Crystal structure of the central region of bovine fibrinogen (E_5 fragment) at 1.4-Å resolution

Joel Madrazo*†, Jerry H. Brown*‡, Sergei Litvinovich†‡§, Roberto Dominguez¶, Sergei Yakovlev†, Leonid Medved†‖, and Carolyn Cohen*‖

*Rosenstiel Basic Medical Sciences Research Center, Brandeis University, Waltham, MA 02454-9110; †Department of Biochemistry, The American Red Cross Holland Laboratory, Rockville, MD 20855; and ¶Boston Biomedical Research Institute, Watertown, MA 02472

Contributed by Carolyn Cohen, August 20, 2001

The high-resolution crystal structure of the N-terminal central region of bovine fibrinogen (a 35-kDa E_5 fragment) reveals a remarkable dimeric design. The two halves of the molecule bond together at the center in an extensive molecular ''handshake'' by using both disulfide linkages and noncovalent contacts. On one face of the fragment, the Aα and Bβ chains from the two monomers form a funnel-shaped domain with an unusual hydrophobic cavity; here, on each of the two outer sides there appears to be a binding site for thrombin. On the opposite face, the N-terminal γ chains fold into a separate domain. Despite the chemical identity of the two halves of fibrinogen, an unusual pair of adjacent disulfide bonds locally constrain the two γ chains to adopt different conformations. The striking asymmetry of this domain may promote the known supercoiling of the protofibrils in fibrin. This information on the detailed topology of the E_5 fragment permits the construction of a more detailed model than previously possible for the critical trimolecular junction of the protofibril in fibrin.

Fibrinogen, the key structural protein in blood clotting, has a unique and complex dimeric structure: the central so-called "E" region, critical for fibrin formation, contains a nexus of chains that bond the two identical halves of the molecule together in a small globular region (Fig. 1). Each monomer of this large (340-kDa) elongated (450-Å-long) molecule consists of three nonidentical chains, Aα, Bβ, and γ, and the N-terminal portions of the six chains are linked together by 11 disulfide bonds at the center. The C termini of each of the three chains also end in globular domains: those of the Bβ and γ chains are located at the ends, or D regions, and those of the Aα chains, the αC domains, appear to interact with each other close to the central E region. Except for an extended flexible portion of the αC domain, the regions between the globular domains in each half-molecule form α-helical coiled-coil structures, so that the E region consists of a globular region with two coiled-coil extensions (for review, see ref. 1).

When clotting occurs, thrombin cleaves two pairs of small negatively charged fibrinopeptides from the central E region, and soluble fibrinogen is converted into a relatively insoluble fibrin molecule, which self assembles to form the clot. In this process, the exposed N-terminal ''knobs'' of the α chains in one molecule of fibrin bind to receptor pockets in the terminal γC domains of adjacent molecules, leading to the formation of two-stranded half-staggered protofibrils (2, 3). The N-terminal knobs of the β chains are also exposed, and interactions between these knobs and receptor pockets in the βC domains may promote assembly of the protofibrils into fibers (4–6). Moreover, release of the fibrinopeptides by thrombin also appears to result in the dissociation of the αC domains from the central E region, and these domains can then promote assembly of protofibrils into fibers (7, 8). Lateral association of the protofibrils that produce thicker fibers does not appear to be as regular as the association of filaments within the two-stranded protofibril itself (9), and the specific lateral contacts made are only partially identified (6). The clot is strengthened by the covalent crosslinking of the end-to-end bonded γC domains (and, at a slower rate, of the αC domains). The clot is later dissolved by the action of plasmin, which produces various fragments of the molecule, including the D–D dimer and the E fragment.

The central E region has previously been visualized only at very low resolution in electron microscopic (10–14) or medium resolution by x-ray crystallographic (15, 16) studies of the proteolytically truncated or native molecules. At best, these data limit reliable chain traces to the relatively simple coiled-coil portions of the E region. To visualize the central region of the molecule in atomic detail, we have prepared a 35-kDa E_5 fragment and determined its crystal structure to 1.4-Å resolution (see *Methods*). The results reveal that this chemical homodimer is conformationally asymmetric and consists of a strongly linked, highly convoluted interface between the two halves; they supplement as well our current picture of the molecular packing in the fibrin clot.

Methods

Protein Preparation and Characterization. A 45-kDa fragment E was purified from a 2-h plasmic digest of bovine fibrinogen, as described previously (17, 18). This fragment was denoted as E_3 because its N termini (AαLeu-23 or AαGln-27, BβLys-61, and γ Tyr-1) were equivalent to those reported for the human fibrinogen E_3 fragment (19). Further digestion of bovine E_3 fragment with chymotrypsin 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*). The complete amino acid sequence of bovine E_5 was determined (Fig. 1*c*), and a comparison of this sequence with that of human fibrinogen E_3 reveals that the generation of E_5 is due to the removal of the N-terminal residues 23–28 from the Aα chain and several residues from the C termini of the Bβ and γ chains. The substantial decrease in the molecular mass as E_3 is degraded to E_5 can be attributed mainly to the stepwise removal of the carbohydrates linked to Asn-52 of each γ chain.

