Amino Acid Sequence Studies on the α Chain of Human Fibrinogen. Characterization of 11 Cyanogen Bromide Fragments†

R. F. Doolittle,* K. G. Cassman,‡ B. A. Cottrell, S. J. Friezner,§ J. T. Hucko, and T. Takagi¶

ABSTRACT: The α chain of human fibrinogen consists of 600 ± 25 amino acid residues, 10–11 of which are methionines. In this regard, we have identified and characterized 11 cyanogen bromide peptide fragments of 2, 3, 26, 28, 28, 37, 51, 56, 60 ± 5, 64 ± 5, and 260 ± 20 residues, respectively. The sequences of five of these and a portion of a sixth have been reported previously. We now report the complete amino acid sequences of another of these fragments (56 residues), partial sequences for four others, and a preliminary characterization of the largest fragment. In a companion study (Doolittle, R. F., Cassman, K. G., Cottrell, B. A., Friezner, S. J., and Takagi, T. (1977), *Biochemistry 16* (following paper in this issue)), we have obtained key overlap sequences from plasmic digests of fibrinogen which allow all but one of these cyanogen bromide peptides to be arranged in order. The sequences of some of these newly reported fragments have revealed an internal homology in the α chain, as well as structural similarities to the corresponding portions of the β and γ chains.

The vertebrate fibrinogen molecule is composed of three pairs of nonidentical polypeptide chains ($\alpha_2\beta_2\gamma_2$). Of these, the α chains are usually the largest, ranging from 60 000–80 000 depending on the particular species (Doolittle, 1973). In the case of human fibrinogen, the molecular weight of these chains, as determined by ultracentrifugation in concentrated guanidinium chloride solutions, is 63 500 (McKee et al., 1966); molecular weights determined by dodecyl sulfate gel electrophoresis are somewhat greater, values of 70 000 and more having been reported (McKee et al., 1970). Covalently bound carbohydrate is not present in α chains (Gaffney, 1972; Pizzo et al., 1972), and the observed molecular weights ought to correspond to a linear sequence of 600 ± 25 amino acids. As in the case of the β and γ chains from fibrinogen, the bulk of amino acid sequence studies to date has been conducted on the molecule derived from human plasma. Thus, the first 78 residues from the amino terminus have been reported (Iwanaga et al., 1969; Takagi and Doolittle, 1975a), as well as a 48-residue midsection skein (Takagi and Doolittle, 1975b; Harfenist and Canfield, 1975) and the 27 residues at the carboxy terminus (Cottrell and Doolittle, 1976). In addition, the first 12 residues of the α-chain piece of a plasmin-generated digestion product—fragment D—have also been reported (Collen et al., 1975).

In this article, we report the characterization of the cyanogen bromide digestion of α chains prepared from human fibrinogen. Eleven different fragments were identified, the sum of whose residues is approximately 600 and calculated molecular weight = 68 000. For the most part, the new sequences reported here were determined by a series of extended runs with a solid-phase automatic sequencer. The results were confirmed and/or extended by the isolation of the tryptic peptides of each of these major fragments.

Materials

The following materials were obtained from the sources cited: human blood plasma (San Diego Blood Bank); human plasmin dissolved in 50% glycerol (Kabi, Stockholm); Tos-PheCH$_2$Cl[1]-treated trypsin and chymotrypsin (Worthington); thermolysin (Daiwa Kasei KK, Osaka); Whatman DEAE- and CM-cellulose (Reeve-Angel Co., Clifton, N.J.); Sephadex G-50, G-75, and G-100 (Pharmacia); cyanogen bromide and [^{12}C]iodoacetic acid (Matheson, Coleman, and Bell); [^{14}C]-iodoacetic acid (New England Nuclear); citraconic anhydride (Aldrich); dithiothreitol (Sigma); aminoethylaminopropyl glass beads (Pierce Chemicals); and phenylene diisothiocyanate (Eastman).

Sequencer Reagents and Solvents. Thioacetylthioglycolic acid (TATG) and thioacetylthioethane (TATE) were synthesized according to procedures described for similar compounds by Marvel et al. (1955) and Jensen and Pedersen (1961). Triethylamine and pyridine were refluxed with phthalic anhydride before being distilled. Trifluoroacetic acid was distilled before use. Constant-boiling HCl (5.7 N) was prepared by distilling concentrated HCl diluted volume for volume with water. All other reagents and solvents used were of analytical grade or better and were used without further treatment.

Analytical Procedures

Sodium dodecyl sulfate–polyacrylamide gel electrophoresis was conducted along the lines described by Weber and Osborn (1969). Samples were dissolved in 4 M urea–1% sodium dodecyl sulfate and incubated at 37 °C in the presence or absence of mercaptoethanol (10 μL/mL) for 16–24 h. The acrylamide concentration of the gels was varied between 4 and 12%, depending on specific requirements; in all cases, the running buffer was 0.1 M phosphate (pH 7.2) containing 0.1% sodium dodecyl sulfate. The gels, which measured 6 × 140 mm,

† From the Department of Chemistry, University of California—San Diego, La Jolla, California 92093. *Received September 29, 1976.* Supported by United States Public Health Service Grants HE-12,759 and GM-17,702.

‡ Present address: Department of Soil Science, University of Hawaii, Honolulu.

§ Present address: Department of Biochemistry, University of Washington, Seattle.

¶ Present address: Biological Institute, Tohoku University, Sendai, Japan.

