# Amino Acid Sequence Studies on the α Chain of Human Fibrinogen. Exact Location of Cross-Linking Acceptor Sites†

B. A. Cottrell, D. D. Strong, K. W. K. Watt,‡ and R. F. Doolittle*

ABSTRACT: Human fibrinogen was clotted under conditions that promote latent factor XIII activity and in the presence of a radioactive substitute cross-linking donor ([$^{14}$C]glycine ethyl ester). The labeled fibrin was reduced and alkylated in the presence of 6 M guanidinium chloride. After dialysis and freeze-drying, the preparation was separated into its constituent polypeptide subunits by chromatography on (carboxymethyl)cellulose in the presence of 8 M urea. Under the incorporation conditions used, the radioactivity was limited to γ chains (one donor molecule/chain) and α chains (two donor molecules/chain). The labeled α chains were digested with cyanogen bromide and fractionated on Sephadex G-50. All the radioactivity was found in a fragment previously designated HαCNI, the largest of the cyanogen bromide fragments in the α chain. The fragment was further fragmented by digestion with plasmin, trypsin, chymotrypsin, and/or staphylococcal protease. The incorporated radioactivity was found to reside in equal amounts at two different sites located 38 residues apart. These were determined to be positions 88 and 126 in HαCNI, which correspond to glutamine-328 and glutamine-366 in the α chain.

The transformation of vertebrate fibrinogen molecules into a fibrin gel is a spontaneous self-assembly process that follows upon the thrombin-catalyzed release of the fibrinopeptides. The resulting gel, which is initially held together by a variety of weak forces and cooperative effects, can be stabilized further by a transglutaminase (factor XIII)-catalyzed introduction of ε-(γ-glutamyl)lysine covalent bonds. These intermolecular cross-links are known to occur reciprocally between γ chains of neighboring units in the fibrin polymer (Chen & Doolittle, 1969, 1970, 1971) and lead to γ–γ dimers after complete disruption of the system (including cleavage of disulfide bonds in the parental molecules). The same kinds of covalent cross-links can also form between α chains, although in this case the involvement leads to the formation of large polymers (multimers), even after complete disruption of the system (McKee et al., 1970). The overall geometry of these multimeric arrays has remained mysterious, although several models have been suggested (Doolittle et al., 1977b; Fretto & McKee, 1978).

The introduction of ε-(γ-glutamyl)lysine cross-links can be inhibited by a variety of small molecular weight amines, which can themselves be incorporated by factor XIII and are thus substitute donors (Lorand & Ong, 1966). These include a variety of glycine and lysine compounds, the most commonly used of which have been glycine ethyl ester (Lorand & Jacobsen, 1964), available as $^{14}$C material, and Dns-cadaverine (Lorand & Ong, 1966) which is fluorescent. The elucidation of the γ-chain cross-linking acceptor site—indeed, the entire reciprocal donor–acceptor unit—depended heavily on the use of incorporated radioactive glycine ethyl ester (Chen & Doolittle, 1969, 1970, 1971), and the same approach was thus undertaken in an initial effort to locate the α-chain acceptor sites (Chen, 1970). Takagi & Doolittle (1975) attempted to isolate the labeled α-chain fragment(s) from plasmic digests of labeled fibrin; a labeled peptide from the midsection region of the α chain was isolated, although the fractional nature of the fibrin digestion by plasmin precluded reasonable quantitation and yield studies. In an earlier report, Lorand (1972) had identified a similar peptide as an acceptor site. Subsequently, during a series of experiments on fully cross-linked fibrin, we found that the cross-linked CNBr[1] fragments did not include the peptide that we had thought was labeled (Doolittle et al., 1977b). Accordingly, we undertook a thorough reinvestigation of the incorporation of [$^{14}$C]GlyOEt into human fibrin with special attention paid to the exact sites of incorporation in α chains.

## Experimental Section

*[$^{14}$C]Glycine Ethyl Ester Labeled Fibrin.* Human fibrin with an incorporated radioactive substitute donor, [$^{14}$C]-GlyOEt, was prepared according to previously published procedures (Lorand & Jacobsen, 1964; Chen & Doolittle, 1969, 1971). The method depends on the thrombin activation of small amounts of residual factor XIII in the starting fibrinogen and the stimulation of activity by calcium ions and a suitable sulfhydryl reagent. Thus, human fibrinogen solutions (10 mg/mL in 0.25 M NaCl) were clotted by the addition of thrombin (Parke Davis; bovine) solutions containing [$^{14}$C]-GlyOEt, cysteine, and calcium ions. The final concentrations were as follows: fibrinogen = 6.5 mg/mL; thrombin = 50 μg/mL (1 unit/mL); NaCl = 0.10 M; $CaCl_2$ = 0.025 M; cysteine = 0.0125 M; GlyOEt = 0.03 M; pH 7.2. A GlyOEt stock solution was prepared fresh each time by dissolving 50 μCi of [$^{14}$C]GlyOEt (New England Nuclear) in a cold 0.15 M solution of GlyOEt. An aliquot of the solution was removed for counting; the radioactivity amounted to 194000 cpm/μmol under the conditions used for scintillation counting.