Crystallization and Structure Determination. E_5 was crystallized in two space groups: $P2_1$ (a = 49.4 Å, b = 66.2 Å, c = 50.7 Å, β = 106.6°) and $P2_12_12_1$ (a = 53.4 Å, b = 58.8 Å, c = 96.8 Å). To ensure that the E_5 fragment was not modified during crystallization, several (washed) crystals were analyzed: the N termini in the crystals were the same as in the starting material, and no free sulfhydryls were detected. By using x-ray diffraction data collected both on an R-AXIS IV detector system mounted on a Rigaku (Tokyo) x-ray generator and at synchrotron facilities (Table 1), the structure of the orthorhombic crystal form was solved (to 1.6-Å resolution) by a combination of single isomorphous replacement (using a trimethyllead-acetate derivative) and density modification. In this process, the α-helices of the coiled-coil regions were first

Data deposition: The atomic coordinates have been deposited in the Protein Data Bank, www.rcsb.org (PDB ID codes 1JY2 and 1JY3).

‡J.H.B. and S.L. contributed equally to this work.

§Present address: The American Chemical Society, 1155 16th Street N.W., Washington, DC 20036.

‖To whom reprint requests should be addressed. E-mail: ccohen@brandeis.edu or medvedl@usa.redcross.org.

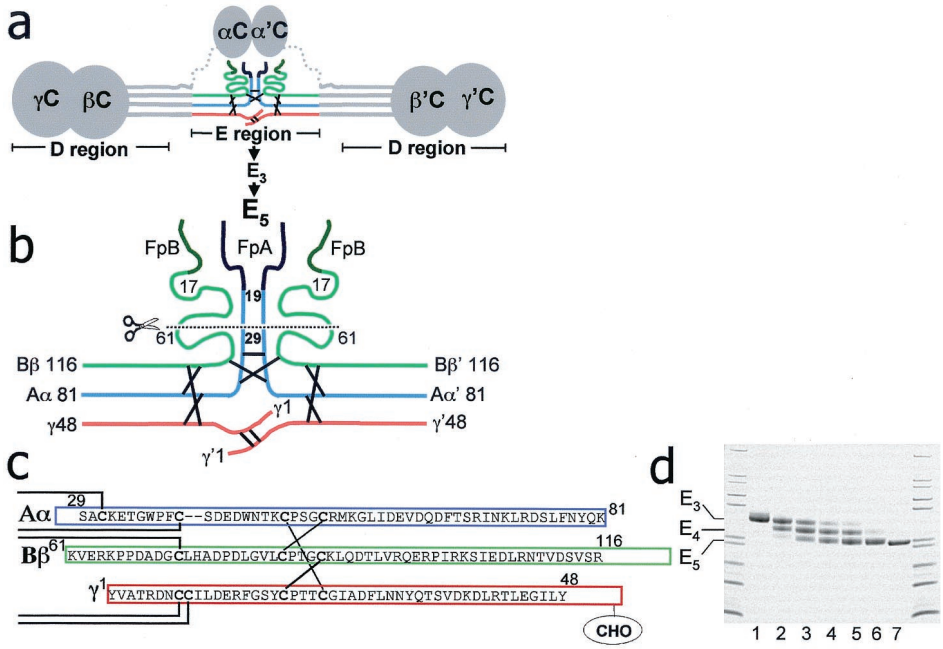

Fig. 1. (*a*) Schematic diagram of the intact fibrinogen dimer highlighting (in color) the central location of the E region. Here, the N-terminal regions of the Aα (blue), Bβ (green), and γ (red) chains from the two halves of the molecule are covalently linked by 11 disulfide bonds (black lines). The C-terminal regions of the chains form globular domains (depicted by circles). Coiled coils are depicted by parallel lines, and disordered segments are dotted. (*b*) Schematic diagram of the central E region showing the N-terminal portions of the Aα and Bβ chains cleaved sequentially by plasmin and chymotrypsin to generate E_5 (below scissors) (see *Methods*). This fragment is missing both the fibrinopeptides (FpA and FpB, shaded dark) and the thrombin-exposed polymerization knobs of fibrin (α19–22 and β17–20) but includes all of the central region's disulfide bonds. (*c*) Amino acid sequence of one half of the E_5 dimer indicates that its generation from E_3 (whose approximate length is represented by the solid rectangles) is due to the removal of several residues from the N termini of the Aα chains and the C termini of the Bβ and γ chains, including the γAsn-52-linked carbohydrates (CHO) (see *Methods*). (*d*) Time course of the chymotryptic digestion of the 45-kDa bovine fibrinogen fragment E_3 reveals that an intermediate 40-kDa "E_4" fragment is produced before the appearance of the 35-kDa E_5 product. The outer lanes are molecular mass markers; lanes 1–6 are the E_3 fragment before and 1, 2, 4, 6, and 10 h after addition of chymotrypsin; lane 7 is purified E_5. (See digestion conditions in supporting information, www.pnas.org.)

identified in the electron density maps; their structures were subsequently used to obtain calculated phases that, when combined with the experimental data, led to completing the chain traces for the less regular central two domains of the fragment. Using this model, phases for the monoclinic crystal form were determined by molecular replacement, and the structure was refined to 1.4-Å resolution. The structure of E_5 is very similar in the two space groups, except for small differences in the coiled-coil domain (see below). The N-terminal residues Aα29–34, Bβ61–63, and γ1 and the C-terminal residues Aα79–81 and Bβ115–116 were disordered in both halves of the molecule in each of the two crystal forms. Refinement statistics are found in Table 2. Additional protein preparation, purification, and crystallographic methods are published as supporting information on the PNAS web site (www.pnas.org).

