Supporting information for Madrazo *et al.* (2001) *Proc. Natl. Acad. Sci. USA* **98** (21), 11967–11972. (10.1073/pnas.211439798)

Materials and Methods

Preparation and Characterization of Bovine Fibrinogen E_5 Fragment. A 45-kDa fragment E was purified from a 2-h plasmic digest of bovine fibrinogen as described (1, 2). This fragment was denoted as E_3 because its NH_2 termini (AαLeu-23 or Gln-27, BβLys-61 and γTyr-1) determined by NH_2-terminal sequence analysis by using a Hewlett--Packard G1000S sequenator were equivalent to those reported for human fibrinogen E_3 fragment (3). Further digestion of bovine E_3 fragment with chymotrypsin (Calbiochem) resulted in the appearance of two new discrete fragments with molecular masses of 40 and 35 kDa, denoted as E_4 and E_5, respectively (Fig. 1*d*). This digestion was performed at room temperature in 20 mM Tris buffer, pH 7.4/0.15 M NaCl. The protein concentration was 2 mg/ml and the enzyme/substrate ratio was 1:100 (wt/wt). The E_5 fragment was purified from a 10-h digest by size-exclusion chromatography by using a 2.5×110-cm column with Biogel P60, dialyzed against 25 mM imidazole buffer, pH 6.5, concentrated to 15-20 mg/ml by ultrafiltration by using YM 10 membrane (Amicon), and stored at –20°C.

The NH_2-terminal sequence analysis of E_5 revealed that the NH_2 termini of its Bβ and γ chains were essentially the same as in E_3, whereas those for the Aα chains starting at Ser-29 were different. To determine the complete sequence of E_5, this fragment was reduced and derivatized by the method described earlier (4), the chains were separated by HPLC on a reverse-phase Nucleosil C-18 column, and individual chains were digested with trypsin, endopeptidase Lys-C, or CNBr. The resulting peptides were purified by HPLC and sequenced through the end. The complete sequence of E_5 is presented in Fig. 1*c*. Comparison of its NH_2 and COOH termini with those established for the human fibrinogen E_3 fragment (3) revealed that the generation of E_5 is due to the removal of NH_2-terminal residues 23-28 from the Aα chain and several residues from the COOH termini of the Bβ and γ chains. A substantial decrease in the molecular mass of the E_5 fragment can be explained by the removal of two carbohydrates linked to Asn-52 of each γ chain; the intermediate E_4 fragment contains, most probably, one carbohydrate, whereas the other one is removed.

Crystallization. Crystals of the E_5 fragment were grown by the vapor-diffusion method in hanging or sitting drops. Drops containing 2-3 μl of protein (at 10 mg/ml) and the same volume of well solution [15.0% polyethylene glycol (PEG) 3350/150 mM $CaCl_2$/3.5% dioxane/3mM NaN_3, in 20 mM Tris•HCl, pH 8.0] were mixed and equilibrated against 0.5 ml of the reservoir well solution. Two crystal forms of different morphologies grew at room temperature in 2-4 days in these conditions. Bipyramidal-shaped crystals grew to a size of ~0.2 mm × 0.3 mm × 0.6 mm and belong to the orthorhombic space group $P2_12_12_1$ with unit cell dimensions of a = 53.4 Å, b = 58.8 Å, and c = 96.8 Å. In addition, needle-shaped crystals grew to ~0.03 mm × 0.05 mm × 0.3 mm, and occasionally large rod-shaped crystals grew to ~0.6 mm × 0.6 mm × 1.0 mm. These crystals belong to the monoclinic space group $P2_1$, and the large rod-shaped crystals have unit cell dimensions

of a = 49.4 Å, b = 66.2 Å, and c = 50.7 Å with β = 106.6ºC. (Under similar crystallization conditions, but at 4ºC, bar-shaped crystals grew to 0.2 mm × 0.2 mm × 0.4 mm after 3-4 weeks. These crystals belong to the same monoclinic space group as those obtained at room temperature.) Crystals were transferred into a cryoprotectant solution (30% PEG 3350/10% glycerol/150 mM $CaCl_2$/3.5% dioxane/3 mM NaN_3, in 20 mM Tris•HCl, pH 8.0), flash-cooled in liquid propane, and stored in liquid nitrogen until data collection.

The orthorhombic crystals were used for the search of heavy atom derivatives, as they usually grew faster and bigger and were easier to handle than the monoclinic crystals. After the screening of a number of different heavy atom compounds, a trimethyllead-acetate [$Pb(CH_3)_3OAc$] derivative was obtained. The heavy atom compound was dissolved to a final concentration of 23 mM in the cryoprotectant solution that did not contain NaN_3. Crystals were transferred into the cryoprotectant solution, soaked with the heavy atom compound for 4 weeks, and then flash-cooled in liquid propane and stored in liquid nitrogen.

