Rev. Genet. 9, 305.
Hood, L., Loh, E., Hubert, J., Barstad, P., Eaton, B., Early, P., Fuhrman, J., Johnson, N., Kronenberg, M., & Schilling, J. (1976) *Cold Spring Harbor Symp. Quant. Biol. 41,* 817.
Inbar, D., Rotman, M., & Givol, D. (1971) *J. Biol. Chem. 246,* 6272.
Karush, F. (1956) *J. Am. Chem. Soc. 78,* 5519.
Krausz, L. M., Grossberg, A. L., & Pressman, D. (1976) *Immunochemistry 13,* 901.
Leon, M. A., & Young, N. M. (1971) *Biochemistry 10,* 1424.
Michaelides, M. C., & Eisen, H. N. (1974) *J. Exp. Med. 140,* 687.
Padlan, E. A., Davies, D. R., Rudikoff, S., & Potter, M. (1976) *Immunochemistry 13,* 945.
Pollet, R., & Edelhoch, H. (1973) *J. Biol. Chem. 248,* 5443.
Pople, J. A., Schneider, W. G., & Bernstein, H. J. (1959) High-Resolution Nuclear Magnetic Resonance, p 223, McGraw-Hill, New York, N.Y.
Potter, M. (1972) *Physiol. Rev. 52,* 631.
Potter, M., & Leon, M. A. (1968) *Science 162,* 369.
Potter, M., & Lieberman, R. (1970) *J. Exp. Med. 132,* 737.
Riesen, W. F., Braun, D. G., & Jaton, J. C. (1976) *Proc. Natl. Acad. Sci. U.S.A. 73,* 2096.
Rudikoff, S., & Potter, M. (1974) *Biochemistry 13,* 4033.
Rudikoff, S., Potter, M., Segal, D. M., Padlan, E. A., & Davies, D. R. (1972) *Proc. Natl. Acad. Sci. U.S.A. 69,* 3689.
Segal, D. M., Padlan, E. A., Cohen, G. H., Rudikoff, S., Potter, M., & Davies, D. R. (1974) *Proc. Natl. Acad. Sci. U.S.A. 71,* 4298.

Complete Covalent Structure of Human β-Thromboglobulin†

Geoffrey S. Begg, Duncan S. Pepper, Colin N. Chesterman, and Francis J. Morgan*

ABSTRACT: The complete primary structure of the platelet-specific protein human β-thromboglobulin has been determined. β-Thromboglobulin consists of identical subunits of 81 amino acids, each with a molecular weight of 8851. The amino acid sequence of the β-thromboglobulin subunit is: Gly-Lys-Glu-Glu-Ser-Leu-Asp-Ser-Asp-Leu-Tyr-Ala-Glu-Leu-Arg-Cys-Met-Cys-Ile-Lys-Thr-Thr-Ser-Gly-Ile-His-Pro-Lys-Asn-Ile-Gln-Ser-Leu-Glu-Val-Ile-Gly-Lys-Gly-Thr-His-Cys-Asn-Gln-Val-Glu-Val-Ile-Ala-Thr-Leu-Lys-Asp-Gly-Arg-Lys-Ile-Cys-Leu-Asp-Pro-Asp-Ala-Pro-Arg-Ile-Lys-Lys-Ile-Val-Gln-Lys-Lys-Leu-Ala-Gly-Asp-Glu-Ser-Ala-Asp. Disulfide bridges link half-cystine-16 to half-cystine-42, and half-cystine-18 to half-cystine-58. The amino acid sequence of β-thromboglobulin shows a marked homology with that of platelet factor 4. When the sequences are aligned for maximum homology, 42 of the 81 residues of β-thromboglobulin are identical with those of platelet factor 4, including the position of the four half-cystines.

The compound β-thromboglobulin (βTG)[1] is a small protein with a molecular weight of approximately 36 000 which is a significant component of the protein secretion from platelets in the release reaction induced by thrombin and other agents (Moore et al., 1975). Many of the contents of platelets are thought not to be of megakaryocytic origin, but to be absorbed from the plasma by the circulating platelet (Adelson et al., 1961). However, there is strong evidence that some clearly defined proteins including βTG (Moore et al., 1975), the anti-heparin activity known as platelet factor 4 (PF4) (Broekman et al., 1975; Niewiarowski et al., 1976), and possibly the platelet growth factor (Ross et al., 1974) are specific for platelets. Both PF4 (Walsh, 1976) and platelet growth factor (Ross & Glomset, 1976) may play a role as important mediators of the role of platelets in the thrombotic process. Radioimmunoassays which may provide methods for measurement of platelet activation both in vivo and in vitro have required detailed characterization of the platelet specific proteins. The complete amino acid sequence of human PF4 has been determined (Deuel et al., 1977; Hermodson et al., 1977; Morgan et al., 1978a). This paper describes the complete amino acid sequence and disulfide bond arrangement of βTG which exhibits a striking and unexpected homology with PF4.