Results

Overall Domain Structure. The dimeric rod-shaped E_5 fragment may be divided into four closely interacting but distinct domains (Fig. 2), two of which are α-helical coiled-coil segments that extend along the long axis of the fragment. The N-terminal ends of the two coiled coils can be seen in the high-resolution E_5 structure to be separated at the center by only 7 Å; the precise N-terminal residues of these coiled coils are AαSer-50, BβThr-85, and γThr-21. Here, the three chains within each monomer can be seen to be covalently linked by the previously predicted ring of three disulfide bonds, Aα48-γ23, γ19-Bβ87, and Bβ83-Aα52 (16, 20–22) (Fig. 3). The remaining N-terminal portions of all six chains are located for the most part surrounding, rather than between, the coiled coils (15) and form two additional domains in the central part of E_5. In contrast to the coiled-coil domains, each of these two central domains includes

Table 1. Data collection statistics

Crystal	$P2_1$ (native)	$P2_12_12_1$ (native)	$P2_12_12_1$ (native)	$P2_12_12_1$ (derivative)†	$P2_12_12_1$ (derivative)†
X-ray source	CHESS	Rigaku	NSLS	Rigaku	NSLS
Wavelength, Å	0.91	1.54	1.087	1.54	1.087
Resolution, Å	1.4	2.29	1.6	2.3	2.3
Unique reflections	61,149	14,312	41,785	14,357	14,332
Total reflections	359,638	115,257	1,321,351	123,405	320,778
Completeness, %	99.2	99.3	99.4	99.0	99.9
R_{sym}* (%)	5.8 (33.4)	8.4 (26.0)	5.9 (14.7)	5.8 (16.3)	9.2 (24.6)

NSLS, National Synchrotron Light Source; CHESS, Cornell High Energy Synchrotion Source.

*$R_{sym} = \Sigma_{hkl} \Sigma_i | I_i - \langle I \rangle | / \Sigma_{hkl} \Sigma_i I$; where $\langle I \rangle$ is the mean intensity of reflection *hkl*. The R_{sym} value for the highest-resolution shell (10% of unique data) is indicated in parentheses.

†The derivative is $Pb(CH_3)_3$ OAc. See *Methods* for additional details.

Table 2. Refinement statistics

Crystal	$P2_1$	$P2_12_12_1$
Resolution range, Å	100.0–1.4	100.0–1.6
No. of protein atoms, waters	2,200, 318	2,229, 312
Number of reflections	58,431	40,453
R_{factor}* (%), R_{free}† (%)	21.6, 23.6	19.4, 22.0
rms bond lengths, Å, angles (°)	0.011, 1.46	0.010, 1.41
rms dihedrals (°), improper (°)	19.4, 1.01	19.7, 1.03
Luzzati coordinate errors, Å	0.20	0.18
Average B values, Å^2	28.8	27.7

*$R_{factor} = \Sigma_{hkl} | |F_{obs}| - |F_{calc}| | / \Sigma_{hkl} |F_{obs}|$; where F_{calc} and F_{obs} are respectively, the calculated and observed structure factor amplitudes for reflections *hkl* included in the refinement. σ cutoff = 0.

†R_{free} is the same as R_{factor} but calculated over a randomly selected fraction (5%) of the reflection data not included in the refinement.

chains from both monomers. The N-terminal portions of the Aα and Bβ chains in these monomers are located approximately on one face of the molecule ("up" in Figs. 2 and 3*a*), where together they form the rims and the walls of a funnel-shaped domain (Fig. 4), which is centered on the molecule's 2-fold axis. The N-terminal portions of both γ chains are located on the opposite face ("down" in Figs. 2 and 3*a*), where they form a distinct region that we call the "γN domain." Four of the five disulfide bonds in these two central domains (Fig. 3*a*) previously predicted to connect the two halves of the molecule (16, 21, 23) are clearly seen in the E_5 electron density maps. These are Aα39-Bβ'72 and Aα'39-Bβ72 in the funnel-shaped domain and γ8-γ'9 and γ'8-γ9 in the γN domain.** [The N-terminal residues Ser-29–Thr-34 of both Aα chains, however, are disordered, and consequently the bond between AαCys-31 and Aα'Cys-31 is not seen, although it is present in the crystals (see *Methods*)]. Numerous apolar and ionic contacts are made between the two coiled-coil domains and the funnel-shaped and γN-domains (as described in supporting information, www.pnas.org), consistent with the early prediction of strong interdomainal interactions within the E fragment from scanning calorimetry studies (17). Nevertheless, the disulfide bonds of the E_5 fragment are either completely intradomainal (i.e., those that connect the two molecular halves in the funnel-shaped or γN domains) or provide a stabilizing cap for the N termini of the coiled coils; this arrangement of covalent linkages is in agreement with the picture that the structure consists of four independently folded domains.