[1] Abbreviations used are: TATG, thioacetylthioglycolic acid; DTE, dithioerythritol; Tos-PheCH$_2$Cl, L-1-tosylamido-2-phenylethyl chloromethyl ketone; Dns, 8-dimethylamino-1-naphthalenesulfonyl; PhNCS, phenyl isothiocyanate; CNBr, cyanogen bromide; DMF, dimethylformamide; DEAE, diethylaminoethyl; CM, carboxymethyl; Tris, 2-amino-2-hydroxymethyl-1,3-propanediol.

were run at a current of 8–10 mA/gel. A variety of proteins was used as reference substances for molecular-weight determination, including lysozyme (14 500), trypsinogen (26 000), soybean trypsin inhibitor (22 000), ovalbumin (33 000), and the light and heavy chains of rabbit immunoglobulin G (25 000 and 55 000, respectively). Gels were stained either by Coomassie brilliant blue for protein or the periodic acid–Schiff reagent for carbohydrate (Zacharius et al., 1969).

Low-voltage paper electrophoresis was conducted on an LKB apparatus at pH 2.0, 5.0, and 6.5. Guide strips were cut from papers and stained with ninhydrin, a stain for arginine (Yamada and Itano, 1966), and a stain for histidine (Smith, 1960). Papers with radioactive samples were scanned for radioactivity with a Packard Model 7201 Strip Counter. All papers were also examined for fluorescent bands, and, where relevant, guide strips stained for tryptophan (Smith, 1960). An estimation of the net charge on peptides was made using the relationship described by Offerd (1966). Elution of bands from papers was usually accomplished with 0.01 M NH_4OH. In exceptional cases, 2% acetic acid was used as an eluting agent. Occasionally, final purification of peptides was accomplished by paper chromatography in 1-butanol–acetic acid–H_2O (4:1:5).

Amino acid analysis was conducted on a Spinco Model 119 amino acid analyzer equipped with a 20-mm cuvette and an Autolab Integrator. Proteins and peptides were hydrolyzed in evacuated tubes with 5.7 N HCl at 108 °C for 24–72 h. Tryptophan was determined by a scaled-down modification of the method of Spies and Chambers (1949).

Dns-Cl–PhNCS Procedure. The Dns-Cl–PhNCS procedure was conducted according to the procedure described by Gray (1972). Identification of Dns-amino acids was accomplished on 7 × 7 cm polyamide plates (Woods and Wang, 1967).

Attachment of Peptides to Glass Beads. Usually, peptides were attached to aminoethylaminopropyl glass beads by a modification of the methods described by Laursen et al. (1972) and Wachter et al. (1973). In this regard, the amino beads were first activated by exposure to a DMF solution containing 100 mg/mL of phenylene diisothiocyanate for 3 h at room temperature. Thereupon, they were washed thoroughly with DMF, water, and, finally, 6 M guanidinium chloride. The peptide dissolved in 6 M guanidinium chloride containing 0.2 M Tris, pH 9.0, was then added to the beads and the mixture was incubated at 40 °C for 60 min. The beads were washed successively with 6 M guanidinium chloride and water. The excess isothiocyanate groups were blocked by the addition of 1.0 M ethanolamine and further incubation for 60 min and then washed with water and methanol before drying. In a few cases, the homoserine-attachment method was used as described by Horn et al. (1973).

Automatic Sequencing. Peptides attached to glass beads were degraded by a thioacetylation procedure (Mross and Doolittle, 1971; Mross and Doolittle, 1977) using thioacetylthioglycolic acid as a coupling agent. An automatic sequencing device of our own design was employed throughout (L. R. Doolittle et al., 1977). In all cases, a program was used in which the sample was coupled, washed, recoupled, washed exhaustively, dried, cleaved, and recleaved. The double-coupling, double-cleaving routine ensured maximum efficiency; indeed, the overall repetitive yield per step averaged greater than 95% in all the automatic degradations reported here. The cleaved residues dissolved in F_3AcOH were collected directly into hydrolysis tubes in a fraction collector, the F_3AcOH was evaporated, and 300 μL of 5.7 N HCl was added. The tubes were evacuated and sealed and the regeneration of free amino

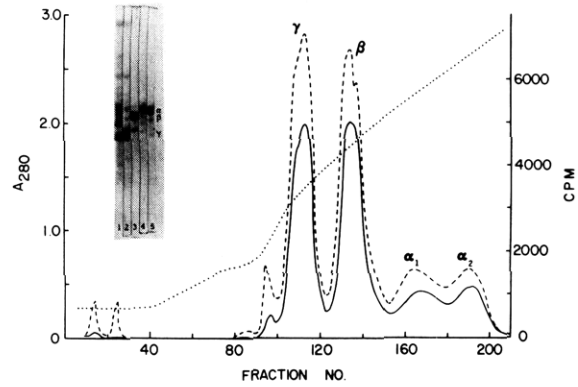

FIGURE 1: Carboxymethylcellulose chromatography (2.5 × 20 cm) of reduced and carboxymethylated human fibrinogen. Starting buffer was 0.005 M sodium acetate, pH 5.2, in 8 M urea; elution gradient was achieved by mixing equal volumes (400 mL each) of starting buffer and a limit buffer of 0.125 M sodium acetate, pH 5.2, in 8 M urea. Fraction size = 5 mL; flow rate = 60 mL/h. The photographic inset is of the corresponding sodium dodecyl sulfate gels (5% acrylamide). Gel 1, fibrinogen preparation before CMC chromatography; gel 2, γ-chain pool; gel 3, β-chain pool; gel 4, α_1-chain pool; gel 5, α_2-chain pool. All gels stained with Coomassie blue.

acids was effected at 125 °C for 4 h. After evaporation of the HCl, identifications were made directly on the amino acid analyzer described above. The recovery of serine and threonine is poor by this method, although serine characteristically yields an identifiable derivative which comes between aspartic acid and threonine. In all cases, serine and threonine residues were confirmed by independent studies on small peptides.