Biochemical Characterization of Crystals of the E_5 Fragment. To ensure that the E_5 fragment was not modified on crystallization, several crystals were briefly washed in the well solution and dissolved in 20 mM Tris buffer, pH 8.5, sequenced, and tested for the presence of free sulfhydryls. The NH_2-terminal sequence analysis confirmed that the NH_2 termini of E_5 from the crystals were the same as in the starting material. A test for the presence of free sulfhydryls was performed as described earlier (5). Briefly, the fragment at 1 μM was mixed with 50 μM pyrene-maleimide, a specific reagent for SH groups, and the fluorescence signal was monitored in time at 395 nm in an SLM 8000-C fluorometer. No changes in fluorescence were observed in 30 min, during which the fluorescence of the control protein containing a single SH group changed substantially, indicating no free SH groups in the crystal-derived E_5 fragment.

Data Collection and Processing. All data sets were collected at cryotemperature (100 K). In-house diffraction data were collected on the Rigaku (Tokyo) RAXIS-IV image plate detector mounted on the Rigaku RU300B rotating anode X-ray generator (CuKα radiation). Higher resolution data sets were collected on beamline X26C at the National Synchrotron Light Source, Brookhaven National Laboratory, New York, and on beamline A1 at the Cornell High Energy Synchrotron Source in Ithaca, NY. Data collection statistics are summarized in Table 1. Data reduction and processing were carried out with the set of programs HKL (6).

Model Building and Refinement. The structure was determined first at 3.0-Å resolution with the data from the orthorhombic crystals by using a combination of Single Isomorphous Replacement and density modification with the program DM (7) from the CCP4 suite (8). By using the program SOLVE (9), three substitution sites for the $Pb(CH_3)_3OAc$ derivative were found independently from the data sets collected in the laboratory and at the synchrotron (see Table 1). Refinement of the heavy atom positions, by using the program MLPHARE from the CCP4 program suite (8), however, produced different heavy atom occupancy values for the two sets of data. As a result, we found that the best quality electron density map could be obtained after combining phases from both

sets of data by using the program MLPHARE, followed by density modification with the program DM (final figure of merit 0.72). This map clearly revealed the positions of the six α-helices of the coiled-coil regions in each half-molecule. The helices were first modeled as poly-Ala by using the program O (10), and their positions were refined with the program package CNS (11). The quality of the electron density map was further improved with the program SIGMAA (12) by combining heavy atom derivative phases with those from the partial helical model and extending the data used to the 1.6-Å resolution limit of the native orthorhombic crystals. At this stage of the refinement, the side chains were assigned in the coiled-coil regions of the model. Multiple cycles of model building (by using the program O) and refinement (by using the program CNS) within the non-coiled-coil central region of the fragment were necessary to complete the model of the E_5 fragment. In the orthorhombic crystal form, the N-terminal residues Aα29-34, Aα′29-34, Bβ61-63, Bβ′61-63, γ1, and γ′1, and the C-terminal residues Aα79-81, Aα′79-81, Bβ115-116, Bβ′115-116, and (in one monomer only) γ48 were disordered and not included in the final model.

By using this model, phases for the monoclinic crystal form were determined by molecular replacement by using the program AMORE (13). The structure was refined to 1.4 Å, initially by using ARP (14) and then CNS. Here, the N-terminal Aα29-34, Aα′29-34, Bβ61-63, Bβ′61-63, γ1-4, and γ′1, and the C-terminal residues Aα78-81, Aα′78-81, Bβ115-116, and Bβ′114 were disordered and not included in the final model. Refinement statistics are summarized in Table 2.

The figures were made with the programs MOLSCRIPT (15), RASTER3D (16, 17), GRASP (18), XTALVIEW (19), and MOLE (Applied Thermodynamics, Hunt Valley, MD).