The Coiled-Coil Domains of E_5. 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 (24) (for review, see ref. 25). In the E_5 fragment, there is one three-residue deletion from the heptad repeat of each chain, located at homologous positions (Aα65, Bβ100, and γ36) midway along the coiled-coil domain (Fig. 1*c*). These deletions, or "stutters," result in local non-close-packed cores as found in certain other coiled coils (26, 27). 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, which promote flexibility, may be related to the functions of fibrinogen.

**Note that the prime (') in this manuscript is used exclusively to distinguish one half of the dimeric molecule from the other half; this same notation is also used elsewhere in the fibrinogen literature to refer to the alternatively spliced C terminus of the γ chain.

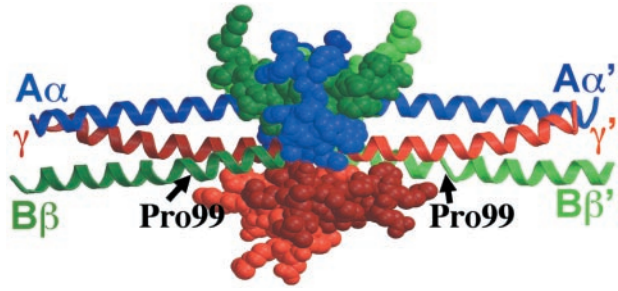

Fig. 2. The two sets of Aα (blue), Bβ (green), and γ (red) chains form four domains in fragment E_5. Each of the coiled-coil domains (ribbons) consists of chains from the same half molecule and are slightly bent at the location of proline Bβ99. The funnel-shaped domain (green and blue space-filling model), composed of the Aα and Bβ chains, and the γN domain (red space filling model) both include chains from the two molecular halves. One monomer is shaded darker than the other.

The Six Chains in the Central Region Produce a Highly Convoluted Dimeric Interface. 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 subunit have distinct structural roles in the formation of the funnel-shaped domain. These segments of the Aα chains are located on the exterior of this domain perpendicular to the long axis of the molecule (Figs. 3*a* and 4). The most N-terminal residues of the Aα chains, Gly-35 and Trp-36, form part of the rim of the domain. The two chains diverge from one another (from residues 36 to 41) and subsequently converge (residues 43 to 49), 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 both coiled-coil domains (Figs. 2 and 4). 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 70–84 of each Bβ chain form a relatively long loop containing a two-stranded antiparallel β-sheet. Residues 78 and 79, near the reverse turn of this loop, interact with the coiled-coil domain of the opposite monomer. One face of each loop is disulfide linked to the Aα chain portion of this domain, whereas the opposite face forms a major portion of the cavity's concave surface. This surface is unusual in being dominated by uncharged and hydrophobic amino acid side chains.

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 also form a loop that interacts with the opposite subunit's coiled coil. 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 are part of a short antiparallel β-sheet and are located just below the funnel-shaped cavity described above.

The four loops formed by the Bβ and γ chains may be pictured as "fingers" that grasp each other and the coiled-coil domains in a firm "handshake" between the two half-molecules (Fig. 5). Such an intertwined structure yields a very large ≈2,500-Å^2 contact area between the two halves of the E_5 fragment (see *Methods*), which is greater than that found in many conventional protein dimers (28). It is apparent that the extensive interactions among the chains produce an exceptionally tight binding between the two molecular halves. In fact, the intertwined conformation of E_5 is reminiscent of "3-D domain-swapped" dimers described by Eisenberg and colleagues (for review, see ref. 29), where two or more protein chains exchange identical elements