Thioacetylation on Fragments in Free Solution. In a few instances, degradations were conducted on fragments which were not coupled to beads. In these cases, thioacetylthioethane was the preferred coupling reagent, and benzene was the principal washing solvent.

Experimental Procedures

Purification of Reduced and Alkylated Chains from Fibrinogen. Human fibrinogen was prepared from blood-bank plasma by a modified cold-ethanol procedure (Doolittle et al., 1967). The individual polypeptide chains were prepared by dissolving about 200 mg of purified fibrinogen in 20 mL of 6 M guanidinium chloride–0.2 M Tris, pH 8.2, containing 0.01 M DTE. The solution was maintained at room temperature for 30 min, following which an equal volume of 6 M guanidinium chloride solution containing 0.03 M radioactive iodoacetic acid was added. The carboxymethylation reaction was alowed to continue for 30 min in the dark. The mixture was then dialyzed against 6 M guanidinium chloride for 60 min in the dark, and then repeatedly against distilled water at 4 °C. After dialysis, the chains were freeze-dried and then redissolved in 8 M urea, 0.005 M sodium acetate, pH 5.20, and chromatographed on CM-cellulose equilibrated with the same buffer. Separation of the chains was achieved with a linear gradient employing 0.125 M sodium acetate, pH 5.20, 8 M urea as a limit buffer. In our early experiments, we pooled the entire α-chain region together; in later experiments, we separated out an α_1 and α_2 and pooled them separately (Figure 1). We have recently shown that α_1 and α_2 chains differ in that the α_1 chains lack a 27-residue segment at the carboxy terminus (Cottrell and Doolittle, 1976). The pooled chains were dialyzed exhaustively against distilled water and then lyophilized.

Amino Acid Analysis of Purified α Chains. The results of

TABLE I: Amino Acid Composition of α Chain from Human Fibrinogen

Amino Acid	Residues/1000			Residues/68 000 (mol wt)		
	This Report[a]	McKee et al.[b]	Cartwright and Kekwick[c]	This Report	McKee et al.[b]	Cartwright and Kekwick[c]
Asp[d]	108.9	117.9	117.7	66	72	72
Thr	76.6	66.3	67.0	47	40	41
Ser	116.6	96.5	94.5	70	59	58
Glu[d]	108.9	111.1	109.8	66	68	67
Pro	62.2	52.8	60.8	37	33	37
Gly	112.3	109.3	109.6	68	67	67
Ala	40.1	42.6	40.8	25	26	25
Cys	13.5[e]	18.4[f]	17.7[g]	8	11	11
Val	47.2	47.2	49.6	29	29	31
Met	16.5	19.8	18.5	10	12	11
Ile	29.1	34.6	43.9	18	21	27
Leu	51.6	55.3	55.8	31	34	34
Tyr	16.0	19.3	19.7	10	12	12
Phe	30.9	30.8	30.0	19	18	18
His	24.0	19.8	19.5	15	12	12
Lys	66.7	72.9	65.2	41	45	40
Arg	67.2	67.4	65.3	41	41	40
Trp	18.2[h]	19.3[i]	ND[j]	11	12	
Total				612	612	604 + Trp

[a] The results are the average of duplicate analyses on 24-, 48-, and 72-h hydrolysates and include suitable extrapolations for serine, threonine, valine, and isoleucine. The chains used for these determinations were not subfractionated with regard to α_1 and α_2. [b] From McKee et al. (1966). Their original data were given as residues per 63 500 molecular weight and have been recast here as residues per 1000 residues and residues per 68 000 molecular weight. [c] From Cartwright and Kekwick (1971). Their original data were presented as residues per 10^5 g and have been recalculated here as residues per 1000 residues and residues per 68 000 molecular weight. [d] Aspartic acid and glutamic acid values include asparagine and glutamine, respectively. [e] Measured as carboxymethylcysteine. [f] Measured as cysteic acid. [g] Measured as cysteine. [h] Measured by modified method of Spies and Chambers (1949). [i] Measured spectrophotometrically. [j] ND = not determined.

our amino acid analyses of purified α chains were similar to those reported previously by others and, for the most part, were within the limits of experimental error (Table I). The exact number of methionine residues remains equivocal, however, our value of 16.5 per 1000 residues contrasting with values of 19.8 and 18.5 per 1000 residues, respectively, reported by McKee et al. (1966) and Cartwright and Kekwick (1971). These values correspond to 9.9, 11.9, and 11.1 per 600 residues, respectively. The exact number of methionine residues is important, of course, since it dictates the number of cyanogen bromide fragments which ought to be generated (Gross and Witkop, 1962).

Cyanogen Bromide Fragmentation of α Chains. Freeze-dried α chains were dissolved in 70% formic acid at concentrations of 5–7 mg/mL; a typical experiment was usually begun with 70 mg of α chain. Cyanogen bromide was added (10 mg/mL) and the reaction was allowed to continue for 16–24 h. At the end of that time, samples were diluted tenfold with ice-cold water and freeze-dried over NaOH pellets. The dried samples were dissolved in 10% acetic acid and then subjected to gel filtration on a Sephadex G-50 column equilibrated and developed with that some solvent (Figure 2).

