Amino Acid Sequence and Location of the Disulfide Bonds in Bovine β2 Glycoprotein I: The Presence of Five Sushi Domains†

Hisao Kato* and Kei-ichi Enjyoji

National Cardiovascular Center Research Institute, Fujishirodai 5, Suita, Osaka 565, Japan

Received July 15, 1991; Revised Manuscript Received September 17, 1991

ABSTRACT: β2 glycoprotein I is a plasma protein with the ability to bind with various kinds of negatively charged substances. The complete amino acid sequence and the location of all the disulfide bonds of bovine β2 glycoprotein I were determined. Bovine β2 glycoprotein I consists of 326 amino acid residues with five asparagine-linked carbohydrate chains. Homology with the human protein was calculated to be 83%. Eleven disulfide bonds in bovine β2 glycoprotein I constitute four characteristic domains, Sushi domains, and one modified form of a Sushi domain.

The β2 glycoprotein I has been shown to be a plasma protein which binds to various kinds of negatively charged substances, such as phospholipid and dextran sulfate (Shousboe & Rasmussen, 1988; Schousboe, 1988) and lipoprotein (Polz & Kostner, 1979). The amino acid sequence of human β2 glycoprotein I has been determined by Lozier et al. (1984). The partial amino acid sequence of rat β2 glycoprotein I was recently elucidated by cDNA analysis of the protein gene (Aoyama et al., 1989). The new function of this plasma protein was recently presented by the finding that a cofactor for the complex formation of cardiolipin and anticardiolipin antibody was identical with β2 glycoprotein I (Matsuura et al., 1990; McNeil et al., 1990; Galli et al., 1990). β2 glycoprotein I may bind to cardiolipin or DNA in plasma, and thus, their autoantibodies will be raised, particularly, in SLE patients. The conformational changes of the protein moiety by the complex formation may be essential for the production of the autoantibodies. We have studied an inhibitor in bovine plasma for dextran sulfate mediated activation of factor XII and prekallikrein and identified the inhibitor to be bovine β2 glycoprotein I (Kato & Enjyoji, 1989). In the present paper, the complete amino acid sequence and the location of all of the disulfide bonds of bovine β2 glycoprotein I were established by amino acid sequence analysis of various peptides isolated from chemical and enzymatic digests of the protein. Results show that β2 glycoprotein I consists of five characteristic domains called Sushi domains.

Experimental Procedures

Materials. TPCK–trypsin and chymotrypsin were products of Worthington Biochemical Corp. (NJ). V8 protease (*Staphylococcus aureus* protease) was a product of Pierce Co. (IL). Lysyl endopeptidase (*Achromobacter* protease I) and thermolysin were purchased from Wako Co. (Osaka, Japan). Carboxypeptidase Y was a product of Oriental Yeast Co., Ltd. (Tokyo, Japan). Sephadex G-100, Sephadex G-150, DEAE-Sepharose CL-6B, and chelating Sepharose were products of Pharmacia LKB. Sulfate Cellulofine was a product of Seikagaku Kogyo Co. (Tokyo, Japan). Cosmosil 5C18-300 column (4.6 × 250 mm) was a product of Nakarai tesque (Kyoto, Japan). Anti β2 glycoprotein I rabbit serum was raised by injecting purified bovine β2 glycoprotein I with complete Freund's adjuvant into a rabbit.

Purification of β2 Glycoprotein I from Bovine Plasma. Bovine plasma (2.7 L) was fractionated by ammonium sulfate, and 50% supernatant was dialyzed against 20 L of 0.02 M Tris-HCl buffer, pH 8.0, for three days by exchange of the dialyzing buffer once a day. The dialysate was applied to a column (6 × 21 cm) of DEAE-Sepharose CL-6B, which had been equilibrated with 0.02 M Tris-HCl buffer, pH 8.0. After the column was washed, protein was eluted by a linear salt gradient formed by each 2 L of the equilibration buffer and the buffer containing 0.3 M NaCl. Each 15-mL fraction was collected. β2 glycoprotein I was detected by immunodiffusion on an agar plate using anti β2 glycoprotein I rabbit serum. [In the early phase of the investigation, β2 glycoprotein I was assayed by the inhibitory activity toward the surface-mediated activation of factor XII and prekallikrein under the conditions described by Kodama et al. (1985).] Fractions 60–130 were pooled and applied to a column (3.5 × 42 cm) of zinc-chelating Sepharose, which had been equilibrated with 0.02 M Tris-HCl buffer, pH 8.0. The nonadsorbed fraction was dialyzed against 20 L of 0.02 M Tris-HCl buffer, pH 8.0, overnight. The dialysate was applied to a column (5 × 10 cm) of Sulfate Cellulofine, which had been equilibrated with 0.02 M Tris-HCl buffer, pH 8.0. After the column was washed, protein was eluted by a linear salt gradient formed by each 1 L of the equilibration buffer and the buffer containing 0.5 M NaCl. Each 10-mL fraction was collected. The fractions containing β2 glycoprotein I (150–175) were pooled and concentrated by ultrafiltration using PM 10 membrane (Amicon Co.). The concentrate was applied to a column (4.6 × 250 cm) of Sephadex G-150, which had been equilibrated with 0.02 M Tris-HCl buffer, pH 8.0, containing 0.15 M NaCl. Each 5-mL fraction was collected. β2 glycoprotein I thus isolated (15.2 mg) gave a single band with a molecular weight of 50 000 in the presence or absence of 2-mercaptoethanol on SDS–PAGE. The absorbance of the protein at 280 nm was found to be 0.97 on the solution of 1 mg/mL. The protein concentration was determined from the dried weight of the protein.

Chemical and Enzymatic Digestions. Pyridylethylation of protein was performed as described by Hermodson et al. (1973). Cyanogen bromide treatment was performed in 70%

† This work was supported in part by a Grant-in Aid for Scientific Research from the Ministry of Education, Science, and Culture of Japan. A preliminary report was presented at the XIIIth Congress of the International Society on Thrombosis and Haemostasis held in Amsterdam, The Netherlands, June 30–July 5, 1991 (Kato & Enjyoji, 1991).

* Address correspondence to this author.