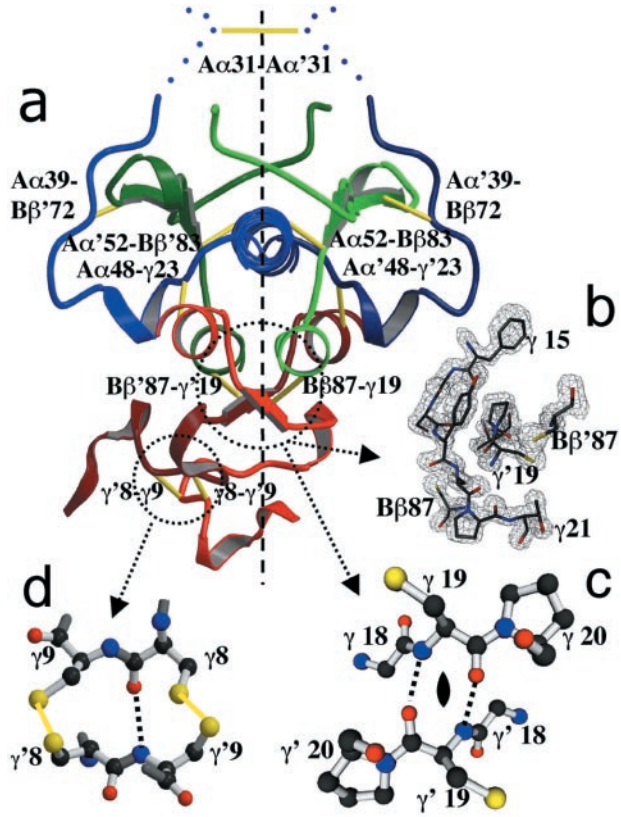

Fig. 3. Locations and conformational effects of disulfide bonds in E_5. (*a*) Cross section of most of the fragment (C-terminal portions of the coiled-coil domains are omitted) viewed along the long axis of the molecule. This view shows that the disulfide bonds (yellow), which brace the N-terminal end of each coiled-coil domain, are located in the interior of the structure relative to the disulfide bonds that connect the two molecular halves. The color coding is the same as in Fig. 2. The N-terminal portion of the γN domain (*Bottom*), which includes the two disulfide bonds between residues 8 and 9 of opposite γ chains (see *d*), is located to the side of the fragment's 2-fold dimeric axis (dashed line). Residues 29–34 of the two Aα chains (*Top*), including the disulfide bond that links them at residue 31, are poorly ordered in the fragment, and their locations (bold dotted lines) are approximate. (*b* and *c*) Magnified views of the symmetrical C-terminal portion of the γN domain. *b* includes residues γ15–21 and γ′19–21, as well as residues Bβ87 and Bβ′87, to which the γ and γ′ chains, respectively, are disulfide linked; also shown is the corresponding 1.6-Å-resolution electron density map produced by using $2F_o - F_c$ coefficients and phases calculated from an E_5 model omitting these residues. High resolution is required to distinguish the closely interacting halves of the fragment. *c* shows the short antiparallel β-sheet in this region; this structure is compatible with the 2-fold axis of the dimer (oval symbol), because identical residues (γ19 and γ′19) in the sheet are located directly opposite each other. Note, for example, the identical chemical environments of the carbonyl oxygens of residues γ19 and γ′19. (The side chains of residues γ18 and γ′18 are omitted for clarity.) (*d*) Atomic model of the disulfide bonds in the N-terminal portion of the γN domain, which locally orient the two γ chains in an antiparallel but asymmetric manner. The covalent linkage of residue 8 from one chain with residue 9 of the other prevents the register between the chains necessary for 2-fold symmetry. As a result, the hydrogen-bonding pattern is different for the two chains. Note here the different chemical environments of the carbonyl oxygens of residues γ8 and γ′8. The conformations of the N-terminal 14 residues of the two γ chains differ from one another.

to form a strongly bound oligomer. These many linkages would be expected to hold the two fibrinogen monomers together even without the intersubunit disulfide bonds.

Fibrinogen Is a Conformational Heterodimer. One of the striking features of the E_5 structure is that only part of the dimeric interface is symmetric. Residues 1–14 of the γ and γ′ chains are

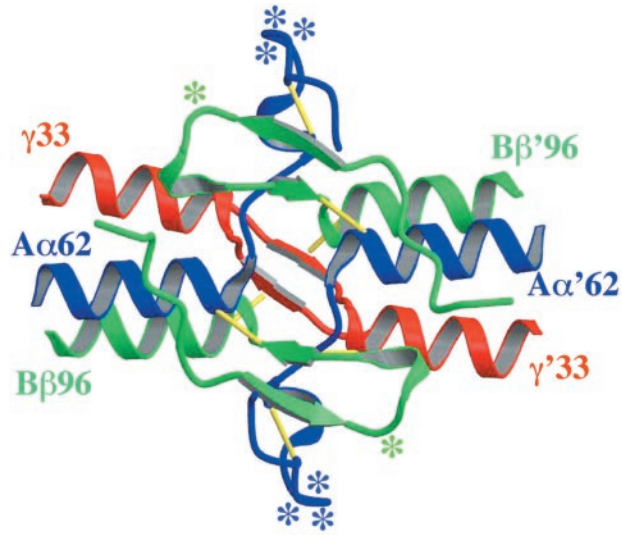

Fig. 4. Fragment E_5 has an unusual funnel-shaped domain with an apolar cavity. The relatively large rims and walls of this cavity are formed by the Aα (blue) and Bβ (green) chains; the small floor is formed by the C-terminal portion of the γN domain (red). Possible binding sites for thrombin (stars) are located predominantly on the rims and exterior sides of the walls of this domain. In this view, the γN domain and the C-terminal portions of the coiled coils are omitted for clarity.

for the most part positioned to the side of the 2-fold axis of the molecule (Fig. 3*a*). Moreover, this portion of the γN domain is itself asymmetric, because the two chemically identical polypeptide segments adopt different conformations. This unusual feature is caused by the two reciprocal disulfide bonds between residues 8 and 9 of the γ and γ′ chains, which, as determined previously (30), locally orient these chains in an antiparallel manner. This design can be accounted for by recognizing that when two antiparallel β-strands are related by a 2-fold symmetric axis (as are residues γ18–20 and γ′18–20; see Fig. 3*c*), the side chains directly across from each other that are closest to the axis (such as γ19 and γ′19) must be the same. This type of register cannot occur in the N-terminal part of the γN domain, however, where residue 8 from one γ chain is covalently linked to residue 9 on the other (Fig. 3*d*). In this arrangement, the main-chain carbonyl oxygen of residue γ8 forms a hydrogen bond with the main-chain nitrogen of γ′9, but the main-chain carbonyl of γ′8 points away from the opposite subunit and forms a hydrogen bond with γ′11, resulting in markedly different conformations