Fractionation of α-Chain CNBr Peptides. During the course of this study, four different batches of α chains (ca. 70 mg each) were treated with cyanogen bromide and fractionated on Sephadex G-50, including one preparation each of purified α_1 and α_2 chains; a typical elution profile is depicted in Figure 2. The effluent was divided into six different pools which were freeze-dried directly. The first three pools were purified further by another gel filtration on either Sephadex G-100, G-75, or G-50, as warranted. The last three pools, which contained the smaller fragments, were purified further by low-voltage paper electrophoresis. All pools were subjected to amino acid and amino-terminal analysis before further purification, as well

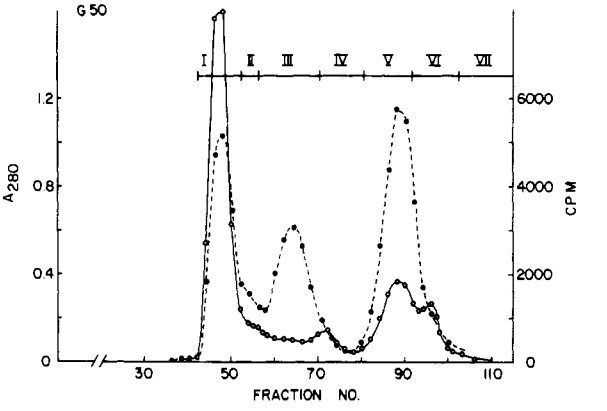

FIGURE 2: Isolation of cyanogen bromide fragments from purified α chains of human fibrinogen. Sephadex G-50 (2.5 × 95 cm) was equilibrated and developed with 10% acetic acid. Solid line = A_{280}; dotted line = radioactivity (from carboxymethylated-cysteine residues). Fraction size = 4 mL; flow rate = 20 mL/h.

as at later stages. A description of each of the pools follows, starting with the small molecular-weight material first.

Pool VII. Paper electrophoresis (pH 2.0) revealed that pool VII contains only two small cyanogen bromide peptides, prolylglutaminylhomoserine (VIIA) and arginylhomoserine (VIIB), the sum of whose (five) amino acids were the only amino acids found upon direct analysis of the pool. Fortuitously, these two peptides lie adjacent in the intact α chain, as has been shown by the isolation of an overlapping plasmin-generated peptide (Takagi and Doolittle, 1975a).

Pool VI. Pool VI contains three peptides, of 26, 28, and 28 residues, respectively. The first of these (VIC, $n = 26$ residues) does not contain homoserine and is the carboxy-terminal segment of the α chain (Table II); we have recently reported

TABLE II: Amino Acid Compositions of 11 Cyanogen Bromide Fragments from α Chain of Human Fibrinogen.

Amino Acid	CNBr Fragment[a]											Sum of Fragments	α Chain[b]
	I-II	III	IVA	IVB	VA	VB	VIA	VIB	VIC[g]	VIIA	VIIB		
Asp[c]	22	8	9	3	7	8	3	3	2			65	66
Thr	32		1	4		3	1	3	2			46	47
Ser	42	6	4	15	5	3	1	2	3			81	70
Glu[c]	23	10	8	8	5	4			3	1		62	66
Pro	19	2		2	3		5	2	1	1		35	37
Gly	45	2	1	6	7	1	1	5	3			71	68
Ala	8	3	2	1	3		1	1	4			23	25
Cys	2	2			4							8	8
Val	10	3	6	1	3	1	3	1	1			29	29
Met[d]	1	1	1	0-1[f]	1	1	1	1		1	1	9-10	10
Ile	3	5	4	2		3		1				18	18
Leu	5	7	6	1	1	4	3	4				31	31
Tyr	3	1	1	2	1	1						9	10
Phe	3		1	5	2	2	1	4				18	19
His	5		1	2	1	1	3		3			16	15
Lys	12	7	4	4	2	4	4		2			39	41
Arg	15	7	7	3	4	1		1	2		1	41	41
Trp	6[e]				2		1					9	11
Total residues	256 (±20)	64 (±5)	56	59-60 (±5)	51	37	28	28	26	3	2	610-611	612

[a] The Roman numeral designation derives from the pools indicated in Figure 2. The values are presented as residues per mole. The composition of fragment I-II is based on a molecular weight of 28 500 observed on sodium dodecyl sulfate gel electrophoresis. The other compositions are based on final amino acid sequences (IVA, VA, VIA, VIC, VIIA, and VIIB) or the best composition available on the basis of partial sequence data (III, IVB, and VIB). [b] Calculated from data in Table I using a chain molecular weight of 68 000. [c] Aspartic acid values include asparagine, and glutamic acid values include glutamine. [d] Methionines determined as homoserine. [e] Tryptophan determined by modified method of Spies and Chambers (1949). [f] Fragment IVB exists in a form without homoserine. It may correlate with the absence of fragment VIC from α_1 chains, which in turn may be the result of an in vivo plasmic cleavage (Cottrell and Doolittle, 1976). [g] Fragment VIC is not present in α_1 chains.