Experimental Procedure

β-Thromboglobulin was prepared from fresh washed platelets as described previously (Moore & Pepper, 1977).

Trypsin treated with L-1-tosylamido-2-phenylethyl chloromethyl ketone was from Worthington, carboxypeptidase Y from Pierce Chemical Co., and staphylococcal protease from Miles Laboratories. Column chromatography, *S*-^{14}C-labeled carboxymethylation and citraconylation of proteins, enzymatic digestions, and amino acid analyses were performed according to standard procedures which have been described in detail elsewhere (Morgan et al., 1978a). Performic acid oxidation was performed according to Hirs (1967).

Estimation of sulfhydryl groups was performed with 5,5′-dithiobis(2-nitrobenzoic acid) (Ellman, 1959). For production

† From St. Vincent's School of Medical Research (G.S.B. and F.J.M.) and University of Melbourne Department of Medicine (C.N.C.), St. Vincent's Hospital, Fitzroy, Melbourne, Victoria 3065, Australia, and Edinburgh & South East Scotland Regional Blood Transfusion Service (D.S.P), Royal Infirmary, Edinburgh, United Kingdom. *Received November 22, 1977.* Supported by the National Heart Foundation and the National Health & Medical Research Council of Australia, and the Herman Trust. A preliminary account of part of this work has been given (Morgan et al., 1978b).

[1] Abbreviations used: βTG, β-thromboglobulin; PF4, platelet factor 4; Theed, *N,N,N′,N′*-tetrakis(2-hydroxyethyl)ethylenediamine; Quadrol (Wyandotte Chemicals Corp.), *N,N,N′,N′*-tetrakis(2-hydroxypropyl)-ethylenediamine; SCM, *S*-carboxymethyl; Pth, phenylthiohydantoin.

0006-2960/78/0417-1739$01.00/0

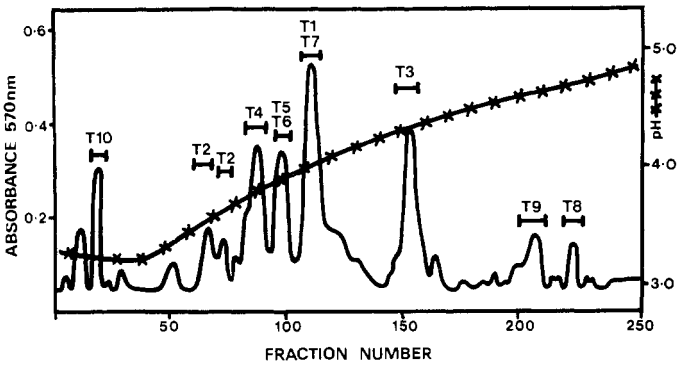

FIGURE 1: Chromatography of the tryptic peptides of *S*-carboxymethyl-β-thromboglobulin (9 mg) on Dowex 50-X2 (0.9 × 55 cm). The column was developed with a gradient from 0.2 M pyridine acetate, pH 3.1, to 2 M pyridine acetate, pH 5.0, at 50 °C, using a gradient maker with two identical chambers each containing 250 mL, at a flow rate of 24 mL per h. The fraction size was 2 mL, of which 50 μL was taken for ninhydrin analysis after alkaline hydrolysis. The peptides contained in each peak have been labeled according to their position in the sequence proposed in Figure 3; pooled regions are shown by the horizontal bars.

of disulfide-containing peptides, βTG (18 mg) was first reacted with CNBr (40 mg) in 70% formic acid (1 mL) for 24 h at 23 °C and then lyophilized. The CNBr-treated βTG was then dissolved in 2 mL of 0.05 M ammonium acetate, pH 4.0, and digested with staphylococcal protease (1/40 by weight) at 37 °C for 18 h (Houmard & Drapeau, 1972). The peptides produced were fractionated on a column of Sephadex G-25 equilibrated with 1 M acetic acid.

Amino acid sequence analysis was performed by the automated phenyl isothiocyanate degradation using a protein sequenator. The sequenator design was essentially as described earlier (Edman & Begg, 1967) except that a larger reaction cup (51 mm i.d., 60 mm high) spinning at 3000 rpm was used (the late P. Edman & G. S. Begg, unpublished results). A programming unit based on a Motorola 6800 microprocessor was fitted to the sequenator allowing the use of degradation programs of 28 stages with times of 0–9000 s variable at 1-s intervals (G. S. Begg, unpublished results); any one of three such degradation programs can be preselected for use at any cycle of a given degradation.