The fibrin clots were allowed to stand for 40 min at room temperature (22 °C), after which the fibrin was wound out on a glass rod, rinsed with water, and then dispersed in 6 M guanidine and 0.2 M Tris, pH 8.2. At that point an equal volume of 0.02 M DTT in 6 M guanidine and 0.2 M Tris, pH 8.2, was added to effect complete reduction of disulfide bonds. After 30 min, a 1.5-fold excess of (cold) iodoacetic acid was

† From the Department of Chemistry, University of California, San Diego, La Jolla, California 92093. *Received July 18, 1979.* Supported by U.S. Public Health Service Grants HE-12,759, HE-18,576, and GM-17,702. A preliminary account of this work was presented at the 11th International Congress of Biochemistry (Toronto), July 12, 1979 (Cottrell et al., 1979).

‡ Present address: Department of Medicine, Harvard Medical School, Boston, MA.

[1] Abbreviations used: CNBr, cyanogen bromide; GlyOEt, glycine ethyl ester; Dns, 5-(dimethylamino)naphthalene-1-sulfonyl; DTT, dithiothreitol; CM-cellulose, (carboxymethyl)cellulose.

0006-2960/79/0418-5405$01.00/0 

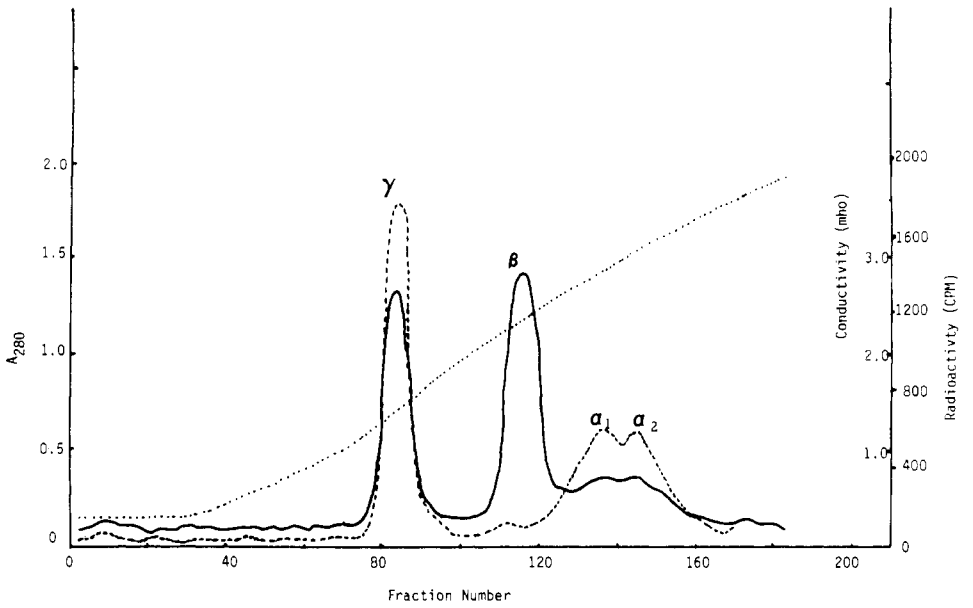


FIGURE 1: (Carboxymethyl)cellulose chromatography (2.5 × 20 cm) of 250 mg of reduced and alkylated human fibrin that had been clotted in the presence of [$^{14}$C]GlyOEt. The starting buffer was 0.005 M sodium acetate, pH 5.2, in 8 M urea; the elution gradient (dotted line = conductivity) was achieved by mixing equal volumes (500 mL each) of starting buffer and a limit buffer of 0.125 M sodium acetate, pH 5.2, in 8 M urea. Fraction size, 6.6 mL; flow rate, 60 mL/h. The solid line depicts the $A_{280}$; the broken line indicates radioactivity measured on 0.2-mL aliquots.