0006-2960/91/0430-11687$02.50/0

Table I: Amino Acid Compositions (in Residues per Molecule) of Pyridylethylated β2 Glycoprotein I (PE-β2GP-I) and Cyanogen Bromide Fragments[a]

amino acid	PE-β2GP-I	PE-CN IV	PE-CN II	PE-CN I	PE-CN III
Asp	15.5 (24)	2.0 (2)	5.6 (5)	11.0 (12)	5.2 (5)
Glu	24.1 (27)	4.6 (5)	5.1 (4)	12.0 (12)	6.4 (6)
Ser	20.0 (24)	2.5 (3)	10.0 (9)	7.5 (7)	4.6 (5)
Gly	22.5 (23)	5.0 (6)	7.9 (7)	7.3 (8)	2.0 (2)
His	7.4 (8)		3.8 (3)	3.4 (3)	2.0 (2)
homoSer	3.0[b] (3)	nq[c] (1)	nq (1)	nq (1)	
Arg	13.1 (12)	5.0 (5)	2.4 (2)	4.8 (5)	
Thr	18.4 (24)	5.0 (6)	6.6 (7)	6.7 (8)	3.0 (3)
Ala	15.4 (17)		7.9 (8)	5.5 (6)	3.0 (3)
Pro	24.4 (30)	5.9 (8)	11.6 (13)	6.3 (8)	1.9 (2)
Tyr	12.5 (12)	1.4 (2)	4.0 (5)	2.9 (4)	0.9 (1)
Val	15.9 (18)	3.1 (4)	5.2 (5)	6.1 (7)	2.0 (2)
PECys	17.0 (22)	4.9 (4)	5.0 (6)	5.1 (7)	4.4 (5)
Ile	12.5 (13)	2.5 (3)	4.1 (4)	3.0 (3)	2.9 (3)
Leu	16.0 (17)	4.3 (5)	6.0 (6)	3.5 (4)	2.0 (2)
Phe	18.2 (20)	2.7 (3)	6.9 (7)	5.1 (6)	3.9 (4)
Lys	24.6 (26)	2.5 (3)	7.2 (7)	6.4 (7)	8.8 (9)
Trp	nd[d] (5)	nd (1)	nd (1)	nd (2)	nd(1)
total	(326)	(61)	(100)	(110)	(55)
position	1–326	1–61	62–161	162–271	272–326
yield(%)		23	27	27	9

[a] Values in parentheses are calculated from the amino acid sequence. [b] Determined as methionine. [c] Not quantitatively determined. [d] Not determined.

formic acid at room temperature for two days. Tryptic, chymotryptic, or V8 protease digestions of cyanogen bromide fragments were performed in 0.1 M ammonium bicarbonate, pH 8.0, at 37 °C for 6 h. For the determination of the position of the disulfide bonds, β2 glycoprotein I was incubated with trypsin in 0.02 M acetate buffer, pH 6.3, V8 protease in 0.1 M ammonium acetate, pH 4.0, and thermolysin in 0.1 M ammonium acetate, pH 6.4, containing 1 mM $CaCl_2$, at 37 °C for 12 h. Digestions of PE[1]-β2 glycoprotein I by lysyl endopeptidase or carboxypeptidase Y were performed in 0.01 M Tris-HCl, pH 8.0, or in 0.1 M pyridineacetic acid, pH 5.5, respectively, at 37 °C. A weight ratio of the enyzmes to the peptides was 1:50.

Purification of Peptides. Isolation of peptides was performed by gel filtration on a column (1.5 × 144 cm) of Sephadex G-100, equilibrated with 10% acetic acid, or by reversed-phase HPLC using a column (4.6 × 250 mm) of Cosmosil 5C18-300 with an acetonitrile gradient in 0.05% trifluoroacetic acid using the Toso HPLC system. Detection of cystine-containing peptides was performed using SBD-F as described by Sueyosi et al. (1987). The yield of peptides isolated was calculated from the amounts of peptide determined by amino acid analysis.

Analytical Methods. Amino acid composition and amino acid sequence were analyzed by Picotag system and a gas-phase sequencer, respectively, as described elsewhere (Maeda et al., 1988).

Nomenclature of Peptides. The pooled fractions and peptides were identified by letters to indicate the type of cleavage performed. Numbers were employed to show the fractions in gel filtration or peptide positions relative to the other peaks in the chromatograms of HPLC. On the second HPLC, letters were employed to identify each peak. The nomenclature used was as follows: CN, cyanogen bromide cleavage; T, trypsin digestion; C, chymotrypsin digestion; E, V8 protease digestion; K, lysyl endopeptidase digestion; Y, carboxypeptidase digestion.

[1] Abbreviations: PE, pyridylethylated; PECys, (pyridylethyl)cysteine; SDS–PAGE, polyacrylamide gel electrophoresis in the presence of sodium dodecyl sulfate.

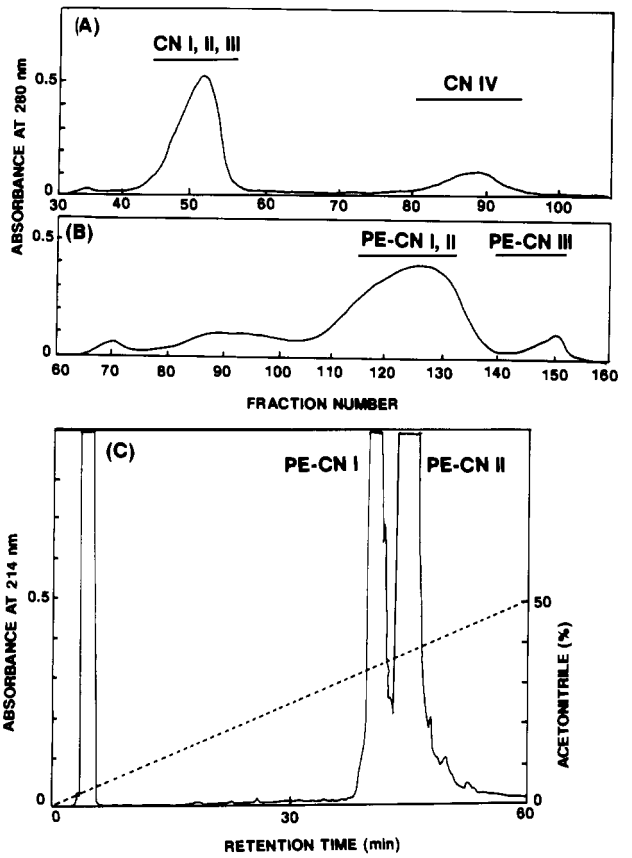

FIGURE 1: Separation of cyanogen fragments of β2 glycoprotein I. (A) Gel filtration of cyanogen fragments of β2 glycoprotein I on a column of Sephadex G-100. Bovine β2 glycoprotein I (220 nmol) was treated with cyanogen bromide, and the digest was applied to a column (1.5 × 144 cm) of Sephadex G-100, equilibrated with 10% acetic acid. Each 1-mL fraction was collected. Fractions were pooled as indicated. (B) Gel filtration of a mixture of pyridylethylated CN I, CN II, and CN III on a column of Sephadex G-100. A mixture of CN I, II, and III from (A) was pyridylethylated and separated by gel filtration under the same conditions as described in (A). Fractions were pooled as indicated. (C) HPLC of a mixture of PE-CN I and PE-CN II. A mixture of PE-CN I and PE-CN II was separated by reverse-phase HPLC as described under Experimental Procedures.