Tetrahydroxyethylethylenediamine (Theed) was used as a nonvolatile buffer in the coupling reaction (Begg & Morgan, 1976) in place of Quadrol. Theed was prepared from ethylenediamine and ethylene oxide. Ethylenediamine was purified by refluxing with NaOH (10% w/v) for 2 h and further refluxing with metallic Na (2% w/v) for 2 h, before distillation. Ethylene oxide was condensed from a cylinder into an ice-cooled flask containing soda lime (10% w/v), from which it was distilled using a Liebig condenser cooled to −5 °C and a receiver at ice temperature. The ethylene oxide was then stored with $KMnO_4$ (1% w/v) at 5 °C for at least 3 weeks. For the preparation of Theed, ethylenediamine (25 mL) was added to 300 mL of water kept at 4 °C with ice and the solution bubbled with N_2 while stirring for 30 min. Ethylene oxide (78 mL), freshly distilled from $KMnO_4$, was added from a dropping funnel over 3 h, and the stirring continued overnight at 5 °C. The temperature was then raised to 20 °C, and stirring continued for a further 1 h. Water and unreacted ethylene oxide were removed by rotary film evaporation (40 °C, 10 mmHg). The viscous Theed obtained was diluted to 250 mL with water and passed through a column of Amberlite IR-120 (2.5 × 25 cm) in the H^+ form. The solution which passed through the column was dried by rotary film evaporation, and the Theed was stored over P_2O_5 in a desiccator. Both starting materials and the Theed product were freed of aldehydes by the Tollens' reaction (Edman & Begg, 1967).

Heptafluorobutyric acid was prepared as previously described (Edman & Begg, 1967). The water content was determined (Meeker et al., 1962) as about 0.15%. Trifluoroacetic anhydride (Eastman) was carefully added until the water content was 0.01%. The water content was checked again after 3 days and, if found to have changed, either trifluoroacetic anhydride or water was added until the water content of 0.01% remained unchanged for 3 days.

The following reagents and solvents were used in the protein program: (solvent 1) benzene (E. Merck, Darmstadt no. 1783); (solvent 2) 15% 1-propanol (Merck no. 997) in ethyl acetate (Merck no. 9623); (solvent 3) butyl chloride; (reagent 1) 1 M Theed in 1-propanol–water (3:4) adjusted to pH 9.0 with trifluoroacetic acid; (reagent 2) 10% phenyl isothiocyanate in *n*-hexane (Merck no. 4372); (reagent 3) heptafluorobutyric acid.

The following reagents and solvents were used in the peptide degradation programs: (solvent 1) 10% 1-propanol in benzene; (solvent 2) ethyl acetate; (solvent 3) butyl chloride; (reagent 1) 0.1 M Theed in 1-propanol–water (3:4) adjusted to pH 9.0 with trifluoroacetic acid; (reagent 2 & 3) as in protein program.

Batches of the commercial solvents specified were used without further purification provided they were negative in the Tollens' reaction. All reagents and solvents for sequence determination not specified above were prepared as previously described (Edman & Henschen, 1975).

Each protein or peptide chain was sequenced at least twice. The reaction conditions for proteins were as previously described (Edman & Begg, 1967) except that where proline was known from preliminary experiments to occur in the sequence, cleavage time was extended threefold to avoid incomplete cleavage of the proline residue (Brandt et al., 1976). Typical degradation programs used are given in the supplementary material to this paper. Amino acid phenylthiohydantoins (Pth) were identified directly by thin-layer chromatography using solvent systems D and H, and Pth-Arg and Pth-His by spot tests (Edman & Henschen, 1975). The identity of Pth-Leu and Pth-Ile was confirmed by hydrolysis to the amino acid in 6 N HCl at 150 °C for 20 h.

Peptide Nomenclature. The following prefixes were used to denote the origin of the various peptides: (T) tryptic digest; (CT) tryptic digest of citraconylated SCM-βTG. Peptides are numbered in order of their position in the final sequence proposed.

Results

Tryptic Digestion of S-Carboxymethyl-β-thromboglobulin. The compositions of the peptides isolated from a tryptic digest of SCM-βTG are shown in Table I. The peptide mixture was fractionated by chromatography on Dowex 50-X2 (Figure 1). Peptides which were not separated by the ion-exchange chromatography were purified by subsequent preparative high voltage paper electrophoresis at pH 3.55 (Table I). The sum of the composition of the peptides agreed very closely with that of intact βTG (Table II) assuming a subunit molecular weight of approximately 9000. Peptide with the composition of T2 was isolated from two adjacent peaks, presumably the result of partial oxidation of methionine to the sulfoxide.

