions of the cysteine residues along their amino acid sequences.

The same strategy used to locate the disulphide bonds permitted us to establish that asparagine residues *N*-glycosylated in platelet GPIIb are the same as the five potential *N*-linked glycosylation sites suggested by Poncz *et al.* (1987) for HEL GPIIb, namely $\alpha 15$, $\alpha 249$, $\alpha 570$, $\alpha 680$ and $\beta 60$. From the amino sugar composition of the glycopeptides isolated here and the sugar composition of GPIIb β (Calvete & González-Rodríguez, 1986), the subunit where a single *N*-glycosylation point was found and no *O*-glycosylation is expected, it is reasonable to predict that the structure of the asparagine-linked oligo-

saccharide units in GPIIb is very likely of the complex type, analogous to that found in human fibrinogen (Townsend *et al.*, 1982). It has recently been found that the high-mannose oligosaccharides are added co-translationally, whereas conversion into the complex type found in the mature glycoprotein, the only form expressed at the megakaryocyte surface, probably takes place in the Golgi cisternae (Duperray *et al.*, 1987). Galactosamine was found in only three of the GPIIb α peptides isolated here: three residues in peptide $\alpha 1$ -285, and two residues each in peptides $\alpha 315$ -489 and $\alpha 704$ -871. There are some discrepancies in the number of amino sugars found in the present study and in our previous work (Calvete & González-Rodríguez, 1986; Eirín *et al.*, 1986). The reasons for these discrepancies could be (i) that we previously overestimated the amino sugar contents of GPIIb and its subunits, (ii) that some amino sugars are located in those peptides missing in our actual fractionation procedures, or (iii), less likely, that some sugar degradation occurred during peptide preparation or fractionation.

About 51% of platelet GPIIb amino acid sequence has already been determined by protein sequencing (Charo *et al.*, 1986; Hiraiwa *et al.*, 1987; Bray *et al.*, 1988; Calvete *et al.*, 1989; the present work and J. J. Calvete, A. Henschen & J. González-Rodríguez, unpublished work). So far the agreement with the amino acid sequence derived from HEL-cell GPIIb cDNA is almost complete. We conclude that most probably they are identical glycoproteins.

Finally some general features of the structure of GPIIb can already be outlined. First, whereas the α -subunit seems to be globular overall, both in solution (Usobiaga *et al.*, 1987) and in the membrane (Phillips *et al.*, 1988), the β -subunit must have at least three topologically well defined domains: cytoplasmic, transmembrane and extracellular. Secondly, the nearest-neighbour cysteine-residue-pairing pattern in GPIIb gives rise to relatively short disulphide-bonded loops in the covalent structure, which suggests that this cross-linking takes place as the GPIIb precursor is synthesized. Although this will have to be investigated further, it seems as if the disulphide bonds are mainly stabilizing the short-range folding, having no important role in the overall conformation of GPIIb. Thirdly, several distinct areas in the topography of GPIIb have already been identified: (i) the area of the C-terminus of GPIIb α and the N-terminus of GPIIb β , which includes the single interchain disulphide bond, the epitopes for monoclonal antibodies M1 ($\beta 4$ -31), M3 ($\alpha 827$ -871) (Calvete *et al.*, 1989) and PMI-1 (Loftus *et al.*, 1987), and the *N*-glycosylation point at β Asn-60; (ii) the chymotryptic cleavage point of GPIIb in whole platelets, somewhere around α Asn-570 (Calvete *et al.*, 1989), which seems to be in an exposed polypeptide loop joining together the N-terminal and the C-terminal domains of GPIIb α ; (iii) the N-terminal domain of GPIIb α , corresponding to the water-soluble Ch-95 kDa fragment, which is detachable from the membrane after cleavage and carries two *N*-glycosylation and two *O*-glycosylation points; (iv) the C-terminal domain, of about 300 amino acid residues, insoluble and non-detachable from the membrane after chymotryptic cleavage and reduction of the interchain disulphide bond, which carries two *N*-glycosylation and at least one *O*-glycosylation points, besides the epitopes for monoclonal antibodies M5 and M6 ($\alpha 608$ -687) and, most probably,

the Bak^a alloantigen (Kieffer *et al.*, 1984; Calvete *et al.*, 1989), and which binds detergents in aqueous solutions of GPIIb α . Preliminary experiments (E. Muñiz, A. Rivera & J. González-Rodríguez, unpublished work) show that M5 recognizes preferentially Bak^a(+) platelets, and that the expression in whole platelet of the epitope for M6, but not that for M3, is EDTA-dependent. These results could indicate that the epitope for M5 and the Bak^a alloantigen are relatively close, and that the Arg-Gly-Asp- and EDTA-dependent epitope for PMI-1 (α 844-860) (Loftus *et al.*, 1987) is apparently more closely related to the epitope for M6 than to the epitope for M3.