Results and Discussion

Amino Acid Sequence of Bovine β2 Glycoprotein I. The amino acid composition of PE-β2 glycoprotein I is shown in Table I. The amino-terminal amino acid sequence was identified as follows: Gly-Arg-X-Cys-Pro-Lys-Pro-Asp-Glu-Leu-Pro-Phe-Ser-X-Val-Val-Pro-Leu-Lys. Bovine β2 glycoprotein I was treated with cyanogen bromide, and one of the four fragments (CN IV) was isolated by gel filtration on a column of Sephadex G-100 (Figure 1A). Mixtures of the other three fragments were pyridylethylated and subjected to a gel filtration on a column of Sephadex G-100 (Figure 1B). Since PE-CN I and PE-CN II were not completely separated by the gel filtration, they were isolated by reversed-phase HPLC (Figure 1C). The amino acid compositions of these four PE fragments (PE-CN I, II, III, and IV) and that of PE β2 glycoprotein I are shown in Table I. From the amino-terminal amino acid sequence of these fragments and PE-β2 glycoprotein I, PE-CN IV was elucidated to be the amino-terminal part of β2 glycoprotein I. The amino acid sequences of PE-CN IV and PE-CN II were determined from the tryptic and chymotryptic fragments. The amino acid compositions of chymotryptic peptides of PE-CN IV and of tryptic peptides

Table II: Amino Acid Compositions (in Residues per Molecule) of Chymotryptic Peptides Derived from PE-CN IV[a]

amino acid	C-16	C-15	C-17	C-7	C-12	C-9	C-24	C-3
Asp	0.9 (1)						1.2 (1)	
Glu	1.7 (1)		3.0 (3)	1.3 (1)				
Ser		1.2 (1)		1.0 (1)	0.9 (1)			
Gly	1.1 (1)		1.0 (1)	1.2 (1)	1.9 (2)		1.0 (1)	
Arg	1.2 (1)		1.0 (1)		3.3 (3)			
Thr	1.0 (1)	1.0 (1)	0.9 (1)			1.1 (1)	1.8 (2)	
Pro	2.7 (3)	1.0 (1)	1.2 (1)	0.9 (1)		1.2 (1)	0.9 (1)	
Tyr			0.7 (1)	0.7 (1)				
Val		1.7 (2)	0.5 (1)		0.9 (1)			
PECys	0.7 (1)			0.7 (1)		0.9 (1)		1.0 (1)
Ile			0.5 (1)		0.8 (1)		1.0 (1)	
Leu	0.8 (1)	1.2 (1)				1.2 (1)	1.9 (2)	
Phe	1.1 (1)		0.9 (1)		0.8 (1)			
Lys	1.0 (1)		0.8 (1)					1.0 (1)
Trp							nd[b] (1)	
homoSer								1.0 (1)
total	12	6	12	6	9	4	9	3
position	1–12	13–18	19–30	31–36	37–45	46–49	50–58	59–61
yield(%)	36	24	36	32	40	20	30	8

[a] Values in parentheses are calculated from the amino acid sequence. [b] Not determined.

Table III: Amino Acid Compositions (in Residues per Molecule) of Tryptic Peptides of PE-CN II[a]

amino acid	T-18a	T-24	T-4	T-20	T-22a	T-18b	T-9	T-6	T-8
Asp	0.3 (1)	0.7 (1)		1.2 (1)	0.3 (1)	0.8 (2)			
Glu	0.6 (1)	1.0 (1)	1.6 (2)	1.9 (2)					
Ser		1.7 (2)		1.5 (1)	0.7 (1)	4.0 (4)			
Gly	2.1 (2)	1.0 (1)	1.3 (1)	1.4 (1)		2.2 (2)			
His		0.6 (1)						2.1 (2)	2.2 (2)
Arg	0.9 (1)								
Thr	0.8 (1)	2.5 (4)	1.1 (1)	1.9 (2)	0.8 (1)				
Ala	1.0 (1)			1.4 (1)	0.9 (1)	1.9 (2)	1.0 (1)	0.8 (1)	0.9 (1)
Pro	1.0 (1)	1.7 (1)		7.5 (8)	7.1 (8)	1.0 (1)		0.8 (1)	0.8 (1)
Tyr		1.5 (3)				1.8 (2)			
Val	2.1 (2)			1.2 (1)	1.0 (1)	1.2 (1)	1.0 (1)		
PECys	0.9 (1)	1.1 (1)	1.2 (1)	2.5 (3)	1.8 (2)			0.7 (1)	0.8 (1)
Ile	0.9 (1)	0.8 (1)		1.9 (2)	2.0 (2)				
Leu	0.9 (1)	0.8 (1)		1.3 (1)	0.9 (1)	2.0 (2)		0.8 (1)	0.8 (1)
Phe	1.0 (1)	1.9 (3)				2.0 (2)	1.0 (1)		
Lys		1.0 (1)	1.0 (1)	2.1 (2)	1.0 (1)	1.7 (2)	1.0 (1)		
homoSer								nd[b] (1)	nd[b] (1)
Trp				nd[b] (1)	nd[b] (1)				
total	(14)	(21)	(6)	(26)	(20)	(20)	(4)	(7)	(7)
position	64–77	78–98	105–110	105–130	111–130	131–150	151–154	155–161	155–161
yield(%)	10	10	11	25	10	14	75	25	50

[a] Values in parentheses are calculated from the amino acid sequence. [b] Not determined.

of PE-CN II are shown in Tables II and III, respectively. The amino acid sequence of PE-CN I was determined from the tryptic fragments and the fragments derived from the V8 protease digest. The amino acid sequence of PE-CN III was determined from chymotryptic fragments and the fragments derived from the V8 protease digests. The amino acid compositions of tryptic peptides of PE-CN I and chymotryptic peptides of PE-CN III are shown in Tables IV and V, respectively. The summary of the amino acid sequences of these fragments derived from four cyanogen bromide fragments is shown in Figure 2. Overlapping peptides between cyanogen bromide fragments (T(68–81)23a, E38–K16, T(68–81)4) were isolated from the digests of PE-β2 glycoprotein I with trypsin, V8 protease, and lysyl endopeptidase as shown in Figure 2. The carboxy-terminal amino acid sequence was determined by the treatment of PE-β2 glycoprotein I with carboxypeptidase Y, as shown in Table VI.