It is clear that under the conditions of the tryptic digestion used (enzyme–substrate ratio of 1/100 by weight for 90 min at 23 °C) hydrolysis between residues 55–56, 67–68, and 72–73 (Figure 3) was largely incomplete. Although peptides

TABLE I: Amino Acid Composition of Tryptic Peptides of *S*-Carboxymethyl β-Thromboglobulin.[a]

Amino acid	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	Total
Asp	1.9 (2)			1.0 (1)	1.1 (1)	1.0 (1)	2.2 (2)			2.2 (2)	9
Thr			2.0 (2)		1.5 (2)						4
Ser	1.6 (2)		1.0 (1)	1.0 (1)	0.3					1.0 (1)	5
Glu	2.5 (3)			2.4 (2)	2.1 (2)				1.0 (1)	1.2 (1)	9
Pro			0.7 (1)				2.0 (2)				3
Gly	1.1 (1)		1.2 (1)	1.2 (1)	1.1 (1)	1.2 (1)				1.2 (1)	6
Ala	0.9 (1)				0.9 (1)		1.2 (1)			2.1 (2)	5
CM-Cys		1.5 (2)			0.5 (1)		0.8 (1)				4
Val				0.8 (1)	1.5 (2)				0.8 (1)		4
Met		0.9 (1)									1
Ile		1.1 (1)	0.9 (1)	1.8 (2)	0.7 (1)		1.1 (1)	0.9 (1)	0.8 (1)		8
Leu	2.6 (3)			0.9 (1)	1.0 (1)		1.1 (1)			1.0 (1)	7
Tyr	0.7 (1)										1
Phe											0
His			0.9 (1)		0.8 (1)						2
Lys	1.0 (1)	1.0 (1)	1.0 (1)	1.0 (1)	1.1 (1)		0.9 (1)	2.0 (2)	2.0 (2)		10
Arg	0.8 (1)					0.9 (1)	1.0 (1)				3
Total	15	5	8	10	14	3	10	3	5	8	81
Yield	20%	17%	50%	13%	8%	20%	40%	14%	28%	31%	
Final purification[b]	PPE	DX	PPE	PPE	PPE	PPE	PPE	DX	DX	DX	
Mobility[c]	0.40	0.25	0.86	0.48	*d*	0.47	0.45	1.14	0.94	0.04	

[a] 24-h hydrolysis value. Peptides are numbered in their final order in the sequence. [b] Some peptides isolated from Dowex 50 chromatography (DX) were further purified by preparative high voltage paper electrophoresis at pH 3.55 (PPE). [c] Mobility is given with respect to arginine at pH 3.55. [d] Remained at origin.

TABLE II: Amino Acid Composition of Tryptic Peptides of Citraconylated *S*-Carboxymethyl-β-thromboglobulin.[a]

Amino acid	CT1	CT2	CT3	CT4	Total	βTG
Asp	2.0 (2)	3.2 (3)	2.4 (2)	1.7 (2)	9	8.96
Thr		3.6 (4)			4	3.70
Ser	1.6 (2)	1.6 (2)	0.1	0.8 (1)	5	4.70
Glu	3.0 (3)	4.2 (4)	0.1	2.1 (2)	9	9.12
Pro		0.9 (1)	2.2 (2)		3	2.87
Gly	1.0 (1)	4.0 (4)	0.1	1.0 (1)	6	6.00
Ala	1.0 (1)	1.0 (1)	1.0 (1)	2.0 (2)	5	5.18
CM-Cys		2.5 (3)	0.9 (1)		4	3.90[d]
Val		2.5 (3)		0.7 (1)	4	3.82
Met		0.7 (1)			1	0.70[e]
Ile		4.6 (5)	1.0 (1)	1.6 (2)	8	8.05
Leu	3.0 (3)	1.9 (2)	1.1 (1)	1.0 (1)	7	6.93
Tyr	0.8 (1)				1	0.80
Phe					0	0
His		1.7 (2)			2	2.08
Lys	1.0 (1)	4.1 (4)	0.9 (1)	4.0 (4)	10	9.66
Arg	0.9 (1)	1.1 (1)	1.0 (1)		3	2.87
Trp					0	0[f]
Total	15	40	10	16	81	
Final purification[b]	PPE	Precipitation[c]	PPE	Sephadex G-25		

[a] Tryptic peptides were hydrolysed for 24 h. The composition of βTG was calculated on the basis of six glycine residues from the average of duplicate hydrolyses for 24, 48, 72, and 120 h; appropriate corrections have been made for changes associated with prolonged hydrolysis. [b] See Figure 2; PPE, preparative paper electrophoresis. [c] See text. [d] As cysteic acid. [e] As methionine sulfone. [f] Hydrolysis with 4 N methanesulfonic acid.

arising from the expected cleavages at these positions were not recovered from the Dowex 50 column in significant yields, small quantities of such peptides and free lysine were identified when the complete tryptic digest mixture was fractionated by high voltage paper electrophoresis. However, no evidence was obtained of any tryptic cleavage following the lysine at residue 2. Tryptic peptides were not submitted to amino acid sequence determination, but their individual compositions served to confirm the correctness of the amino acid sequence assignments made by automated degradation of larger polypeptide chains.