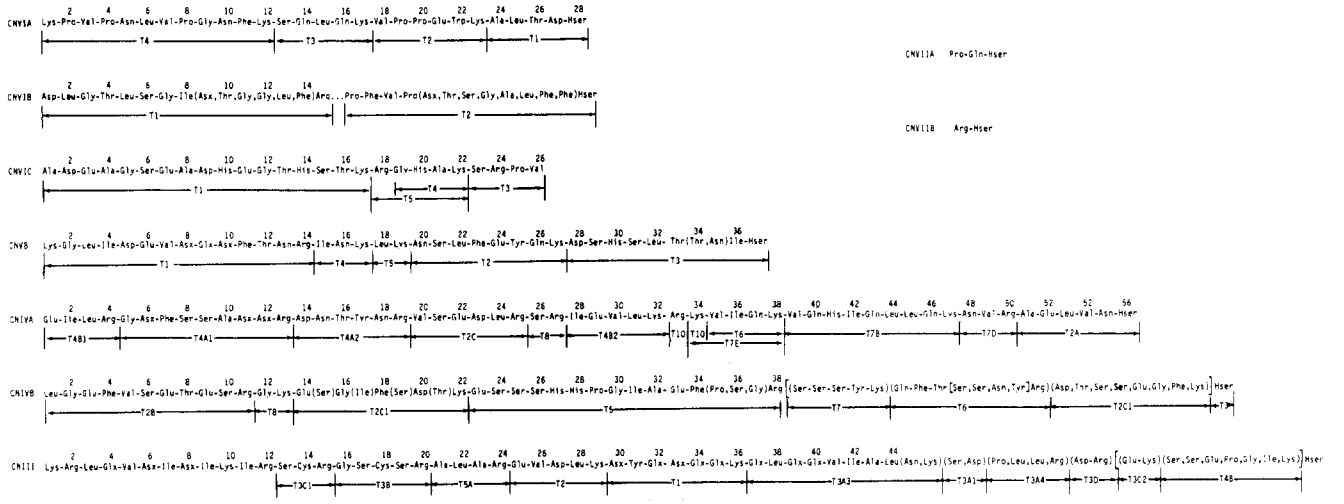

FIGURE 3: Amino acid sequence summaries of nine α chain CNBr peptides. The sequences of VIIA, VIIB, VIA, and VIC have been reported previously (Takagi and Doolittle, 1975b; Cottrell and Doolittle, 1976). The 51-residue amino-terminal fragment (CNV) previously reported by Iwanaga et al. (1969) is not shown, nor is the large molecular-weight fragment, CNI. Some of the sequences shown depend on data presented in the following article in this issue (Doolittle et al., 1977a).

its complete amino acid sequence (Cottrell and Doolittle, 1976). Both tryptic and thermolysin-derived peptides were used to confirm results obtained with the solid-phase sequencer (Figure 3). The fragment was not found in CNBr digests of α_1 chains, and the yield was low in digests of α chains which were not subdivided into pools α_1 and α_2.

We have also reported the entire sequence of the 28-residue peptide (VIA); it lies next to the two small pool VII peptides in a middle section of the chain (Takagi and Doolittle, 1975b). The third peptide in this pool (VIB, n = 28 residues) has been only partially characterized at this point, and its relative location in the chain has not been established. In fact, the parent peptide is difficult to purify, and its amino acid composition is mainly predicated on the sum of two recovered tryptic peptides (Table II). Strong evidence for the exclusive existence of these peptides in pool VI was obtained by coupling 2 mg of the unfractionated pool directly to beads and performing three rounds of thioacetylation. The only amino acids found were fully accounted for by the peptides described above.

Pool V. Pool V contains two CNBr fragments. One of these is the amino-terminal 51-residue fragment originally reported

by Iwanaga et al. (1969). We have confirmed their original sequence to the extent that we have isolated and analyzed all the constituent tryptic and chymotryptic peptides. This peptide, which begins with the 16-residue fibrinopeptide A, also contains four cysteines—which account for all the radioactivity in pool V—and two tryptophans, which contribute greatly to the ultraviolet absorbance of this pool. Pool V also contains a 37-residue fragment (CNVB), the first 27 residues of which we have reported previously in a study of the α-chain portion of fragment E (Takagi and Doolittle, 1975a). We have now completed most of the remainder of the sequence (Figure 3). In this regard, 40 nmols of peptide were attached to amino-derivatized glass beads by the homoserine-attachment method (Horn et al., 1973) and degraded through 28 residues. In addition, the five tryptic peptides of the fragment were isolated and the previously unstudied carboxy decapeptide was sequenced through its first seven residues by the Dns-Cl-PhNCS method. It was also attached to polystyrene beads by the homoserine method and subjected to ten rounds of thioacetylation. Finally, this tryptic peptide was digested with thermolysin and three subpeptides were isolated.

As in the case of pool VI, the unfractionated pool (2 mg) was also attached to glass beads and subjected to three cycles of thioacetylation. The only amino acids found were those corresponding to CNV A and B.

Pool IV. The pool IV region contains two major fragments which are separated only with difficulty. Neither fragment contains tryptophan and each contains only a single tyrosine; as a result, the absorbance at 280 nm is very weak. Neither peptide has cysteine, so no advantage could be made of radioactive carboxymethylation. The material, which is only sparingly soluble at neutral pH, could be partially resolved by gel filtration on Sephadex G-75 equilibrated with 1% acetic acid and scanned at 230 nm. The two incompletely separated pools obtained were both freeze-dried and then attached to amino-derivatized glass beads. In one case, 2 mg of αCNIVB were attached to 400 mg of amino beads by the homoserine method and the first seven residues were established. The experiment was then repeated using 2 mg of the same preparation attached to 300 mg of amino beads by the lysine-attachment method; this time 45 degradative cycles were performed. In the case of αCNIVA, 2 mg of peptide were attached to beads by the lysine procedure and 42 degradative cycles were carried out. In each case, a significant shadow sequence for the (other) contaminating peptide was apparent. The pools were also subjected to tryptic digestion and the resulting peptides were purified by paper electrophoresis. When these data were combined with the results of our studies on the isolated α chain from fragment D (following paper in this issue), an unequivocal sequence was obtained for peptide IVA ($n = 56$ residues) and a partial sequence for the other (IVB, $n = 60 \pm 5$ residues) (Figure 3). In both cases, some of the tryptic peptides were also subjected to degradation with the Dns-Cl-PhNCS procedure. Peptides T2C, T8, T6, and T2A from αCNIVA were entirely determined by this method, the results being completely in agreement with the sequencer data.