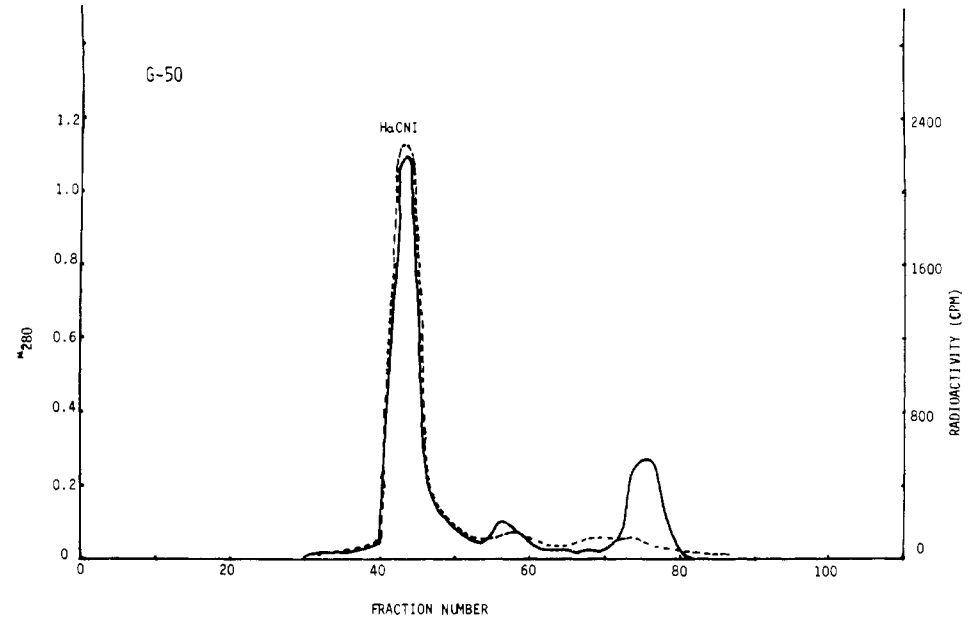


FIGURE 2: Gel filtration chromatography on Sephadex G-50 (2.5 × 150 cm) of a cyanogen bromide digest of $\alpha_2$ chains from human fibrin derivatized with [$^{14}$C]GlyOEt. The column was equilibrated and developed with 10% acetic acid. Solid line, $A_{280}$; broken line, radioactivity from [$^{14}$C]GlyOEt. Fraction size, 6.5 mL; flow rate, 78 mL/h.

added to carboxymethylate the exposed sulfhydryls. After 30 min, the preparation was dialyzed for 1 h against 6 M guanidine and then exhaustively against multiple changes of distilled water in the cold. Finally, the suspended material was transferred to flasks and freeze-dried.

*Isolation of Labeled α-Chain Fragments.* The lyophilized reduced and alkylated [$^{14}$C]GlyOEt-labeled fibrin was dissolved in starting buffer (8 M urea and 0.005 M sodium acetate, pH 5.2) for chromatography on (carboxymethyl)cellulose according to a previously described regimen (Doolittle et al., 1977a). The column was monitored at λ = 280 nm, and aliquots of fractions were counted for the presence of incorporated [$^{14}$C]GlyOEt. Both the γ-chain and α-chain peaks were radioactive, there being about twice as much total radioactivity in the latter (Figure 1). β chains were not radioactive. Appropriate fractions were pooled, dialyzed exhaustively against water in the cold, and then freeze-dried. It should be noted that human fibrinogen yields two kinds of α chains under these conditions, designated $\alpha_1$ and $\alpha_2$. We have previously shown (Cottrell & Doolittle, 1976) that $\alpha_1$ chains lack the carboxy-terminal 27 residues.

The labeled α chains were subjected to CNBr digestion (Gross & Witkop, 1962) in 70% formic acid for 16 h. The protein concentration was 7 mg/mL and the CNBr concentration 14 mg/mL. The digest was freeze-dried after a 10× dilution with cold water. The lyophilized material was subsequently dissolved in 10% acetic acid and chromatographed on Sephadex G-50. Virtually all of the radioactivity emerged in the first peak of material absorbing at λ = 280 nm (Figure 2). This CNBr fraction, the complete characterization of