The homology of the amino acid sequence of bovine β2 glycoprotein I with that of the human protein (Lozier et al., 1984) and that deduced from rat cDNA (Aoyama et al., 1989) is shown in Figure 3. The amino-terminal portion of rat β2 glycoprotein I was not included for comparison because it was quite different from those of human and bovine proteins.[2] Among 326 amino acid residues of bovine protein, 83% of the amino acid residues are identical with those of the human protein. Among 22 cysteine residues in the bovine protein, one cysteine residue is not consistent with those of the human protein; that is, Cys 169 in bovine protein was assigned to be Asn. Furthermore, Ser 102 in the bovine protein was assigned to be Cys in the human protein. These residues are indicated by arrows in Figure 3. However, if we compare the amino acid sequence of bovine protein with that of rat protein which was elucidated from cDNA study, all the cysteine residues are homologous between two proteins. The positions of the cysteine residues in human protein should be reexamined because it is reasonable to speculate that all the cysteine residues between three species should be homologous.[3]

[2] The amino acid sequence of the amino-terminal portion of the purified rat β2 glycoprotein I was homologous to those of human and bovine proteins (H. Kato, unpublished experiment). Therefore, cDNA data in the amino-terminal portion will be in error.

[3] After the submission of this paper, the cDNA sequence for human β2 glycoprotein I was reported (Steinkasserer et al., 1991). On the deduced amino acid sequence, the amino acid residues 102 and 169 were assigned to be serine and cysteine, respectively.

Table IV: Amino Acid Compositions (in Residues per Molecule) of Tryptic Peptides of PE-CN I[a]

amino acid	T-18	T-2	T-26	T-20	T-30	T-16	T-3	T-13	T-15	T-9	T-5
Asp	2.1 (3)		4.0 (4)	2.0 (2)	5.4 (6)	0.6 (1)				0.7 (1)	
Glu	2.6 (3)	0.9 (1)		3.9 (4)	4.5 (4)	0.9 (1)			1.7 (2)	1.0 (1)	
Ser			1.2 (1)	1.9 (2)	3.2 (3)	2.2 (2)	1.0 (1)	1.0 (1)			
Gly	2.3 (2)		1.3 (1)	2.0 (2)	3.0 (3)	0.9 (1)			1.3 (1)		
His	1.0 (1)		1.0 (1)	1.0 (1)	2.5 (2)						
Arg	1.2 (1)	1.1 (1)	1.2 (1)		1.4 (1)			1.2 (1)	1.4 (1)		
Thr	3.7 (4)			2.8 (3)	2.4 (3)				1.0 (1)		
Ala			1.0 (1)	0.9 (1)	2.3 (2)	0.9 (1)	1.2 (1)		1.2 (1)	1.2 (1)	
Pro	1.0 (1)		4.0 (5)	0.9 (1)	4.4 (6)	0.9 (1)					
Tyr			1.6 (2)	0.8 (1)	2.3 (3)				0.9 (1)		
Val	0.9 (1)	1.0 (1)	1.9 (2)	1.1 (1)	2.7 (3)				1.0 (1)	1.4 (1)	
PECys	1.8 (2)		0.9 (1)	1.9 (2)	2.6 (3)	0.7 (1)	1.1 (1)				
Ile								1.1 (1)	1.0 (1)	1.3 (1)	
Leu	0.8 (1)		0.9 (1)	0.9 (1)	1.7 (2)			1.1 (1)			
Phe	0.9 (1)		1.8 (2)	1.2 (1)	2.6 (3)	0.8 (1)					1.0 (1)
Lys			1.1 (1)	1.1 (1)	2.4 (2)	0.6 (1)	1.0 (1)	0.8 (1)		0.9 (1)	0.9 (1)
homoSer											
Trp	nd[b] (1)					nd[b] (1)					
total	(21)	(3)	(23)	(23)	(46)	(11)	(4)	(5)	(9)	(6)	(2)
position	162–182	183–185	186–208	209–231	186–231	232–242	243–246	247–251	252–260	261–266	267–268
yield (%)	9	52	7	8	57	9	10	10	20	13	26

[a] Values in parentheses are calculated from the amino acid sequence. [b] Not determined.

FIGURE 2: Summary of the amino acid sequence of bovine β2 glycoprotein I. Amino acid sequences determined by Edman degradation of the peptides are given below the summarized amino acid sequence. Those not identified are indicated by dashes. The asterisk indicates the carbohydrate attachment site.