The Coiled-Coil Domains of E_5 (Expanded). The α-helical coiled-coil domains of E_5, consisting of residues Aα50-78, Bβ85-114, and γ21-48, have two noncanonical structural features. The sequences of coiled coils are characterized by their so-called "heptad repeat," where every third then fourth residue is usually apolar and close-packed in the core (20); (for review, see ref. 21). In the E_5 fragment, there is one three-residue deletion from the heptad repeat of each chain, located roughly at homologous positions (Aα65, Bβ100, and γ36) midway along the coiled-coil domain (Fig. 1). These deletions, or "stutters," result in local non-close-packed cores as found in certain other coiled coils (22, 23). In the homotrimeric coiled coil of hemagglutinin (24), similar stutters and non-close-packed cores have been described (22, 23). The conservation of the stutter in the three chains of fibrinogen, along with other sequence similarities, is consistent with the evolution of at least part of the molecule from a homotrimeric precursor (25). In addition, there is a proline residue in this stutter region of the Bβ chains at position 99 (Fig. 2). The location of this residue coincides with a bend in the Bβ-chain helix. The degree of bending varies (between ~12 and 18°) in the two halves of the dimer and in the two crystal forms. The stutter and the proline residue are conserved among a number of vertebrate species, suggesting that these features that promote flexibility may be related to the functions of fibrinogen.

Detailed Description of the Structures of the Six Chains in the Two Central Domains of Fragment E_5 (Expanded). As seen in the crystal structure of fragment E_5, the N-terminal portions of the Aα (residues 35-49) and Bβ (residues 64-84) chains from each half-molecule (or "monomer") have distinct structural roles in the formation of the funnel-shaped domain. The segments of the Aα chains are located almost entirely within the central plane that bisects the long axis of the molecule (Fig. 4). The most N-terminal residues of the Aα chains, Gly-35 and Trp-36, form part of the rim of the funnel-shaped domain. The two chains diverge from one another (from residues 36 to 41) and then converge in a partially helical (residues 43 to 47) and partially extended (residues 48 and 49) conformation, so that they wrap around the Bβ chains.

In contrast to the Aα chains, the Bβ chains in the funnel-shaped domain extend along the long axis of the molecule and interact extensively with the coiled-coil domains. The most N-terminal residues of the Bβ chains, 64-69, are in extended conformations and form the remainder of the central cavity's rim. Residues 64 and 68 of each of the Bβ chains are also stabilized by salt bridges to their respective monomers' coiled-coil domains at residues Aα60 and Aα53. Residues 70-84 of each Bβ chain form a relatively long loop, in which residues 70-73 and 80-83 form a two-stranded antiparallel β-sheet. This loop extends toward the coiled-coil domain of the opposite half-molecule: Bβ Leu-79 near the reverse turn forms hydrophobic contacts with Aα′Met54, Aα′Leu57, and γ′Ala26 of the opposite (denoted by ′) monomer's coiled-coil domain, and BβAsp-78 forms a salt bridge with Bβ′Arg-64 (the reciprocal contacts occur as well). One face of each loop forms a major portion of the cavity's internal surface (or walls), whereas BβCys-72 on the opposite face of the loop is disulfide linked to Cys-39 of the Aα chain portion of this domain. The cavity's surface in this domain is unusual in being dominated by uncharged and hydrophobic amino acid side chains (Fig. 8).

Each of the two γ chains in the γN-domain (residues 1-20), like the Bβ chains of the funnel-shaped domain, also contribute to the formation of a convoluted dimeric interface (Fig. 2). Following residues 4-7, which form short helices, and the disulfide-forming cysteines at positions 8 and 9, residues 10-16 of each γ chain form a loop that, like the longer loops of the Bβ chains, extends toward and interacts with the opposite monomer's coiled coil. Here, γArg-14 forms a salt bridge with Bβ′Glu-97, and γPhe-15 makes hydrophobic contacts with γ′Phe-28, Bβ′Leu-93, and Bβ′Val-94. Residues γ17-21 and γ′17-21 then fold back toward their respective coiled-coil domains, crisscrossing *en route* at residues γ19 and γ′19, which participate in a short antiparallel β-sheet and form the floor of the funnel-shaped cavity described above. The positions of the N-terminal 14 residues of the γ and γ′ chains are significantly different from one another (Fig. 7); this asymmetry in the γN domain is stabilized by the disulfide bonds between residues 8 and 9 (Fig. 3*d*).

The Locations and Functions of the Central Domains in the Protofibril of Fibrin (Expanded). The two-stranded protofibril of fibrin formed after reaction with thrombin can now be modeled by taking into account the domain structure of E_5 (Fig. 6). In E_5, residues Aα35 and Aα′35 (the most N-terminal residues traced) are located within 21 Å