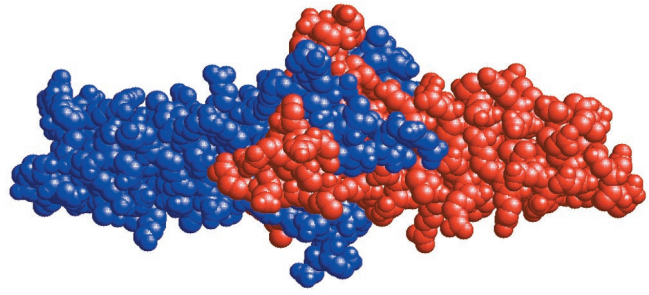

Fig. 5. The two halves of fragment E_5 (shown in red and blue) form an intertwined dimer and contain many stabilizing contacts within a relatively small region. The convoluted nature of the dimeric interface results primarily from the pairs of N-terminal Bβ and γ chains (best seen in Figs. 4 and 2, respectively). These chains appear as if they had exchanged, in evolution, identical elements between the two halves of the molecule, and the structure is reminiscent of ''3-D domain-swapped'' dimers (see text and ref. 29).

for the N-terminal 14 residues of the two chains (Fig. 3*a*). Moreover, this region contacts the stutter segment of the coiled-coil domain, which can bend differently in the two halves of the molecule. This heterodimeric structure of the E region reveals that fibrinogen, in a formal sense, is a polar molecule with respect to its long axis. This asymmetry may be significant for the formation of twisted protofibrils and fibers in fibrin (see below).

Discussion

Comparison to Previous Structures. The crystallographic analyses of fibrinogen and its fragments have been pursued for a number of decades, but only in the past 5 years have atomic or near-atomic resolution results been achieved. The resolution of the E_5 fragment crystal structure, at 1.4 Å, is the highest of any portion of fibrinogen or fibrin determined to date. Previously, structures of C-terminal fragments have also been determined to fairly high resolution (2.1–2.9 Å): these include the human γC domain (31) and fragments D and D–D (32–34). These results were critical for phasing the medium- to low-resolution data (4.0 and 5.5 Å) obtained, respectively, from crystals of the proteolytically truncated bovine (15) and intact chicken (16) fibrinogen molecules. Chain traces for the E region were also reported in the chicken fibrinogen molecule, but the conformations of the noncoiled-coil domains are significantly different from that of the bovine E_5 fragment described here. We believe that these discrepancies, especially in the highly conserved γN chains, arise from the overinterpretation of the 5.5-Å resolution map of chicken fibrinogen.

Central Role for the E Region in Fibrin. The structure of the E_5 fragment of bovine fibrinogen provides information on the binding site for thrombin and the topology of the fibrin clot. The removal of Aα-chain fibrinopeptides A in fibrinogen by thrombin creates the N-terminal α knobs, consisting of gly-pro-arg (GPR) residues at positions 19–21 of the α chains, that fit into receptor pockets in the C-terminal γ domains during self assembly. Similarly, removal of Bβ-chain fibrinopeptides B creates the gly-his-arg (GHR) β knobs at positions 15–17 that appear to fit into pockets in the C-terminal β domains (see ref. 6). These residues, however, as well as additional parts of the native E region implicated in binding to thrombin [i.e., Bβ22–49 (35)], have been removed in the preparation of the E_5 fragment (see *Methods* and Fig. 1). Nevertheless, because much of the central region has now been traced, we can locate the significant remaining portion of the thrombin-binding site; we can also estimate the general locations of the thrombin-exposed polymerization knobs more closely than previously (15) and thus improve the current model of the molecular packing in fibrin.

The Thrombin-Binding Sites in E_5 Are Located on the Funnel-Shaped Domain. Results from various studies on native and abnormal human fibrinogens indicate that the chain segments present in the E_5 fragment implicated in the binding of thrombin include (in bovine numbering) Aα38–46 (35, 36) and Bβ75 (37). In each half of the E_5 structure, these residues are located on the outer wall of the funnel-shaped domain and are positioned adjacent to one another, as one would expect for a composite binding site. Of the five surface-oriented residues in this Aα-chain segment (at positions 38, 40, 41, 42, and 45), Phe-38 may be of special note: this residue is closest to the rim of the funnel (in the presumed direction of the Aα knobs; see below), is the only apolar side chain (and is highly solvent accessible in E_5), and is nearest Ala-75 of the Bβ chain. The binding site for thrombin does not appear to extend beyond Bβ75 (to the portion of the Bβ chain near the coiled-coil domains), because site-directed mutagenesis studies show that substitutions of Pro-77 and Leu-79 do not affect the kinetics of fibrinopeptide release by thrombin (38). The dimeric E_5 structure also shows that the two composite