We thank Dr. J. M. Wilkinson for reading the manuscript, Mr. G. Rivas for his help in the preparation of GPIIb, Mrs. B. Gross and Mrs. C. Martín de Loeches for technical assistance, Mrs. M. L. Ruiz Pineda for typing the manuscript, and the Blood Banks of Centro Ramón y Cajal, La Paz, and Doce de Octubre (Madrid) for providing us with outdated platelet concentrates. This work was supported by the Comisión Interministerial de Ciencia y Tecnología, the Secretaría de Estado para Universidades e Investigación, and the Consejo Superior de Investigaciones Científicas.

REFERENCES

- Bray, P. F., Rosa, J.-P., Johnston, G. I., Shiu, D. T., Cook, R. G., Lau, C., Kan, Y. W., McEver, R. P. & Shuman, M. A. (1988) *J. Clin. Invest.* **80**, 1812-1817
- Calvete, J. J. & González-Rodríguez, J. (1986) *Biochem. J.* **240**, 155-161
- Calvete, J. J., Alvarez, M. V., Rivas, G., Hew, C.-L., Henschen, A. & González-Rodríguez, J. (1989) *Biochem. J.* **261**, 551-560
- Charo, I. F., Fitzgerald, L. A., Steiner, B., Rall, S. C., Jr., Bekeart, L. S. & Phillips, D. R. (1986) *Proc. Natl. Acad. Sci. U.S.A.* **83**, 8351-8355
- Duperray, A., Berthier, R., Chagnon, E., Ryckewaert, J. J., Ginsberg, M., Plow, E. & Marguerie, G. (1987) *J. Cell Biol.* **104**, 1665-1673
- Edman, P. & Henschen, A. (1975) in *Protein Sequence Determination*, 2nd edn. (Needleman, S. B., ed.), pp. 232-279, Springer, Berlin
- Eirín, M. T., Calvete, J. J. & González-Rodríguez, J. (1986) *Biochem. J.* **240**, 147-153
- Fitzgerald, L. A., Poncz, M., Steiner, B., Rall, S. C., Jr., Bennett, J. S. & Phillips, D. R. (1987) **26**, 8158-8165
- Henschen, A. (1986) in *Advanced Methods in Protein Microsequence Analysis* (Wittman-Liebold, B., Salnikow, J. & Erdmann, V. A., eds.), pp. 244-255, Springer, Berlin
- Hiraiwa, A., Matsukage, A., Shiku, H., Takahashi, T., Naito, K. & Yamada, K. (1987) *Blood* **69**, 560-564
- Kieffer, N., Boizard, B., Didry, D., Wautier, J.-L. & Nurden, A. T. (1984) *Blood* **64**, 1212-1219
- Kunicki, T. J., Pidard, D., Rosa, J. P. & Nurden, A. T. (1981) *Blood* **58**, 268-278
- Loftus, J. C., Plow, E. F., Frelinger, A. L., III, D'Souza, S. E., Dikxon, D., Lacy, J., Sorge, J. & Ginsberg, M. H. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 7114-7118
- Marguerie, G. A., Ginsberg, M. H. & Plow, E. F. (1987) in *Platelets in Biology and Pathology* (MacIntyre, D. E. & Gordon, J. L., eds.), vol. 3, pp. 95-125, Elsevier, Amsterdam
- Nurden, A. T., George, J. N. & Phillips, D. R. (1986) in *Biochemistry of Platelets* (Phillips, D. R. & Shuman, M. A., eds.), pp. 159-224, Academic Press, New York
- Phillips, D. R., Charo, I. F., Parise, L. V. & Fitzgerald, L. A. (1988) *Blood* **71**, 831-843
- Poncz, M., Eisman, R., Heindenreich, R., Silver, S. M., Vilaire, G., Surrey, S., Schwartz, E. & Bennett, J. S. (1987) *J. Biol. Chem.* **262**, 8476-8482
- Townsend, R. R., Hilliker, E., Li, Y.-T., Laine, R. A., Bell, W. R. & Lee, Y. C. (1982) *J. Biol. Chem.* **257**, 9704-9710
- Usobiaga, P., Calvete, J. J., Saiz, J. L., Eirín, M. T. & González-Rodríguez, J. (1987) *Eur. Biophys. J.* **14**, 211-218
- Uzan, G., Frachet, P., Lajmanovich, A., Prandini, M.-H., Dernarier, E., Duperray, A., Loftus, J., Ginsberg, M., Plow, E. & Marguerie, G. (1988) *Eur. J. Biochem.* **171**, 87-93

Received 16 August 1988/23 January 1989; accepted 1 March 1989