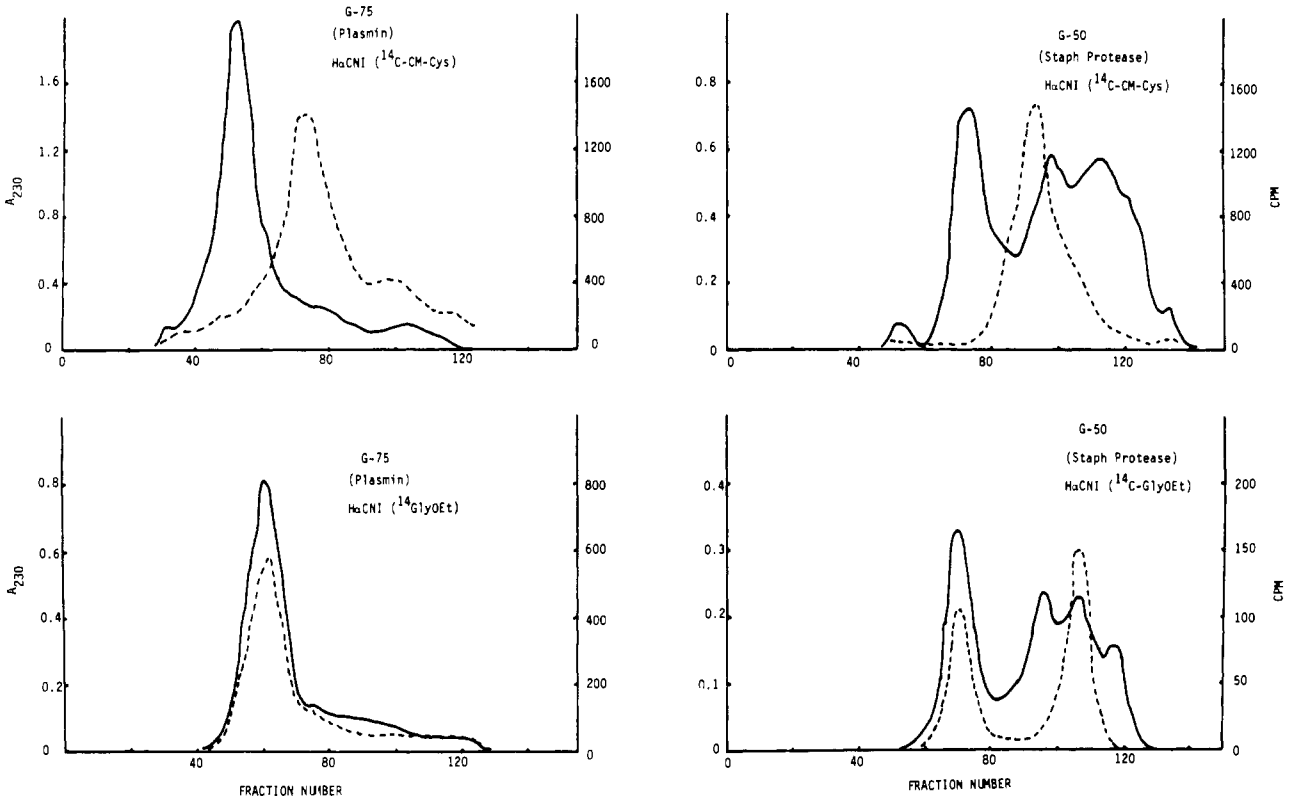


FIGURE 3: Comparison of the results of digesting two different kinds of radioactive HαCNI with plasmin (left panels) or staphylococcal protease (right panels). In the upper set, the digestions of HαCNI prepared from [$^{14}$C]CM-Cys α chains are shown; in the lower set the radioactivity comes from [$^{14}$C]GlyOEt. The chromatography conditions are comparable and are as described in Strong et al. (1979). Only about half as much material was digested in the case of the [$^{14}$C]GlyOEt-labeled HαCNI, however.

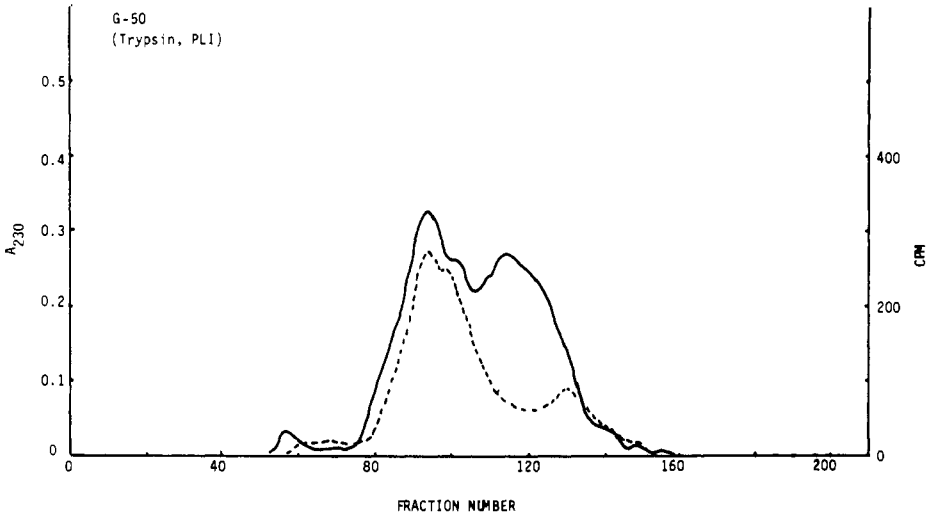


FIGURE 4: Gel filtration of tryptic digest of the plasmic fragment PLI derived from [$^{14}$C]GlyOEt-labeled HαCNI on Sephadex G-50 (2.5 × 140 cm) equilibrated and eluted with 0.1 M ammonium bicarbonate. Flow rate, 60 mL/h; fraction size, 5.0 mL. Solid line, $A_{230}$; broken line, radioactivity.

which appears in the preceding paper in this issue (Strong et al., 1979), is designated HαCNI.

*Enzymatic Digestion of [$^{14}$C]GlyOEt-Labeled HαCNI.* HαCNI is the largest of the 11 CNBr peptides that compose the α chain of human fibrinogen. It contains 236 residues, and under appropriate conditions it can be cleaved into two discrete fragments by plasmin. When [$^{14}$C]GlyOEt-labeled HαCNI was digested by plasmin, all the radioactivity was associated with the larger piece, PLI, known to represent the amino-terminal 184 residues (Figure 3). On the other hand, when [$^{14}$C]GlyOEt-labeled HαCNI was digested with staphylococcal protease, the radioactivity appeared in two discrete peaks, indicating the existence of two labeled peptides (Figure 3) and suggesting the existence of two cross-linking acceptor sites, an observation in accord with the amount of radioactivity initially incorporated in the α chains (Table I).