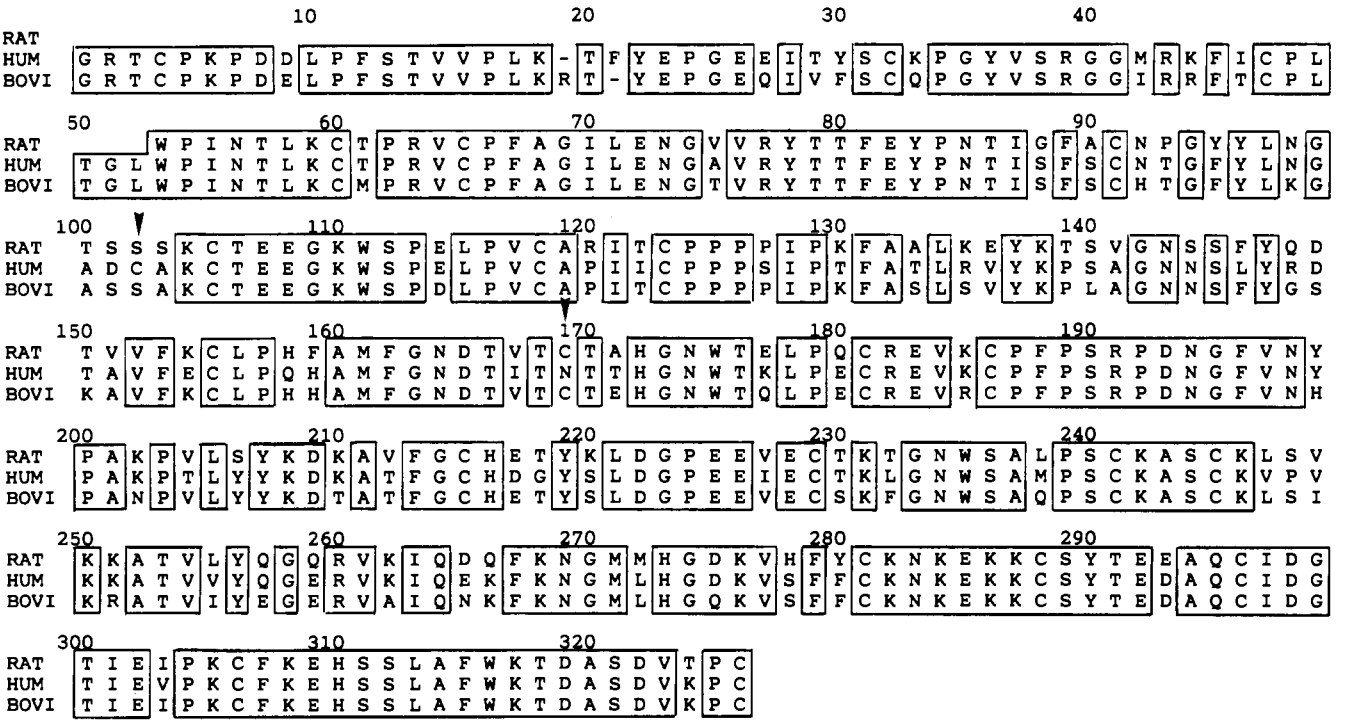

FIGURE 3: Homology of the amino acid sequence of bovine β2 glycoprotein I (BOVI) with those of human (HUM) (Lozier et al., 1984) and rat (RAT) (Aoyama et al., 1989) β2 glycoprotein I. The amino-terminal portion of rat β2 glycoprotein I was excluded because it was completely different from those of human and bovine β2 glycoprotein I (see text). The identical amino acid residues betwen three kinds of proteins are bracketed. The arrowhead indicates cysteine residues of human protein that are not consistent with those of bovine and rat proteins.

Table V: Amino Acid Compositions (in Residues per Molecule) of Chymotryptic Peptides from PE-CN III[a]

amino acid	C-17b	C-16	C-4	C-17a	C-5
Asp		2.6 (3)			2.0 (2)
Glu	1.2 (1)	4.3 (4)	1.1 (1)		
Ser	1.2 (1)	1.0 (1)	2.0 (2)		1.1 (1)
Gly	1.2 (1)	1.1 (1)			
His	1.0 (1)		1.1 (1)		
Thr		1.7 (2)			1.0 (1)
Ala		1.1 (1)		1.1 (1)	1.0 (1)
Pro		1.0 (1)			0.9 (1)
Tyr		1.0 (1)			
Val	1.0 (1)				0.9 (1)
PECys		3.5 (4)			1.0 (1)
Ile		2.3 (3)			
Leu	1.0 (1)		1.0 (1)		
Phe	2.0 (2)	1.0 (1)		0.6 (1)	
Lys	0.8 (1)	4.1 (5)	1.0 (1)		2.0 (2)
Trp					
total	(9)	(27)	(6)	(2)	(10)
position	272–280	281–307	308–313	314–315	317–326
yield (%)	13	85	50	25	60

[a] Values in parentheses are calculated from the amino acid sequence.

Table VI: Amino Acids Liberated from Pyridylethylated β2 Glycoprotein I by Carboxypeptidase Y[a]

amino acid	5 min	10 min	30 min	60 min	120 min
PECys	11.3	14.8	27	25	35
Pro	2.9	3.6	8	9.9	17.6
Lys	1.0	3.5	10	12	25.9
Val	2.6	3.7	8	9.9	19
Asp	2.6	6.5	9	11.5	27.7
Ala	0	3.3	8	10.5	53.8

[a] Each value represents amounts of amino acids (in picomoles) liberated from 60 pmol of PE-β2 glycoprotein I, using a Pico-Tag amino acid analyzer.

The positions of asparagine-linked carbohydrate in bovine β2 glycoprotein I are shown in Figure 2 by an asterisk (Asn 73, 143, 164, 174, and 234). The glycosylation of these asparagine residues was confirmed by the absence of PTH-amino acids on amino acid sequence analyses and by the presence of an extra aspartic acid on amino acid analyses for each peptide: PE-CNII-C-13 for Asn 73, PE-CNII-C-10b for Asn 143, PE-CNI-E-11 for Asn 164, PE-CNI-E-19 for Asn 174, and PE-CNI-T-16 for Asn 234. These asparagine residues are consistent with the probable site of carbohydrate attachment, Asn-X-Ser, or Thr.

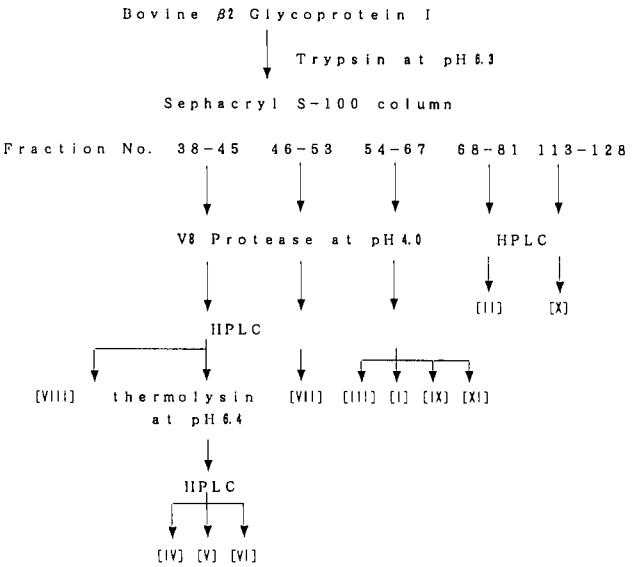

FIGURE 4: Summary of the purification procedures of cystine-containing peptides from bovine β2 glycoprotein I. Cystine-containing peptides were located by SBD-F reagent and by amino acid analysis. They were named by Roman numerals starting from the disulfide bond at the amino-terminal portion.