Tryptic Digestion of Citraconylated S-Carboxymethyl-β-thromboglobulin. βTG contains three arginine residues (Table II) and thus four peptides should result from tryptic cleavage of citraconylated SCM-βTG. When citraconyl groups were removed by acidification to pH 2.5 following tryptic digestion of citraconylated SCM-βTG, an insoluble residue remained which was removed by centrifugation. This material, which contained a single large peptide in pure form, was soluble in 20% acetic acid, and was desalted on Sephadex G-50 in this solvent before amino acid analysis. Three peptides were purified from the soluble portion by a combination of Sephadex

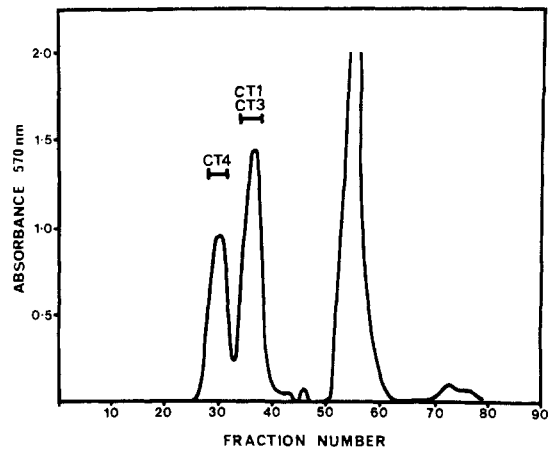

FIGURE 2: Chromatography of the soluble portion of a tryptic digest of citraconylated *S*-carboxymethyl-β-thromboglobulin (18 mg) on Sephadex G-25 fine (1.5 × 94 cm) equilibrated with 1 M acetic acid at 20 °C. Fraction size was 2.5 mL, of which 100 μL was taken for ninhydrin analysis after alkaline hydrolysis. Peptides found in each peak have been labeled in accordance with their position in the sequence proposed in Figure 3. CT4 was rechromatographed under identical conditions. Peptides CT1 and CT3 were separated by preparative high voltage paper electrophoresis at pH 3.55. The large peak in the region of the total column volume arose from the salts and buffers used and did not contain any peptide material.

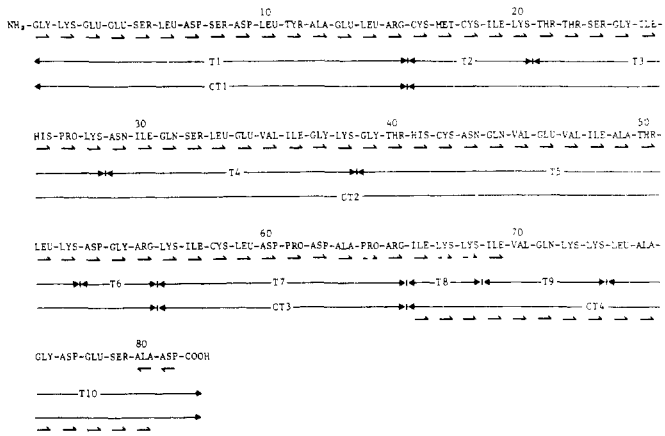

FIGURE 3: Amino acid sequence of human β-thromboglobulin subunit. The component peptides are indicated by double-headed arrows. Symbols are: (half-arrows pointing to the right) sequences obtained using the sequenator with direct identification of Pth-amino acids by thin-layer chromatography; (broken half-arrows pointing to the right) Pth-amino acid not identified directly in the automated degradation; (half-arrows pointing to the left) sequences obtained by the use of carboxypeptidase Y. Peptide nomenclature is given in Experimental Procedure.

G-25 chromatography (Figure 2) and preparative high voltage paper electrophoresis. The compositions of the four peptides (CT1, CT2, CT3, & CT4), the sum of which accounts for the intact βTG subunit, are given in Table II.

Amino Acid Sequence of β-Thromboglobulin. Automated amino acid sequence analysis of SCM-βTG (3 mg) established the sequence from the NH_2-terminal residue through to residue 69 (Figure 3). The average repetitive yield calculated from the radioactivity recovered at steps 16, 18, 42, and 58 was 97%. Unless noted, unambiguous identification of each amino acid residue was made from thin-layer chromatography. Residues which may present difficulty by this procedure because of low yields of the Pth-amino acid are serine and proline (Hermodson et al., 1972; Edman & Henschen, 1975). Serine at residues 5, 8, and 23 was clearly identifiable. Serine was also identifiable in lower yield at residue 32, and the unambiguous identification

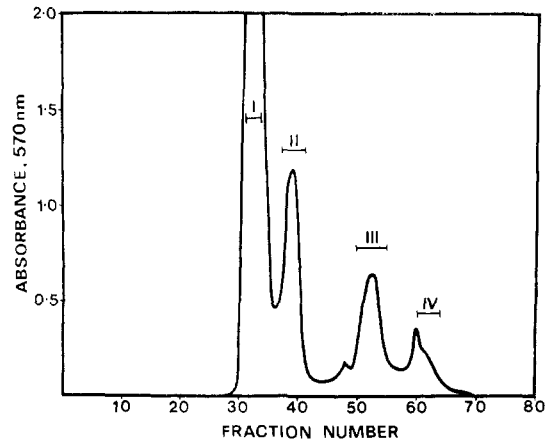

FIGURE 4: Chromatography of peptides of βTG (18 mg) following CNBr treatment and staphylococcal protease digestion. Conditions are those of Figure 2.