The observed amino acid compositions of IVB preparations did not agree with the proposed sequence (Figure 3) in every detail, even allowing for contamination with CNIVA. Indeed, a portion of the fragment could be isolated by paper electrophoresis which lacks carboxy-terminal homoserine, a feature which may be explained by incidental plasmin degradation; in fact, the yields of peptides T6 and T2C1 were rather low. It may be the absence of some of the carboxyl region which contributes to the anomalous amino acid compositions. In any event, the absence of homoserine in a significant fraction of this peptide leads us to believe that this is the penultimate CNBr fragment. This fragment is of considerable interest because of its involvement in α-chain cross-linking (Doolittle et al., 1977b).

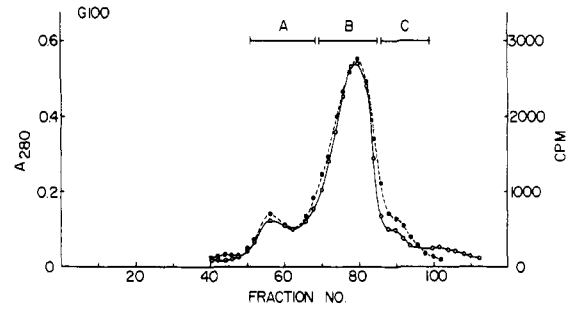

FIGURE 4: Gel filtration of pool I (Figure 2) on Sephadex G-100 (2.5 × 90 cm) equilibrated and developed with 10% acetic acid. The material in the region marked B was pooled, concentrated, and analyzed. Flow rate = 10 mL/h.

Pool III. Pool III contains one major cyanogen bromide fragment. Although it too lacks tryptophan, it contains two cysteines, and its purification was aided considerably by monitoring these radioactive carboxymethylated residues. An amino-terminal tetrapeptide sequence was first established by the manual stepwise degradation procedure using thioacetylthioethane to treat 2.5 mg of peptide. Thereafter, the first 44 residues of this fragment were established with the automatic solid-phase sequencer using 2 mg of peptide attached to 300 mg of glass beads by the lysine-coupling procedure. The sequence was confirmed by the isolation and characterization of the constituent tryptic peptides, once again starting with 2 mg of the fragment. Not all the tryptic peptides isolated were overlapped by the sequencer run, and the sequence and peptide arrangement at the carboxy terminus remain uncertain. One of these latter peptides contains a bond which is labile at low pH (evidently Asp-Pro) and yields peptides T3A1 and T3A4 (Figure 3). This same peptide also occurs in the α-chain portion of fragment D, the sequence of which is reported in the following paper of this issue (Doolittle et al., 1977a).

Pool I–II. Originally, we separated this region of the G-50 effluent into two fractions (I and II), but amino acid analysis and amino-terminal identification revealed that pool II was only material trailing from pool I. The pool contains one major fragment with a molecular weight determined by sodium dodecyl sulfate gel electrophoresis to be 28 000–30 000. The material was further purified by rechromatography on Sephadex G-100 (Figure 4). This fraction contains two cysteine residues and a disproportionate number of tryptophans (Table II). Amino acid analysis revealed that this fragment is very rich in glycine and serine (Table II). A similar fragment has been reported by Gollwitzer and Timpl (1975) and Hessel (1975). The amino-terminal sequence determined by thioacetylation stepwise degradation procedures was Glu-Leu-Glu-Arg. Digestion of this material with plasmin releases all of the radioactivity (Figure 5) and converts the parent material into a major fragment of 20 000 molecular weight (Figure 6), indicating that both cysteines are near an end of the parent fragment.

Discussion

In this article, we have reported the identification and characterization of 11 cyanogen bromide fragments from the α chain of human fibrinogen, the sum of whose amino acid

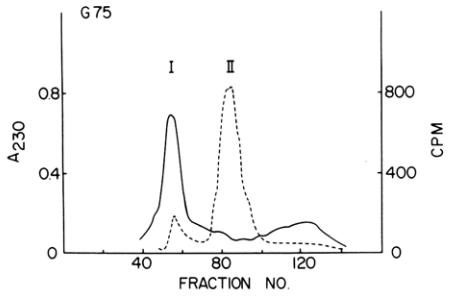

FIGURE 5: Gel filtration of plasmin-digested pool IB material on Sephadex G-75 (2.5 × 94 cm) equilibrated and developed with 0.1 M ammonium bicarbonate. Solid line = A_{230}; dotted line = radioactivity (from carboxymethylated cysteines). Flow rate = 25 mL/h; fraction size = 4 mL.

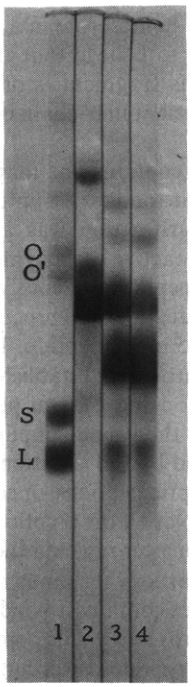

FIGURE 6: Sodium dodecyl sulfate gels (8% acrylamide) of α-chain CNBr fragment before and after treatment with plasmin at room temperature. Gel 1, references (O = ovalbumin; S = soybean trypsin inhibitor; L = lysozyme); gel 2, αCNIB before plasmin; gel 3, αCNIB after 20-min plasmin digestion; gel 4, αCNIB after 60-min plasmin digestion.