Location of Disulfide Bonds in Bovine β2 Glycoprotein I. To elucidate the location of disulfide bonds, we used a strategy to isolate the peptides containing disulfide bonds, as shown in Figure 4. Bovine β2 glycoprotein I was treated with trypsin at pH 6.3, and the digest was subjected to a gel filtration on a column of Sephacryl S-100. Peptides containing disulfide

Table VII: Amino Acid Composition (in Residues per Molecule) of Cystine-Containing Peptides from Bovine β2 Glycoprotein I[a]

amino acid	I	II	III	IV	V	VI	VII	VIII	IX	X	XI
Asp	1.2 (2)						3.4 (5)	0.9 (2)	0.9 (2)		0.9 (2)
Glu	0.8 (1)	3.5 (4)	2.2 (3)		0.5 (1)	1.5 (2)	2.7 (3)	1.3 (2)	1.6 (2)		0.7 (1)
Ser	0.9 (1)	1.9 (2)		1.0 (1)	0.7		2.8 (3)	2.3 (2)	1.0 (1)	0.7 (1)	1.9 (2)
Gly	1.1 (1)	2.1 (2)	2.1 (2)		1.1 (1)	1.9 (2)	2.2 (2)	2.0 (2)	1.2 (1)		
His					1.2 (1)	1.0 (1)	1.0 (1)	0.7 (1)			
Arg		1.6 (2)					1.2 (1)				
Thr	4.2 (5)	1.0 (1)	0.9 (1)		2.9 (3)		1.1 (1)	1.7 (2)	1.0 (1)		2.0 (2)
Ala			1.1 (1)	1.0 (1)			1.1 (1)	1.8 (2)	2.0 (2)		1.1 (1)
Pro	5.7 (6)	3.1 (3)	1.0 (1)	1.7 (1)	5.1 (5)	1.8 (2)	5.7 (6)	1.1 (1)	1.0 (1)		1.1 (1)
Tyr		1.7 (2)					2.7 (3)				1.0 (1)
Val	1.6 (2)	1.8 (2)	1.0 (1)	1.0 (1)	1.2 (1)		3.2 (3)			1.0 (1)	1.2 (1)
Met		0.5 (1)									
1/2Cys	0.9 (2)	0.3 (2)	0.8 (2)	1.2 (2)	0.8 (2)	1.1 (2)	0.7 (2)	0.8 (2)	1.4 (2)	1.0 (2)	1.7 (2)
Ile	1.0 (1)	1.0 (1)	1.0 (1)		2.0 (2)				3.0 (3)		
Leu	4.2 (5)		1.0 (1)			2.2 (2)	2.0 (2)				
Phe	1.9 (2)	1.0 (1)	1.0 (1)	1.0 (1)			2.0 (2)	1.7 (2)		2.7 (3)	
Lys	2.3 (3)		0.8 (1)		0.8 (1)		0.9 (1)	0.7 (1)	1.8 (2)		1.1 (1)
Trp	nd[b] (1)							nd[b] (1)			
total	(32)	(23)	(15)	(7)	(17)	(11)	(36)	(20)	(17)	(7)	(14)
position	3–19	21–39	64–72	89–91	121–130	155–159	186–207	209–217	243–246	277–281	287–292
	45–59	60–63	105–110	117–120	167–173	177–182	218–231	232–242	293–305	306–307	318–326
recovery (%)	5	3	19	9	9	12	11	5	47	43	49

[a] Values in parentheses are calculated from the amino acid sequence. [b] Not determined.

Peptide	Position	Amino Acid Sequence	Positions of Disulfide Bonds
I	3-19	Thr Cys Pro Lys Pro Asp Glu Leu Pro (Phe, Ser, Thr, Val, Val, Pro, Leu, Lys)	
		91 62 99 106 29 34 23	
	45-59	Phe Thr Cys Pro Leu Thr Gly (Leu, Trp, Pro, Ile, Asn, Thr, Leu, Lys)	4-47
		284 113 57 25	
II	21-39	Thr Tyr Glu Pro Gly Glu Gln Ile Val Phe Ser Cys Gln Pro Gly Tyr Val (Ser, Arg)	
		374 370 397 262 226 166 122 116 92 47 46 20 13 4	
	60-63	Cys Met Pro Arg	32-60
		314 229 122	
III	64-72	Val Cys Pro Phe Ala Gly Ile Leu Glu	
		600 417 350 269 329 117 81 47	
	105-110	Cys Thr Glu Glu Gly (Lys)	65-105
		595 814 393 238	
IV	89-91	Phe Ser Cys	
		660	
	117-120	Val Cys Ala Pro	91-118
		622 145 91	
V	121-130	Ile Thr Cys Pro Pro Pro Pro Ile (Pro, Lys)	
		337 182 140 46 30 12	
	167-173	Val Thr Cys Thr Glu (His, Gly)	123-169
		352 313	
VI	155-159	Cys Leu Pro His His	
		179 110	
	177-182	Gln Leu Pro Glu Cys Arg	155-181
		148 483 299 102 79	
VII	186-207	Cys Pro Phe Pro Ser Arg Pro Asp Asn Gly Phe Val Asn His Pro Ala Asn Pro (Val, Leu, Tyr, Tyr)	
		508 578 420 327 533 598 340 260 268 233 294 351 170 172 240 100	
	218-231	Thr Tyr Ser Leu Asp Gly Pro Glu Glu Val Glu Cys Ser Lys	186-229
		556 354 614 259 265 370 210 289 56	
VIII	209-217	Asp Thr Ala Thr Phe Gly Cys His (Glu)	
		278 150 104 112 75	
	232-242	Phe Gly Asn Trp Ser Ala Gln Pro (Ser, Cys, Lys)	215-241
		198 180 104 78 51	
IX	243-246	Ala Ser Cys Lys	
		2070 1126	
	293-305	Asp Ala Gln Cys Ile Asp Gly Thr Ile Glu (Ile, Pro, Lys)	245-296
		3345 2235 1952 619 1275 510 176 179	
X	277-281	Val Ser Phe Phe (Cys)	
		539 258 106	
	306-307	Cys Phe	281-306
		322	
XI	287-292	Lys Cys Ser Tyr Thr Glu	
		252 704 629	
	318-326	Thr Asp Ala Ser Asp Val Lys Pro (Cys)	288-326
		2601 1388 2452 1196 910 539	

FIGURE 5: Amino acid sequence of 11 cystine-containing peptides from bovine β2 glycoprotein I. Each value under the amino acid residue identified (shown by arrows) shows the amounts of PTH-amino acid (in picomoles) recovered using a gas-phase sequencer. Serine and threonine were not quantitatively determined.