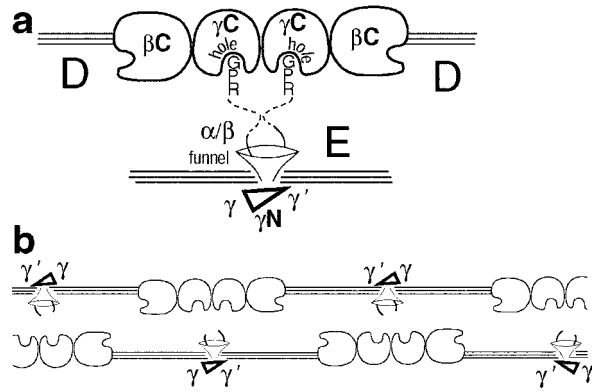

Fig. 6. Schematic model of the basic protofibrillar unit of polymerized fibrin (3, 15, 32) (which consists of two half-staggered filaments of end-to-end bonded fibrin molecules), now taking into account the domain structure of bovine E_5. (*a*) In this model of the DDE interface, the γC domain receptor pockets (holes) for the N-terminal Aα knobs and the funnel-shaped domain face each other. The γN domain is located on the exterior of the protofibril (see text). [Note that, for simplicity, the offset between γC domains at the D–D interface (32) is not shown, and the distance between the two filaments is relatively arbitrary.] (*b*) If the γN domain does not influence the formation of individual protofibrils (as its location in this protofibril model suggests), then fibrin molecules may not pack regularly with respect to the asymmetry in this domain. The exterior of the protofibril would then not display a regular repeat, a feature that could affect subsequent lateral associations between protofibrils. (Additional nonuniformity may also arise from the offset at the D–D interface.)

thrombin-binding sites on the two molecular halves are well separated, by ≈35 Å, on opposite sides of the funnel-shaped domain (Fig. 4), such that two thrombin molecules can be accommodated simultaneously, in agreement with the measured stoichiometry (39).

The Locations and Functions of the Central Domains in the Protofibril of Fibrin. 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 (15) 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 (dotted lines in Fig. 3*a*). The two AαGPR 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*). 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 so that it might influence associations between protofibrils (see below).

Implications of γN-Domain Asymmetry. The unusual structure of this γN domain illustrates the special role of a dimeric interface in protein folding and suggests how certain topological features of the fibrin fiber may be generated. Identical polypeptide chains generally fold into identical conformations. When two such chains are in a dimer, however, the interactions at their interface may have important conformational effects (40). Two of the most common dimerization motifs found in proteins—the antiparallel β-sheet and the parallel α-helical coiled-coil—generally form symmetrical structures with identically folded halves. It has

recently been shown, however, that core alanines in the tropomyosin coiled coil disrupt the in-register symmetrical alignment of the chemically identical α-helices (41). In fragment E_5, we now see how certain covalent disulfide bridges break the symmetry between two chemically identical antiparallel β-strand-like chains and, together with a number of noncovalent interactions, produce a compact asymmetric domain at the center of the molecule (see *Results*).

A functional role for the γN domain has not yet been established. Nevertheless, the asymmetry of this domain, which contacts the coiled-coil domains, may play a role in fibrin assembly. Molecules, protofibrils, and fibers of fibrin appear to display twisted conformations (9, 13, 42), and this asymmetry may be one feature that gives rise to the twisting. Moreover, the (probable) location of such a symmetry-breaking domain on the outside of the two-stranded protofibril (described above) may affect the uniformity of fibrin packing. It does not appear that the asymmetry would be propagated all the way to the positions of the α and β knobs, and the most direct implication of the location of the γN domain is that it should have little influence on the formation of individual protofibrils. If this were indeed the case, then successive fibrin molecules along the protofibril would not be arranged in a regular way with respect to this asymmetry (i.e., for some molecules γ would be situated to the "left" and γ' to the "right;" for other molecules, the polarity could be reversed), and the exterior surfaces of any two extended protofibrils would generally differ (Fig. 6*b*). In this way, the asymmetry in the γN domain could contribute to disorder seen in the side-to-side packing between protofibrils (9, 43), depending on the actual contacts this domain may make.

Tying up Loose Ends. With the crystallographic analysis of E_5 in hand, the conformation of the entire "backbone" of the fibrinogen molecule has now been reliably described. However, the precise locations and structures of those residues that extend away from the molecule's long axis, including the αC domains and the N-terminal residues of the Aα and Bβ chains adjacent to the funnel-shaped domain, have yet to be determined; these segments have either been excised or are relatively disordered in the various crystallographic studies. Moreover, the function of the unusual largely apolar cavity in the funnel-shaped domain of E_5 is unknown, but this region may turn out to interact with the adjacent N-terminal Aα- and/or Bβ-chain segments. The αC domains [poorly visualized in the native chicken fibrinogen structure (16)] appear to associate with each other to form a dimeric domain that is also located near the center of the E region in fibrinogen (7, 11, 12), but this binding appears to be most strongly mediated by the Bβ-chain segment that is released on thrombin cleavage (8). Knowledge of the functions of the apolar cavity thus awaits further studies and may require, for example, a high-resolution structure of a more complete E region fragment of fibrinogen. Even more revealing would be the structure of the trimeric DDE complex isolated from the fibrin clot. Knowledge of the detailed interactions in this complex will also test the basic protofibrillar model, including the relative orientations and possible functional roles of all four domains described here for E_5.

We thank K. Ingham, S. Lord, and J. Weisel for critical reading of the manuscript, and M. Love, D. Himmel, and the staffs of the Brookhaven National Laboratory and the Cornell High Energy Synchrotron Source for assistance with data collection. This work was supported by grants from the National Institutes of Health (AR17346 to C.C. and HL-56051 to L.M.).