of each other, and weaker electron density seen in the bovine fibrinogen map (26) suggests that the disulfide bond between residues Aα31 and Aα′31 is positioned above the funnel-shaped cavity nearly coinciding with the 2-fold axis of the dimer (dashed lines in Fig. 3*a*). The two Aα GPR knobs (only 10-12 residues away along the sequence at positions 19-21) are thus constrained to be located roughly on the same side of the molecule as the funnel-shaped domain. In a closed half-staggered protofibril of fibrin, we therefore expect that this domain of the E region from one filament would face the two closely situated γ-domain receptor pockets for these knobs on the adjacent filament (Fig. 6*a*). [Residues Bβ64 and Bβ′64 of E_5 (the most N-terminal residues observed for these chains) are located ~35 Å from each other on opposite sides of the funnel-shaped domain, displaced from one another along the long axis of the molecule (Fig. 4)]. Given the current (incomplete) packing models of fibrin, the positions of these Bβ residues are consistent with the binding of the β knobs [each located ~50 residues away along the sequence) to βC-domain receptor pockets within (27) and/or between (28, 29) protofibrils.] The E_5 structure also indicates that the γN domain would be situated on the exterior side of the two-stranded protofibril and thus be positioned for possibly influencing associations between protofibrils. [Such interactions using the γN domain, however, have not yet been reported (27)].

1. Medved, L. V., Platonova, T. N., Litvinovich, S. V. & Lukinova, N. I. (1988) *FEBS Lett.* **232,** 56-60.

2. Privalov, P. L. & Medved, L. V. (1982) *J. Mol. Biol.* **159,** 665-683.

3. Olexa, S. A., Budzynski, A. Z., Doolittle, R. F., Cottrell, B. A. & Greene, T. C. (1981) *Biochemistry* **20,** 6139-6145.

4. Litvinovich, S. V., Henschen, A. H., Krieglstein, K. G., Ingham, K. C. & Medved, L. V. (1995) *Eur. J. Biochem.* **229,** 605-614.

5. Brown, R. D. & Matthews, K. S. (1979) *J. Biol. Chem.* **254,** 5128-5134.

6. Otwinowski, Z. & Minor, W. (1997) *Methods Enzymol.* **276,** 307-326.

7. Cowtan, K. D. (1994) *Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography* **31,** 34-38.

8. Collaborative Computational Project, Number 4 (1994) *Acta Crystallogr. D* **50,** 760-763.

9. Terwilliger, T. C. & Berendzen, J. (1999) *Acta Crystallogr. D* **55,** 849-861.

10. Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991) *Acta Crystallogr. A* **47,** 110-119.

11. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., *et al.* (1998) *Acta Crystallogr. D* **54,** 905-921.

12. Read, R. J. (1986) *Acta Crystallogr. A* **42,** 140-149.

13. Navaza, J. (1994) *Acta Crystallogr. A* **50,** 157-163.

14. Perrakis, A., Morris, R. M. & Lamzin, V. S. (1999) *Nat. Struct. Biol.* **6,** 458-463.

15. Kraulis, P. J. (1991) *J. Appl. Crystallogr.* **24,** 946-950.

16. Bacon, D. J. & Anderson, W. F. (1988) *J. Mol. Graphics* **6,** 219-220.

17. Merritt, E. A. & Murphy, M. E. P. (1994) *Acta Crystallogr. D* **50,** 869-873.

18. Nicholls, A., Sharp, K. A. & Honig, B. (1991) *Proteins* **11,** 281-296.

19. McRee, D. E. (1992) *J. Mol. Graphics* **10,** 44-46.

20. Crick, F. H. C. (1953) *Acta Crystallogr.* **6,** 689-697.

21. Cohen, C. & Parry, D. A. D. (1990) *Proteins Struct. Funct. Genet.* **7,** 1-15.

22. Lupas, A., Müller, S., Goldie, K., Engel, A. M., Engel, A. & Baumeister, W. (1995) *J. Mol. Biol.* **248,** 180-189.

23. Brown, J. H., Cohen, C. & Parry, D. A. D. (1996) *Proteins Struct. Funct. Genet.* **26,** 134-145.

24. Wilson, I. A., Skehel, J. J. & Wiley, D. C. (1981) *Nature (London)* **5796,** 366-373.

25. Doolittle, R. F., Everse, S. J. & Spraggon, G. (1996) *FASEB J.* **10,** 1464-1470.

26. Brown, J. H., Volkmann, N., Jun, G., Henschen-Edman, A. & Cohen, C. (2000) *Proc. Natl. Acad. Sci. USA* **97,** 3907-3912. (First Published March 28, 2000; 10.1073/pnas.080065697)

27. Yang, Z., Mochalkin, I. & Doolittle, R. F. (2000) *Proc. Natl. Acad. Sci. USA* **97,** 14156-14161.

28. Laurent, T. C. & Blombäck, B. (1958) *Acta Chem. Scand.* **12,** 1875-1877.

29. Hantgan, R. & Hermans, J. (1979) *J. Biol. Chem.* **254,** 11272-11281.