bonds, [II] and [X], were isolated from two fractions (fractions 68–81 and 113–128) by HPLC, respectively. The other three fractions were treated with V8 protease at pH 4.0, and the digests were subjected to HPLC. From fractions 46–53, a peptide containing disulfide bonds [VII] was isolated. From fraction 54–67, four peptides containing disulfide bonds, [III], [I], [IX], and [XI], were isolated, respectively. From fractions 38–45, a peptide containing disulfide bonds [VIII], was isolated. Three peptides containing disulfide bonds, [IV], [V], and [VI], were isolated from the digests of the fraction with thermolysis at pH 6.4. The amino acid compositions and amino acid sequences of these 11 peptides are shown in Table

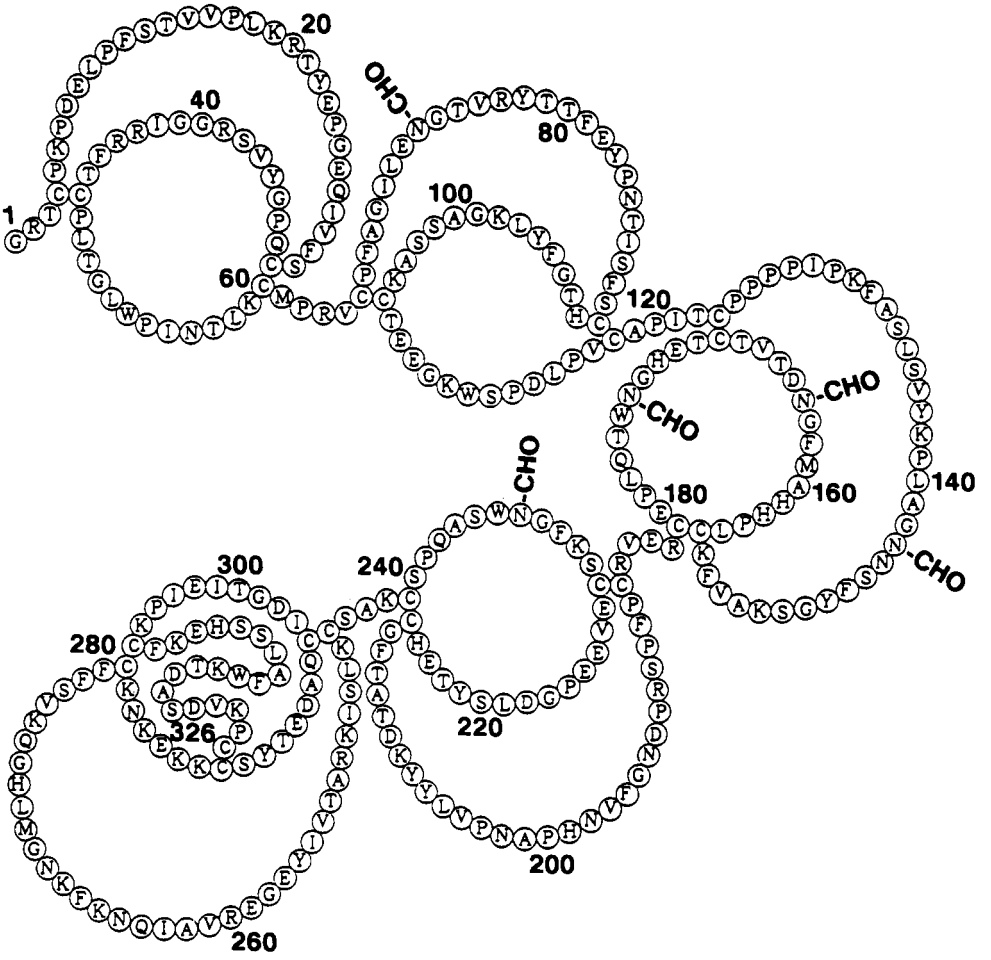

FIGURE 6: Amino acid sequence and location of the disulfide bonds of bovine β2 glycoprotein I. -CHO indicates a carbohydrate chain.

VII and Figure 5. The location of the disulfide bonds in bovine β2 glycoprotein I is shown in Figure 6. The repeated amino acid sequence in human β2 glycoprotein I has been shown, and the location of six disulfide bonds in human protein has been described (Lozier et al., 1984). However, one of six disulfide bonds is not consistent with those of bovine protein. Results indicate that bovine β2 glycoprotein I consists of four characteristic domains and one modified form of the domain. These characteristic domains were initially called a GP-I structure (Davie et. al., 1986) because they were first identified in β2 glycoprotein I. The domain structure is now called a Sushi domain (Ichinose et al., 1990). Therefore, the presence of five Sushi domains in β2 glycoprotein I was demonstrated by this study. The Sushi domains are found in various proteins from mammalian origins and from horseshoe crab (Muta et al., 1991). The function of the domains is not known. In β2 glycoprotein I, it can be speculated that the basic amino acid residues distributed in each domain (38 residues of arginine and lysine) play the important role for binding with negatively charged substances. The distribution of the basic amino acids in each domain is uneven: 8 in domain 1; 5 in domains 2, 3, and 4; and 15 in domain 5. It is not clear yet whether all these domains have equal ability to bind with negatively charged substances or whether they have the ability in a concerted way. Although we could not isolate each of five domains respectively, only domain 1 was isolated by cyanogen bromide treatment of β2 glycoprotein I. The domain 1 fragment had the same ability to inhibit the surface-mediated activation of factor XII and prekallikrein as the native protein did (H. Kato, unpublished data). The inhibitory ability of domain 1 may be due to the basic amino acids arranged by the specific conformation which was fixed by two disulfide bonds. It should be pointed out that domain 3 has the characteristic structure among five domains, which contains 3 carbohydrate chains and four consecutive proline residues. These carbohydrate chains and proline residues will give the specific conformation to domain 3.

ACKNOWLEDGMENTS

We are grateful to Dr. Toshiyuki Miyata for helpful discussions in the preparation of the manuscript.

REFERENCES

Aoyama, Y., Chan, Y.-L., & Wool, I. G. (1989) *Nucleic Acids Res. 17*, 6401.

Davie, E. W., Ichinose, A., & Leytus, S. (1986) *Cold Spring Harbor Symp. Quant. Biol. 51*, 509–514.

Galli, M., Comfurius, P., Maassen, C., Hemker, H. C., Debaets, M. H., Van Breda-Vriesman, P. J. C., Barbui, T., Zwaal, R. F. A., & Bevers, E. M. (1990) *Lancet* (June 30), 1544.

Hermodson, M. A., Ericsson, L. H., Neurath, H., & Walsh, K. A. (1973) *Biochemistry 12*, 3146–3153.