The labeled plasmic fragment, PLI, was further degraded with trypsin and fractionated on Sephadex G-50 (Figure 4). Under these conditions, all the radioactivity was still associated with large peptide material, although the double peak near the void volume suggested the existence of two labeled peptides. In fact, paper electrophoresis followed by autoradiography indicated the existence of four labeled peptides. Upon purification these turned out to be tryptic peptides TR4 and TR5 and their subpeptides, TR4a and TR5a, the latter resulting from the anomalous partial cleavage of two different arginyl–

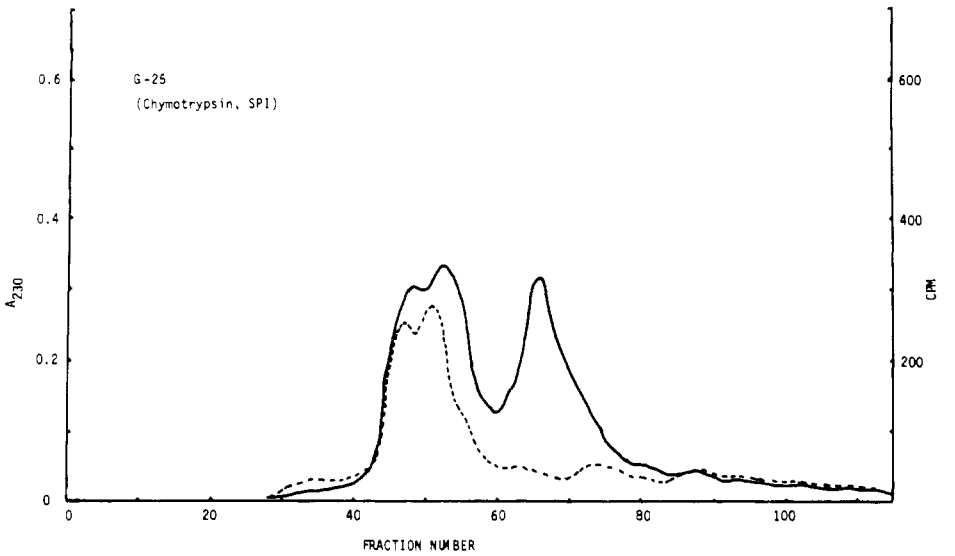


FIGURE 5: Gel filtration of a chymotryptic digest of staphylococcal protease derived fragment, SPI, on Sephadex G-25 (2.5 × 90 cm) equilibrated and developed with 0.1 M ammonium bicarbonate. Flow rate, 60 mL/h; fraction size, 4.0 mL. Solid line, $A_{230}$; broken line, radioactivity.

Table I: Recovery of Material and Radioactivity from Digestion of [$^{14}$C]Glycine Ethyl Ester Labeled Fibrin[a]

| | cpm | mass (mg) | $M_r$ | nmol | cpm/nmol | cpm/site[b] |
|---|---|---|---|---|---|---|
| reduced and alkylated fibrin applied to CM-cellulose | $10 \times 10^5$ | 256 | 170 000[c] | 1500[c] | 667 | 222 |
| CNBr-digested α chains applied to Sephadex G-50 | $3.9 \times 10^5$ | 75 | 66 125 | 1100 | 355 | 178 |
| CNI recovered from Sephadex G-50 (SP digest applied to Sephadex G-50) | $3.0 \times 10^5$ | 23 | 23 949 | 900 | 330 | 165 |
| SPI recovered from Sephadex G-50 | $6.2 \times 10^4$ | (3.3)[d] | 7 861 | 413 | 150 | 150 |
| CHI recovered from Sephadex G-25 | $4.6 \times 10^4$ | (0.7)[d] | 2450 | 270 | 169 | 169 |

[a] The specific activity of the initial [$^{14}$C]GlyOEt labeling solution was 194 cpm/nmol. [b] Fibrin contains two γ-chain acceptor sites and four α-chain acceptor sites per 340 000 molecular weight. This reduces to one per γ chain and two per α chain. [c] The number of nanomoles of fibrin is expressed per half-molecule (molecular weight 170 000). [d] Mass calculated from amino acid analyses.

Table II: Amino Acid Compositions of [$^{14}$C]GlyOEt-Labeled Peptides

| | residues/molecule[a] | | |
|---|---|---|---|
| | acceptor 1 | acceptor 2 | |
| | TR4 CH1b | TR5 CH4d | SP3d3 |
| aspartic acid[b] | 2.9 (3) | | |
| threonine | 3.1 (4) | 1.7 (2) | 1.1 (1) |
| serine | 5.8 (6) | 4.8 (5) | 4.5 (5) |
| glutamic acid[b] | 1.2 (1) | 2.4 (2) | 2.3 (2) |
| proline | 2.9 (3) | | |
| glycine | 7.7 (8)[c] | 3.1 (3)[c] | 3.2 (3)[c] |
| alanine | | | |
| valine | | 0.9 (1) | 0.9 (1) |
| isoleucine | | | |
| leucine | | | |
| tyrosine | | | |
| phenylalanine | | | |
| histidine | | | 0.9 (1) |
| lysine | | | |
| arginine | 1.0 (1) | | |
| tryptophan | +(1) | +(1) | +(1) |
| total: | (27)[c] | (14)[c] | (14)[c] |
| amino terminal: | Asn | Thr | Ser |

[a] Determined after 24-h total acid hydrolysis. Values in parentheses are integers expected from sequence studies. [b] Aspartic and glutamic acid values include asparagine and glutamine, respectively. [c] Includes the glycine residue from [$^{14}$C]GlyOEt. [d] Determined by fluorescence on paper.

proline bonds (Strong et al., 1979). The tryptic peptides were degraded further with chymotrypsin and two labeled peptides were isolated (Table II). One of these peptides (CH1b) could also be purified from chymotryptic digests of the large staphylococcal protease peptide (SPI = 84 residues) (Figure 5).