TABLE III: Cyanogen Bromide Fragments Isolated from Human Fibrinogen α Chains.

Pool[a]	Amino-Terminal Sequence	No. of Residues	Proposed Position in Chain[b]
I–II	Glu-Leu-Glu-Arg . . .	256 ● 20	H
III	Lys-Arg-Leu Glx . . .	64 ± 6	D
IVA	Glu-Ile-Leu-Arg . . .	56	C
IVB	Leu-Gly-Glu-Phe . . .	60 ± 5	I
VA	Ala-Asp-Ser-Gly . . .	51	A
VB	Lys-Gly-Leu-Ile . . .	37	B
VIA	Lys-Pro-Val-Pro . . .	28	E
VIB	Asp-Leu-Gly-Thr	28 ● 3	?
VIC[c]	Ala-Asp-Glu-Ala . . .	26	J
VIIA	Pro-Gln-Met[d]	3	F
VIIB	Arg-Met[d]	2	G
Totals:	10 methionines, 11 peptides	611 ± 34	1–611

[a] Roman numeral indicates pool designation from G50 column (Figure 2). [b] Sequential position in chain designated from A to J, where A = amino-terminal peptide and J = carboxy-terminal peptide. [c] Peptide VIC is not found in cyanogen bromide digests of α_1 chain (Figure 1). [d] Methionines identified as homoserines.

compositions is in quite good agreement with that of unfragmented chains if a molecular weight of 68 000 is presumed (Table II). We would cautiously submit that these peptides represent the entire chain, but a complete set of suitable overlaps which might prove that contention is not yet available (Doolittle et al., 1977a). On the other hand, although our own methionine values correspond to 10 residues/600 residues, previous reports (McKee et al., 1966; Cartwright and Kekwick, 1971) have noted methionine values about 10% higher (Table I), and there may indeed be another cyanogen bromide peptide which we have not found.

Of the 11 identified fragments, one is quite large (n = 260 ± 20 residues) and accounts for 40–45% of the α chain. The fragment has been isolated by several groups studying the cyanogen bromide fragmentation of fibrinogen or fibrin. Thus, Gollwitzer and Timpl (1975) have reported its molecular weight, amino acid composition, and immunological properties; our molecular weight determination and amino acid composition are in reasonable agreement with their results. These same authors have reported that this fragment contains a methionine which is resistant to cyanogen bromide cleavage, and they find fractional amounts of methionine in their analyses. They contend that the remainder of that methionine is converted to homoserine during hydrolysis. Although we find only trace amounts of methionine in our CNI preparations, it is possible that an uncleaved methionyl bond exists in this material, and that the conversion to homoserine or other by-products occurs as suggested by Gollwitzer and Timpl (1975). Our homoserine recoveries are seldom quantitative. The presence of another methionine in the fragment would tend to vitiate any discrepancy in reported methionine values (Table I) and numbers of fragments. Blombäck et al. (1974) reported an equivalent cyanogen bromide fragment which they designated Hi2-DSK, and Hessel (1975) has reported the first 11 residues of that peptide. All parties are in agreement that the fragment contains one internal disulfide bond (two cysteines). We have now shown that the disulfide loop is in a region which is selectively removed by plasmin, the parent fragment being reduced in molecular weight from 30 000 to 20 000 in the process.

Of the remaining ten cyanogen bromide fragments, the sequences of five had been reported previously, including the amino-terminal 51-residue piece (Iwanaga et al., 1969), the carboxy-terminal 26-residue peptide (Cottrell and Doolittle, 1976), and three adjacent peptides of 27, 3, and 2 residues, respectively, which comprise a midsection segment (Takagi and Doolittle, 1975).

The five new fragments reported in this article are all in the range of 3000 to 8000 in molecular weight. One of these, αCNIVA, has 56 residues, all of which we have positioned. Another, αCNVB, has 37 residues, all but two of which have been unequivocally located. A third, CNIII, has 64 ± 5 residues, the first 40 of which have been sequenced, and a fourth, αCNIVB, has had approximately half of its structure completed. The final peptide has been more elusive, although more than one-third of its postulated 28 residues has been positioned.

We have been able to make a tentative arrangement for most of these peptides based on a number of considerations. For

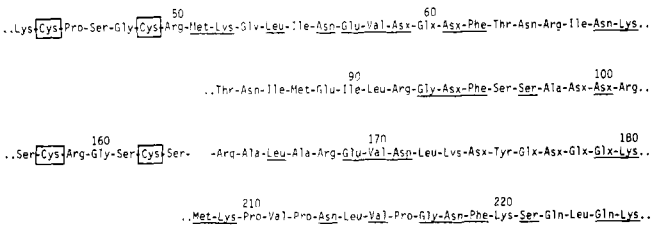

FIGURE 7: Aligned amino acid sequences from portions of various cyanogen bromide fragments showing internal homology in various portions of α chain.

example, the amino- and carboxy-terminal peptides are self-evident (Table III). Fragment CNVB overlaps the sequence of the α-chain part of fragment E (Takagi and Doolittle, 1975a), and fragments CNIVA and III are overlapped by the α chain of fragment D (Doolittle et al., 1977a). Four other midsection peptides were arranged previously (Takagi and Doolittle, 1975b). Also, evidence is presented in the following paper in this issue (Doolittle et al., 1977a) that CNIVA can be positioned by homology with the γ chain. Finally, the fact that a fraction of fragment CNIVB lacks homoserine leads us to believe that it is the carboxy-terminal penultimate CNBr fragment. If this reasoning is correct, only peptide CNVIB remains to be positioned (Table III).