of all other residues in peptide T4 confirmed its assignment to this position. Similarly, although proline was not seen at position 64, the composition of peptide T7 provided the evidence to permit its assignment. Residues 67 and 68 were not identified in the degradation of intact SCM-βTG, but the composition of peptide T8 suggested the sequence Lys-Lys, which was confirmed by direct sequence determinations of peptide CT4. The position of the half-cystine residues was confirmed by measuring the radioactivity of a portion of each Pth-amino acid released. Degradation of intact βTG thus established the sequence from peptide T1 to T8, or from CT1 to CT3, and provided an overlap with CT4.

Automated sequence degradation of peptide CT4 (0.5 μmol) using the peptide program with 0.1 M Theed buffer and reduced extraction times established the sequence from residue 66 through to residue 80. The amino acid compositions of peptide CT4 and T10 indicated that the sole remaining residue was aspartic acid or asparagine. The COOH-terminal assignment was confirmed by carboxypeptidase Y digestion of SCM-βTG with an enzyme to substrate ratio of 1/125 at 37 °C which gave the following results expressed as residues of amino acid released per mole of βTG subunit: 10 min, Asp 0.29; 15 min, Asp 0.36, Ala 0.16; 20 min, Asp 0.46, Ala 0.23; 30 min, Asp 0.60, Ala 0.39. No other amino acids were detectable at these times of digestion.

Disulfide Bonds of β-Thromboglobulin. No free sulfhydryl groups could be detected in βTG with 5,5′-dithiobis(2-nitrobenzoic acid). Figure 4 illustrates the elution from Sephadex G-25 of the peptides resulting from CNBr treatment and staphylococcal protease digestion of βTG. Amino acid analysis showed that peaks I and II contained cystine. Peak III was devoid of cystine, and peak IV contained no peptide material. When examined by high voltage paper electrophoresis at pH 3.55, peak II was shown to contain a single peptide migrating with a mobility of 0.80 with respect to arginine. The composition of this peptide (Table III) corresponds to residues 14–17 plus 35–46 which would be expected from cleavages at methionyl and glutamyl residues. The peptide was recovered in a yield of 50%. This peptide was subjected to performic acid oxidation, and two peptides only, designated IIA and IIB, were isolated from high voltage paper electrophoresis of the reaction product (Table III). Their compositions confirmed that the cystine-containing peptide was composed of residues 14–17 and 35–46. These findings demonstrated a disulfide bond between half-cystine-16 and half-cystine-42. The second disulfide therefore pairs half-cystine-18 and half-cystine-58.

TABLE III: Amino Acid Compositions of a Disulfide-Containing Peptide of β-Thromboglobulin.[a]

Amino acid	Peak II	IIA	IIB
Asp	1.2 (1)	1.0 (1)	0
Thr	1.1 (1)	0.9 (1)	0
Ser	0	0	0.1
Glu	2.0 (2)	2.1 (2)	0
Pro	0	0	0
Gly	2.2 (2)	2.0 (2)	0.1
Ala	0	0	0.1
½-cystine	0.8 (2)	1.0[d] (1)	1.1[d] (1)
Val	1.8 (2)	1.8 (2)	0
Met[b]	1.2 (1)	0	1.1 (1)
Ile	0.8 (1)	0.8 (1)	0
Leu	1.1 (1)	0	1.1 (1)
Tyr	0	0	0
Phe	0	0	0
His	1.0 (1)	0.9 (1)	0
Lys	1.0 (1)	1.0 (1)	0
Arg	1.0 (1)	0	1.0 (1)
Total	16	12	4
Mobility[c]	0.80	0.48	0.63

[a] Peak II of Figure 4; IIA and IIB are the performic acid oxidation products of peak II (see text). Figures in parentheses are the theoretical compositions derived from the sequences of residues 14-17 and 35-46 (Figure 3). [b] As homoserine. [c] Mobility with respect to arginine at pH 3.55. [d] As cysteic acid.

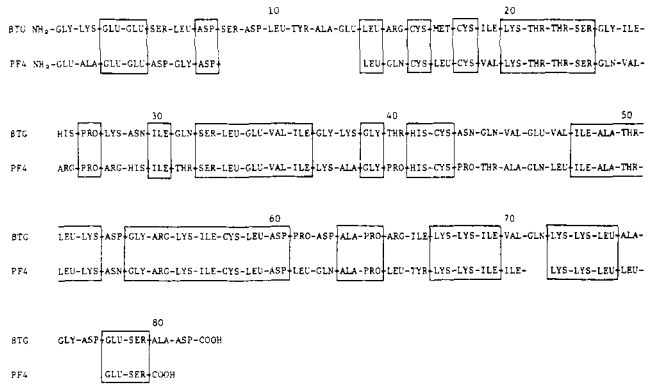

FIGURE 5: Comparison of the amino acid sequences of human β-thromboglobulin and human platelet factor 4 (Hermodson et al., 1977; Morgan et al., 1978a). Solid boxes enclose identical residues.