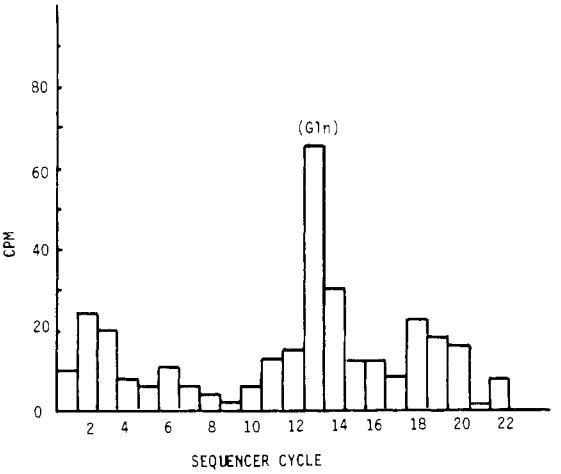


FIGURE 6: Radioactivity released at successive steps of a 22-cycle automatic degradation of peptide TR4a CH1b (Table II), attached to glass beads. Duplicate 10% aliquots of the extracted thiazolinone were counted at each step. Residue 13 is known to be a glutamine residue.

The other (CH4d) overlapped the smaller of the two labeled staphylococcal protease peptides (SP3d3) (Table II). The sequences of these peptides, which were determined independently on their underivatized counterparts (Strong et al., 1979), revealed that each contained only one glutamine residue, which must be the acceptor site for the substitute donor. In the case of acceptor site 1, it was also possible to attach the labeled peptide (TR4 CH1b) to glass beads and, by subjecting the material to 22 cycles in an automatic sequencer (Doolittle et al., 1977), to demonstrate that the radioactive derivative was located at residue 13 as expected (Figure 6).

α#1 ...Ser-Gly-Thr-Gly-Ser-Thr-Gly-Asn-Gln*-Asn-Pro-Gly-Ser-Pro-Arg-Pro-Gly-Ser-Thr-Gly-Thr-Trp..

α#2 ...Ser-Ser-Val-Ser-Gly-Ser-Thr-Gly-Gln*-Trp-His-Ser-Glu-Ser-Gly-Ser-Phe-Arg-Pro-Asp-Ser-Pro..

γ-Chain ...Asn-Arg-Leu-Thr-Ile-Gly-Glu-Gly-Gln*-Gln-His-His-Leu-Gly-Gly-Ala-Lys†-Gln-Ala-Gly-Asp-Val[C]

FIGURE 7: Comparison of amino acid sequences in the neighborhood of two α-chain cross-linked acceptor sites compared with that of the γ-chain cross-linking peptide. (*) Glutaminyl acceptor; (†) lysyl donor.

## Discussion

Previous studies on human fibrin cross-linking systems have been in agreement that the α chains contain two different sites which can accept substitute donors (Chen, 1970; Pisano et al., 1971; Doolittle et al., 1972). Moreover, the locations of these acceptor sites, whether for [$^{14}$C]GlyOEt (Takagi & Doolittle, 1975) or Dns-cadaverine (Ferguson et al., 1975), have been localized to the carboxy-terminal two-thirds of α chains and, more precisely, to its midsection region. Indeed, Finlayson & Mosesson (1973), in a definitive study involving the plasminolysis and cross-linking of fibrin, deduced that at least one portion of the authentic cross-linking system existed between residues α-220 and α-280. Subsequently, Fretto et al. (1978) delimited the chief Dns-cadaverine sites to the middle "17%" of the α chain, a span embracing approximately 100 residues. We have now pinpointed the two [$^{14}$C]GlyOEt sites to glutaminyl residues α-328 and α-366. These must also be the primary Dns-cadaverine acceptor sites, as well as the authentic cross-link acceptor sites, since they are the *only* glutaminyl residues in the middle region of the α chain (Strong et al., 1979). Moreover, a large plasmic fragment of HαCNI, PLI, can be incorporated into fibrin by factor XIII and is an effective inhibitor of fibrin stabilization.[2]

In the present study no [$^{14}$C]GlyOEt was detected in fragments containing glutamine α-237, a residue previously reported as an acceptor site for this substitute donor (Takagi & Doolittle, 1975). In that earlier study the incorporation conditions were much more sustained, however, and the nature of the recovery process was less rigorously controlled. Indeed, Fretto et al. (1978) found that about 20% of their Dns-cadaverine was incorporated in a fragment corresponding to that midsection piece, whereas 80% was in the 100-residue span existing nearer the carboxy terminus. In the present study we have kept a detailed inventory of radioactivity (Table I) in order to ascertain that the two peptides identified accounted for the principal acceptor sites.