The sequences of these peptides have revealed several possibly homologous segments within the α chain (Figure 7). In particular, there is a set of cysteines in αCNIII which is spaced in the exact same fashion as the set of cysteines at α 45–49 in αCNVA. The coincidence of this repeat, which also occurs in γ chains at the corresponding location,[2] is elaborated upon in the following article in this issue (Doolittle et al., 1977a). Suffice it to say here that these two cysteines are evidently involved in the same kind of three-chain disulfide arrangement as exists for the set at α 45–49 in coordination with equivalent sets of matched cysteines in the β and γ chains.

Acknowledgments

We thank Marcia Riley for preparing all the fibrinogen used in these studies and Danielle Goldbaum and Dennis Trovato for assistance in the operation of the automatic sequencer. We are also very grateful to Larry Doolittle for constructing the automatic sequencer.

Supplementary Material Available

Supplementary data are available for all three papers of this series in this issue (63 pages). Ordering information is given on any current masthead page.

References

Blombäck, B., Blombäck, M., Finkbeiner, W., Holmgren, A., Kowalska-Loth, B., and Olovson, G. (1974), *Thromb. Res. 4,* 55.

Cartwright, T., and Kekwick, R. G. O. (1971), *Biochim. Biophys. Acta 236,* 550.

Collen, D., Kudryk, B., Hessel, B., and Blombäck, B. (1975), *J. Biol. Chem. 250,* 5808.

Cottrell, B. A., and Doolittle, R. F. (1976), *Biochem. Biophys. Res. Commun. 71,* 754.

Doolittle, L. R., Mross, G. A., Fothergill, L. A., and Doolittle, R. F. (1977), *Anal. Biochem.* (in press).

Doolittle, R. F. (1973), *Adv. Protein Chem. 27,* 1.

Doolittle, R. F., Cassman, K. G., Cottrell, B. A., Friezner, S. J., and Takagi, T. (1977a), *Biochemistry 16* (second in a series of three in this issue).

Doolittle, R. F., Cassman, K. G., Cottrell, B. A., and Friezner, S. J. (1977b), *Biochemistry 16* (third in a series of three in this issue).

Doolittle, R. F., Schubert, D., and Schwartz, S. A. (1967), *Arch. Biochem. Biophys. 118,* 456.

Gaffney, P. J. (1972), *Biochim. Biophys. Acta 263,* 453.

Gollwitzer, R., and Timpl, R. (1975), *FEBS Lett. 58,* 269.

Gray, W. (1972), *Methods Enzymol. 25,* 333.

Gross, E., and Witkop, B. (1962), *J. Biol. Chem. 237,* 1856.

Harfenist, E. J., and Canfield, R. E. (1975), *Biochemistry 14,* 4110.

Hessel, B. (1975), *Thromb. Res. 7,* 75.

Horn, M., and Laursen, R. A. (1973), *FEBS Lett. 36,* 285.

Iwanaga, S., Wallen, P., Gröndahl, N. J., Henschen, A., and Blömback, B. (1969), *Eur. J. Biochem. 8,* 189.

Jensen, K. A., and Pedersen, C. (1961), *Acta Chem. Scand. 15,* 1087.

Laursen, R. A., Horn, M. J., and Bonner, A. G. (1972), *FEBS Lett. 21,* 67.

Marvel, C. S., de Radzitzky, P., and Brader, J. J. (1955), *J. Am. Chem. Soc. 77,* 5997.

McKee, P. A., Mattock, P., and Hill, R. L. (1970), *Proc. Natl. Acad. Sci. U.S.A. 66,* 738.

McKee, P. A., Rogers, L. A., Marler, E., and Hill, R. L. (1966), *Arch. Biochem. Biophys. 116,* 271.

Mross, G. A., and Doolittle, R. F. (1971), *Fed. Proc., Fed. Am. Soc. Exp. Biol. 30,* 1241.

Mross, G. A., and Doolittle, R. F. (1977), in Protein Sequence Determination, Vol. 2, Needleman, S. B., Ed., New York and Heidelberg, Springer-Verlag (in press).

Offerd, R. E. (1966), *Nature (London) 211,* 591.

Pizzo, S. V., Schwartz, M. L., Hill, R. L., and McKee, P. A. (1972), *J. Biol. Chem. 247,* 636.

Smith, I. (1960), Chromatographic and Electrophoretic Techniques, Vol. 1, New York, N.Y., Interscience.

Spies, J. R., and Chambers, D. C. (1949), *Anal. Chem. 21,* 1249.

Takagi, T., and Doolittle, R. F. (1975a), *Thromb. Res. 7,* 813.

Takagi, T., and Doolittle, R. F. (1975b), *Biochemistry 14,* 5149.

Wachter, E., Machleidt, W., Hofner, H., and Otto, J. (1973), *FEBS Lett. 35,* 97.

Weber, K., and Osborn, M. (1969), *J. Biol. Chem. 244,* 4406.

Woods, K. R., and Wang, K. T. (1967), *Biochim. Biophys. Acta 133,* 369.

Yamada, S., and Itano, H. A. (1966), *Biochim. Biophys. Acta 130,* 538.

Zacharius, R. M., Zell, T. E., Morrison, J. H., and Woodlock, J. J. (1969), *Anal. Biochem. 30,* 148.

[2] Doolittle, R. F., Takagi, T., Bouma, H., Cottrell, B. A., and Friezner, S. J. (unpublished observations).