Discussion

The results presented here permit a proposal for the complete amino acid sequence of human βTG (Figure 3). The βTG subunit consists of a polypeptide chain of 81 amino acids containing no phenylalanine or tryptophan. The subunit contains two disulfide bridges linking residues 16 to 42, and 18 to 58. These results confirm the suggestion that Theed is an alternative to Quadrol in the automated sequence determination of proteins (Begg & Morgan, 1976) and also is useful as a nonvolatile buffer for peptide degradations. Staphylococcal protease was found to cleave βTG at the expected glutamyl residues (Houmard & Drapeau, 1972) to produce disulfide-containing peptides. Its optimum activity at pH 4 allows reaction conditions which minimize the possibility of disulfide interchange (Ryle & Sanger, 1955), and staphylococcal protease is therefore a useful alternative to pepsin which has commonly been used to obtain disulfide-containing peptides.

On the basis of the amino acid sequence the βTG subunit posseses a molecular weight of 8851. βTG has a molecular weight of 35 800 as determined by ultracentrifugation at pH 8.8 (Moore et al., 1975) but dissociation to a smaller molecular size occurs under acid conditions without reduction of disulfide bonds (Moore & Pepper, 1977). As the number and amino acid composition of the peptides isolated from the tryptic digest were consistent with the above proposal for the βTG subunit sequence, it seems that the native βTG molecule consists of four identical noncovalently bound subunits.

The most striking feature of the βTG structure is its obvious homology with PF4 (Figure 5). When three short gaps, of six, one and two residues, are inserted in the PF4 molecule to maximize homologies, 42 of the 81 residues of βTG are identical with those in the PF4 molecule. A further 11 residues, most of which represent single base mutations, can be regarded as highly favored replacements by the criteria of Dayhoff et al. (1972a). The four half-cystine residues occupy the same relative positions in the βTG molecule and in PF4, suggesting that both proteins possess the same disulfide bond structure. βTG and PF4 thus clearly represent homologous molecules arising from duplication of an ancestral gene. The position of the gaps accords with the general finding that additions and deletions tend to occur at the ends rather than in the interiors of protein chains (Dayhoff et al., 1972b). A comparison of the sequence of human PF4 with 1050 amino acid sequences by computer search of a data collection (Dayhoff et al., 1977) has shown no evidence of any significant relationship between PF4 and the other sequenced proteins in the collection.

The striking homology was unexpected from previous data. Sensitive and specific radioimmunoassays have been developed for both PF4 (Bolton et al., 1976b; Chesterman et al., 1977; Kaplan et al., 1978) and βTG (Bolton et al., 1976a). In a PF4 assay there was essentially no cross-reactivity between βTG and PF4, a 10 000-fold excess of βTG over PF4 being required to give equal displacement of tracer in the assay (Chesterman et al., 1977). Similarly between 500- and 1000-fold excess of PF4 over βTG was necessary for equal immunoreactivity in a radioimmunoassay for βTG (C. N. Chesterman, unpublished results) using a βTG antiserum previously described (Bolton et al., 1976a). These findings are consistent with a minor degree of cross-contamination of the preparations of PF4 and βTG. However, the lack of immunological cross-reactivity is not inconsistent with the present degree of homology; for example, α-lactalbumin and human lyzozyme which have extensive homologies and are derived from duplication of an ancestral gene possess no cross-reactivity in the native state (Arnon & Maron, 1970) and it has been shown that the calculation of sequence differences by measurement of immunological distance using microcomplement fixation becomes inapplicable where sequence differences of greater than 40% exist (Prager & Wilson, 1971). Comparison of the structure for βTG presented here with that of PF4 seems to strengthen the argument that the binding of heparin to PF4 depends on specific three-dimensional features rather than a simple charge interaction. It has been pointed out since there is a remarkable clustering of lysine residues at the C-terminal region of PF4 (Deuel et al., 1977) and, since modification of lysine in PF4 abolishes heparin binding (Handin & Cohen, 1976), that the lysine-rich COOH-terminal region may contain an important site for the binding of heparin. However, it is clear that most of this clustering of lysine is preserved in βTG, which has not been reported to neutralize heparin, and binds very weakly to agarose-heparin columns compared with PF4 (Moore & Pepper, 1977).

Recent studies have indicated that βTG is secreted in parallel with PF4 and the platelet-specific growth factor and that

these three proteins may be contained within the α granule (Witte et al., 1977). In view of the findings presented above, it would be of great interest to compare amino acid sequence data from the platelet-specific growth factor which bears some superficial structural resemblances to βTG and PF4.

Acknowledgment

The authors are indebted to Ms. D. Dorow for assistance with peptide purifications and amino acid analyses.

Supplementary Material Available

Sequenator programs (2 pages). Ordering information is given on any current masthead page.