The sequences around the two α-chain sites have been compared with that of the γ-chain cross-linking site (Figure 7). No significant homologies exist, although the second α-chain acceptor site does have the sequence ...Gly-Gln*-Trp-His..., where the γ chain has ...Gly-Gln*-Gln-His... (the asterisk indicates the acceptor site). Although the three chains of fibrinogen are clearly evolved from a common ancestor (Doolittle, 1976), the two α-chain acceptor sites exist in a region that is absent from—presumably as the result of a deletion—β and γ chains (Doolittle et al., 1979).

The question arises, does the precise location of the acceptor sites reveal anything about the nature of the lysyl donors and the overall geometry of α-chain cross-linking? It has been proposed that α-chain cross-links are mainly involved in holding fundamental polymers, which are themselves of bimolecular width, together laterally (Doolittle, 1973). The system which ensues results in a multimeric network of α chains (McKee et al., 1970) that contributes significantly to the overall stabilization of the fibrin clot. Moreover, the multimeric α-chain network survives fragmentation by cyanogen bromide, two different CNBr peptides being involved in the resulting polymer (Doolittle et al., 1977b). Because the acceptor sites are exclusively located in the larger of these fragments (HαCNI), the donors ought to be in the other peptide (HαCNIVB), which contains 67 residues and extends from α-518 to α-584 (Doolittle et al., 1979). The latter fragment contains five lysine residues in a 44-residue stretch (at α-539, α-556, α-562, α-580, and α-583), some and/or any two of which ought to be the cross-link donors; the fragment contains no glutamine. The polymeric system demands, however, that the two acceptors in a given chain must be joined to lysines that originate in two different α chains. Thus, an interactive network exists in which two acceptor sites are located in a 38-residue span in the middle of α chain, and potential donor sites are situated near the carboxy terminus more than 200 residues away. Under the circumstances, the arrangement of interlocked chains is most likely antiparallel.

[2] B. A. Cottrell, D. D. Strong, and R. F. Doolittle, unpublished experiments.

## References


Chen, R. (1970) Ph.D. Dissertation, University of California, San Diego.

Chen, R., & Doolittle, R. F. (1969) *Proc. Natl. Acad. Sci. U.S.A. 63*, 420.

Chen, R., & Doolittle, R. F. (1970) *Proc. Natl. Acad. Sci. U.S.A. 66*, 472.

Chen, R., & Doolittle, R. F. (1971) *Biochemistry 10*, 4486.

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

Cottrell, B. A., Strong, D. D., Watt, K. W. K., & Doolittle, R. F. (1979) *Proc. Int. Congr. Biochem., 11th*, Abstract No. S24.

Doolittle, L. R., Mross, G. A., Fothergill, L. A., & Doolittle, R. F. (1977) *Anal. Biochem. 78*, 491.

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

Doolittle, R. F. (1976) *Fed. Proc., Fed. Am. Soc. Exp. Biol. 35*, 2145.

Doolittle, R. F., Cassman, K. G., Chen, R., Sharp, J. J., & Wooding, G. L. (1972) *Ann. N.Y. Acad. Sci. 202*, 114.

Doolittle, R. F., Cassman, K. G., Cottrell, B. A., Friezner, S. J., Hucko, J. T., & Takagi, T. (1977a) *Biochemistry 16*, 1703.

Doolittle, R. F., Cassman, K. G., Cottrell, B. A., & Friezner, S. J. (1977b) *Biochemistry 16*, 17.

Doolittle, R. F., Watt, K. W. K., Cottrell, B. A., Strong, D. D., & Riley, M. (1979) *Nature (London) 280*, 464.

Ferguson, E. W., Fretto, L. J., & McKee, P. A. (1975) *J. Biol. Chem. 250*, 7210.

Finlayson, J. S., & Mosesson, M. W. (1973) *Thromb. Res. 2*, 467.

Fretto, L. J., & McKee, P. A. (1978) *J. Biol. Chem. 253*, 6614.

Fretto, L. J., Ferguson, E. W., Steinman, H. M., & McKee, P. A. (1978) *J. Biol. Chem. 253*, 2184.

Gross, E., & Witkop, B. (1962) *J. Biol. Chem. 237*, 1856.
Lorand, L. (1972) *Ann. N.Y. Acad. Sci. 202*, 6.
Lorand, L., & Jacobsen, A. (1964) *Biochemistry 3*, 1939.
Lorand, L., & Ong, H. H. (1966) *Biochemistry 5*, 1747.
McKee, P. A., Mattock, P., & Hill, R. L. (1970) *Proc. Natl. Acad. Sci. U.S.A. 66*, 738.
Pisano, J. J., Finlayson, J. S., Peyton, M. P., & Nagai, Y. (1971) *Proc. Natl. Acad. Sci. U.S.A. 68*, 770.
Strong, D. D., Watt, K. W. K., Cottrell, B. A., & Doolittle, R. F. (1979) *Biochemistry* (first of three papers in this issue).
Takagi, T., & Doolittle, R. F. (1975) *Biochemistry 14*, 5149.