Ichinose, A., Bottenus, R. E., & Davie, E. W. (1990) *J. Biol. Chem. 265*, 13411–13414.

Kato, H., & Enjyoji, K. (1989) *Thromb. Haemostasis 62*, 16.

Kato, H., & Enjyoji, K. (1991) *Thromb. Haemostasis 65*, 1246.

Kodama, K., Kato, H., & Iwanaga, S. (1985) *J. Biochem. 97*, 139–151.

Lozier, J., Takahashi, N., & Putnam, F. W. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 3640–3644.

Maeda, H., Matsumura, Y., & Kato, H. (1988) *J. Biol. Chem. 263*, 16051–16054.
Matsuura, E., Igarashi, Y., Fujimoto, M., Ichikawa, K., & Koike, T. (1990) *Lancet* (July 21), 177–178.
McNeil, H. P., Simpson, R. J., Chesterman, C. N., & Krilis, S. A. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 4120–4124.
Muta, T., Miyata, T., Misumi, Y., Tokunaga, F., Nakamura, T., Toh, Y., Ikehara, Y., & Iwanaga, S. (1991) *J. Biol. Chem. 266*, 6554–6561.
Polz, E., & Kostner, G. M. (1979) *FEBS Lett. 102*, 183–186.
Schousboe, I. (1988) *Eur. J. Biochem. 176*, 629–636.
Schousboe, I., & Rasmussen, M. S. (1988) *Int. J. Biochem. 20*, 787–792.
Steinkasserer, A., Estaller, C., Weiss, H., Sim, R. B., & Day, A. J. (1991) *Biochem. J. 277*, 387–391.
Sueyoshi, T., Miyata, T., Hashimoto, N., Kato, H., Hayashida, H., Miyata, T., & Iwanaga, S. (1987) *J. Biol. Chem. 262*, 2768–2779.

Reconstitution of Catalytically Competent Human ζ-Thrombin by Combination of ζ-Thrombin Residues A1–36 and B1–148 and an *Escherichia coli* Expressed Polypeptide Corresponding to ζ-Thrombin Residues B149–259

Zhong-Ru Gan, Sidney D. Lewis, James R. Stone,[‡] and Jules A. Shafer*
Department of Biological Chemistry, Merck Sharp & Dohme Research Laboratories, West Point, Pennsylvania 19486
Received June 11, 1991; Revised Manuscript Received September 5, 1991

ABSTRACT: Human ζ-thrombin, a catalytically competent serine proteinase, arises from a single chymotryptic cleavage at Trp-148 in α-thrombin to generate two nonconvalently associated polypeptide segments designated ζ1-thrombin (the 36-residue A-chain disulfide linked to B-chain residues B1–148) and ζ2-thrombin (B149–259). We report here the expression of recombinant ζ2-thrombin in *Escherichia coli* and the reconstitution of catalytically competent ζ-thrombin by combination of ζ1-thrombin with recombinant ζ2-thrombin. A DNA fragment encoding ζ2-thrombin was cloned into a pATH2 expression vector as a *trpE*–ζ2 fusion gene, in which a factor Xa cleavage site was inserted between the *trpE* and the ζ2-thrombin gene. High-level expression of this fusion protein was achieved under the control of the *E. coli trp* promoter. The expressed ζ2-thrombin was liberated from the fusion protein by factor Xa cleavage, reduced with DTT, and purified to homogeneity by reverse-phase HPLC. Oxidation of the reduced ζ2-thrombin in the presence of 80 μM $CuSO_4$ and 6 M urea at pH 8.15 yielded material that was indistinguishable on HPLC from ζ2-thrombin isolated by resolution of human ζ-thrombin. Catalytically active ζ-thrombin was generated by combination of recombinant ζ2-thrombin with ζ1-thrombin that was isolated by resolution of human ζ-thrombin. Recombinant ζ-thrombin displayed catalytic activities, toward a small chromogenic substrate and fibrinogen, that were similar to those of α-thrombin prepared from human blood plasma and ζ-thrombin obtained by treatment of α-thrombin with chymotrypsin. This result indicates that the information for formation of a catalytically competent conformation resides in the primary structure of ζ-thrombin and suggests that studies of variants of ζ-thrombin produced by site-directed mutagenesis of ζ2-thrombin could facilitate identification of the structural and functional determinants of the interactions of thrombin that are important in blood coagulation.

α-Thrombin is a serine protease that plays a central role in hemostasis. It converts (via limited proteolysis) fibrinogen to fibrin monomers that polymerize spontaneously to form the insoluble fibrin matrix of blood clots (Blomback, 1978; Shafer & Higgins, 1988). α-Thrombin also catalyzes conversion of factor XIII to factor XIIIa, a transglutaminase that stabilizes fibrin clots by cross-linking fibrin to itself and other plasma proteins (Lorand & Konishi, 1964; Takagi & Doolittle, 1974; Lewis et al., 1987). Additionally, thrombin regulates the reaction cascade responsible for its generation. α-Thrombin activates factor V, factor VIII, and platelets so as to induce an explosive increase in the rate of generation of thrombin during blood coagulation (Coleman, 1969; Nesheim & Mann, 1979; Hoyer & Trabold, 1981; Mann et al., 1988; Berndt et al., 1986). When thrombin enters the microcirculation (where the concentration of the endothelial cell-surface receptor, thrombomodulin, becomes substantial), thrombin exerts a negative regulatory effect by forming a complex with thrombomodulin which in turn activates protein C (Kisiel, 1979; Esmon et al., 1986). Activated protein C together with its cofactors modulates the thrombin-generating cascade by proteolytically inactivating factor Va and factor VIIIa (Esmon, 1987). Besides its involvement in blood coagulation, thrombin displays a variety of effects in other biological systems. It functions as (i) a growth factor (Bar-shavit et al., 1986), (ii) a promoter of endothelial cell adhesion (Bar-shavit et al., 1991), and (iii) an activator of prostacyclin release from endothelial cells (Pearson et al., 1983).

Various derivatives of thrombin have been obtained from the products of either autolysis or limited proteolysis of α-thrombin, a protein comprised of a 36-residue A-chain disulfide linked to a 259-residue B-chain. Cleavage of α-thrombin at Arg-73, Ala-150, and Trp-148, yields β-, ϵ-, and ζ-thrombin,

[‡]Current address: Department of Biological Chemistry, University of Michigan, Ann Arbor, MI 48109.

0006-2960/91/0430-11694$02.50/0