References

Adelson, E., Rheingold, J. J., & Crosby, W. H. (1961) *Blood 17,* 767.

Arnon, R., & Maron, E. (1970) *J. Mol. Biol. 51,* 703.

Begg, G. S., & Morgan, F. J. (1976) *FEBS Lett. 66,* 243.

Bolton, A. E., Ludlam, C. A., Moore, S., Pepper, D. S., & Cash, J. D. (1976a) *Br. J. Haematol. 33,* 233.

Bolton, A. E., Ludlam, C. A., Pepper, D. S., Moore, S., & Cash, J. D. (1976b) *Thromb. Res. 8,* 51.

Brandt, W. F., Edman, P., Henschen, A., & Von Holt, C. (1976) *Hoppe-Seyler's Z. Physiol. Chem. 357,* 1505.

Broekman, M. J., Handin, R. I., & Cohen, P. (1975) *Br. J. Haematol. 31,* 51.

Chesterman, C. N., McGready, J. R., & Morgan, F. J. (1977) *Thromb. Haemostasis 38,* 211.

Dayhoff, M. O., Eck, R. V., & Park, C. M. (1972a) in *Atlas of Protein Sequence and Structure* (Dayhoff, M. D., Ed.) Vol. 5, p 89, National Biomedical Research Foundation, Washington, D.C.

Dayhoff, M. O., Hunt, L. T., McLaughlin, P. J., & Jones, D. D. (1972b) in *Atlas of Protein Sequence and Structures* (Dayhoff, M. D., Ed.) Vol. 5, p 17, National Biomedical Research Foundation, Washington, D.C.

Dayhoff, M. O., Hunt, L. T., Barker, W. C. & Schwartz, R. M. (1977) *Protein Sequence Data File,* October 1977, National Biomedical Research Foundation, Washington, D.C.

Deuel, T. F., Keim, P. S., Farmer, M., & Heinrikson, R. L. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74,* 2256.

Edman, P., & Begg, G. (1967) *Eur. J. Biochem. 1,* 80.

Edman, P., & Henschen, A. (1975) in *Protein Sequence Determination* (Needleman, S. B., Ed.) 2nd ed, p 232, Springer-Verlag, New York, N.Y.

Ellman, G. L. (1959) *Arch. Biochem. Biophys. 82,* 70.

Handin, R. I., & Cohen, H. J. (1976) *J. Biol. Chem. 251,* 4273.

Hermodson, M. A., Ericsson, L. H., Titani, K., Neurath, H., & Walsh, K. A. (1972) *Biochemistry, 11,* 4493.

Hermodson, M., Schmer, G., & Kurachi, K. (1977) *J. Biol. Chem. 252,* 6276.

Hirs, C. H. W. (1967) *Methods Enzymol. 11,* 197.

Houmard, J., & Drapeau, G. R. (1972) *Proc. Natl. Acad. Sci. U.S.A. 69,* 3506.

Kaplan, K. L., Nossel, H. L., Drillings, M., & Lesznik, G. (1978) *Br. J. Haematol.* (in press).

Ludlam, C. A., Moore, S., Bolton, A. E., Pepper, D. S., & Cash, J. D. (1975) *Thromb. Res. 6,* 543.

Meeker, R. L., Critchfield, F., & Bishop, E. T. (1962) *Anal. Chem. 34,* 1510.

Moore, S., & Pepper, D. S. (1977) in *Platelets in Biology and Pathology* (Gordon, J. L., Ed.) p 293, North-Holland, Amsterdam.

Moore, S., Pepper, D. S., & Cash, J. D. (1975) *Biochim. Biophys. Acta 379,* 370.

Morgan, F. J., Begg, G. S., & Chesterman, C. N. (1978a) *Thromb. Haemostasis* (in press).

Morgan, F. J., Begg, G. S., Chesterman, C. N., & Pepper, D. S. (1978b) *Thromb. Res.* (in press).

Niewiarowski, S., Lowery, C. T., Hawiger, J., Millman, M., & Timmons, S. (1976) *J. Lab. Clin. Med. 87,* 720.

Prager, E. M., & Wilson, A. C. (1971) *J. Biol. Chem. 246,* 5978.

Ross, R., & Glomset, J. A. (1976) *N. Engl. J. Med. 295,* 369.

Ross, R., Glomset, J., Kariya, B., & Harker, L. (1974) *Proc. Natl. Acad. Sci. U.S.A. 71,* 1207.

Ryle, A. P., & Sanger, F. (1955) *Biochem. J. 60,* 535.

Walsh, P. N. (1976) in *Heparin: Chemistry and Clinical Usage* (Kakkar, V. V., & Thomas, D. P., Eds.) p 125, Academic Press, London.

Witte, L. D., Kaplan, K. L., Nossel, H. L., Lages, B. A., Weiss, H. J., & Goodman, DeW. S. (1977) *Circ. Res.* (in press).