# Amino Acid Sequence Studies on the $\alpha$ Chain of Human Fibrinogen. Overlapping Sequences Providing the Complete Sequence†

K. W. K. Watt,‡ B. A. Cottrell, D. D. Strong, and R. F. Doolittle*

ABSTRACT: The complete amino acid sequence of the $\alpha$ chain of human fibrinogen has been determined. It contains 610 amino acid residues and has a calculated molecular weight of 66 124. The chain has 10 methionines, and fragmentation with cyanogen bromide yields 11 peptides [Doolittle, R. F., Cassman, K. G., Cottrell, B. A., Friezner, S. J., Hucko, J. T., & Takagi, T. (1977) *Biochemistry 16*, 1703]. The arrangement of the 11 fragments was determined by the isolation of peptide overlaps from plasmic and staphylococcal protease digests of fibrinogen and/or $\alpha$ chains. In addition, certain of the cyanogen bromide fragments, preliminary reports of whose sequences have appeared previously, have been reexamined in order to resolve several discrepancies. The $\alpha$ chain is homologous with the $\beta$ and $\gamma$ chains of fibrinogen, although a large repetitive segment of unusual composition is absent from the latter two chains. The existence of this unusual segment divides the sequence of the $\alpha$ chain into three zones of about 200 residues each that are readily distinguishable on the basis of amino acid composition alone.

The $\alpha$ chain is the largest of the three nonidentical subunits that constitute the vertebrate fibrinogen molecule ($\alpha_2\beta_2\gamma_2$). In the case of human fibrinogen, the $\alpha$-chain molecular weight has been found to be 63 500 by ultracentrifugation (McKee et al., 1966) and 70 000 by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (McKee et al., 1970). We have now completed the amino acid sequence of the chain; it contains 610 amino acid residues and has a calculated molecular weight of 66 124.

Although the completion of the sequence of the largest cyanogen bromide fragment (Strong et al., 1979) actually represented the final step in the determination, there are a number of data bearing on the arrangement of the CNBr[1] fragments and details of sequences reported in brief and preliminary communications that have not been published previously. Also, there are some discrepancies with other reports that need resolving. In this article we provide the final data from this laboratory bearing on the amino acid sequence of the $\alpha$ chain of human fibrinogen. The sequence has provided a wealth of information about the structure, function, and evolution of the chain itself and the fibrinogen molecule as a whole, including features dealing with fibrin stabilization and fibrinolysis.

## Experimental Section

*Materials and Methods.* All the procedures used in this study have been fully described in previous publications from this laboratory (for a listing of these references, see the preceding two articles in this issue).

*Amino Acid Composition of $\alpha$ Chains and CNBr Fragments.* In an earlier article (Doolittle et al., 1977a), we published the results of amino acid analysis of human fibrinogen $\alpha$ chains and compared them with results reported from other laboratories. Although the agreement with the final sequence was quite good (we had noted 611 residues), the serine value was low and values for the other amino acids were slightly inflated as a result. Accordingly, we undertook another set of analyses, taking care to use only $\alpha_2$ chains, which extend all the way to the authentic carboxy terminus of the chain (Cottrell & Doolittle, 1976). The results are in excellent agreement with the reported sequence (Table I).

In that earlier publication (Doolittle et al., 1977a) we also reported the amino acid compositions of the 11 CNBr peptides based on the sequences observed to that point. In this article we now present the actual amino acid compositions determined on the individual purified fragments and compare them with the number of residues expected from the observed sequences (Table II). Although the agreement is generally excellent, a few notable exceptions do occur. In particular, fragment CNI contains two more alanine residues than were identified in the sequence operations.

*Arrangement of CNBr Fragments.* Previously, we had suggested an arrangement for 10 of the 11 CNBr fragments (Doolittle et al., 1977a), based, for the most part, on inferences derived from earlier and concurrent studies on the plasmic digestion of fibrinogen (Takagi & Doolittle, 1975a,b; Doolittle et al., 1977b). The position of the 11th peptide was subsequently established (Cottrell & Doolittle, 1978), also as the result of isolating an overlapping plasmic peptide from fibrinogen. In this report we summarize the data used as a basis for CNBr fragment arrangement (Table III), including for the first time the results of digesting $\alpha$ chains with staphylo-

† From the Department of Chemistry, University of California, San Diego, La Jolla, California 92093. *Received July 18, 1979.* Supported by U.S. Public Health Service Grants HE-12,759, HE-18,576 and GM-17,702.

‡ Present address: Department of Medicine, Harvard Medical School, Boston, MA.

[1] Abbreviations used: CNBr, cyanogen bromide; Dns, 5-(dimethylamino)naphthalene-1-sulfonyl; PhNCS, phenyl isothiocyanate.

0006-2960/79/0418-5410$01.00/0 