1. Doolittle, R. F., Everse, S. J. & Spraggon, G. (1996) *FASEB J.* **10,** 1464–1470.
2. Stryer, L., Cohen, C. & Langridge, R. (1963) *Nature (London)* **197,** 793–794.
3. Weisel, J. W., Phillips, G. N., Jr. & Cohen, C. (1983) in *Molecular Biology of Fibrinogen and Fibrin*, eds. Mosesson, M. W. & Doolittle, R. F. (NY Acad. Sci., New York), pp. 367–379.
4. Hantgan, R., McDonagh, J. & Hermans, J. (1983) in *Molecular Biology of Fibrinogen and Fibrin*, eds. Mosesson, M. W. & Doolittle, R. F. (NY Acad. Sci., New York), pp. 344–366.
5. Weisel, J. W., Veklich, Y. & Gorkun, O. (1993) *J. Mol. Biol.* **232,** 285–297.
6. Yang, Z., Mochalkin, I. & Doolittle, R. F. (2000) *Proc. Natl. Acad. Sci. USA* **97,** 14156–14161.
7. Veklich, Y. I., Gorkun, O. V., Medved, L. V., Nieuwenhuizen, W. & Weisel, J. W. (1993) *J. Biol. Chem.* **268,** 13577–13585.
8. Gorkun, O. V., Veklich, Y. I., Medved, L. V., Henschen, A. H. & Weisel, J. W. (1994) *Biochemistry* **33,** 6986–6997.
9. Weisel, J. W., Nagaswami, C. & Makowski, L. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 8991–8995.
10. Hall, C. E.& Slayter, H. S. (1959) *J. Biophys. Biochem. Cytol.* **5,** 11–17.
11. Mosesson, M. W., Hainfeld, J., Wall, J. & Haschemeyer, R. H. (1981) *J. Mol. Biol.* **153,** 695–718.
12. Erickson, H. P. & Fowler, W. E. (1983) in *Molecular Biology of Fibrinogen and Fibrin*, eds. Mosesson, M. W. & Doolittle, R. F. (NY Acad. Sci., New York), pp. 146–163.
13. Williams, R. C. (1983) in *Molecular Biology of Fibrinogen and Fibrin*, eds. Mosesson, M. W. & Doolittle, R. F. (NY Acad. Sci., New York), pp. 180–193.
14. Weisel, J. W., Stauffacher, C. V., Bullitt, E. & Cohen, C. (1985) *Science* **230,** 1388–1391.
15. Brown, J. H., Volkmann, N., Jun, G., Henschen-Edman, A. & Cohen, C. (2000) *Proc. Natl. Acad. Sci. USA* **97,** 85–90.
16. Yang, Z., Mochalkin, I., Veerapandian, L., Riley, M. & Doolittle, R. F. (2000) *Proc. Natl. Acad. Sci. USA* **97,** 3907–3912. (First Published March 28, 2000; 10.1073/pnas.080065697)
17. Privalov, P. L. & Medved, L. V. (1982) *J. Mol. Biol.* **159,** 665–683.
18. Medved, L. V., Platonova, T. N., Litvinovich, S. V. & Lukinova, N. I. (1988) *FEBS Lett.* **232,** 56–60.
19. Olexa, S. A., Budzynski, A. Z., Doolittle, R. F., Cottrell, B. A. & Greene, T. C. (1981) *Biochemistry* **20,** 6139–6145.
20. Henschen, A. & McDonagh, J. (1986) in *Blood Coagulation*, eds. Zwaal, R. F. A. & Hemker, H. C. (Elsevier, Amsterdam), pp. 171–241.
21. Blombäck, B., Hessel, B. & Hogg, D. (1976) *Thromb. Res.* **8,** 639–658.
22. Doolittle, R. F., Goldbaum, D. M. & Doolittle, L. R. (1978) *J. Mol. Biol.* **120,** 311–325.
23. Huang, S., Cao, Z. & Davie, E. W. (1993) *Biochem. Biophys. Res. Commun.* **190,** 488–495.
24. Crick, F. H. C. (1953) *Acta Crystallogr.* **6,** 689–697.
25. Cohen, C. & Parry, D. A. D. (1990) *Proteins Struct. Funct. Genet.* **7,** 1–15.
26. Lupas, A., Müller, S., Goldie, K., Engel, A. M., Engel, A. & Baumeister, W. (1995) *J. Mol. Biol.* **248,** 180–189.
27. Brown, J. H., Cohen, C. & Parry, D. A. D. (1996) *Proteins Struct. Funct. Genet.* **26,** 134–145.
28. Janin, J., Miller, S. & Chothia, C. (1988) *J. Mol. Biol.* **204,** 155–164.
29. Schlunegger, M. P., Bennett, M. J. & Eisenberg, D. (1997) *Adv. Protein Chem.* **50,** 61–122.
30. Hoeprich, P. D. & Doolittle, R. F. (1983) *Biochemistry* **22,** 2049–2055.
31. Yee, V. C., Pratt, K. P., Côte, H. C., Trong, I. L., Chung, D. W., Davie, E. W., Stenkamp, R. E. & Teller, D. C. (1997) *Structure (London)* **5,** 125–138.
32. Spraggon, G., Everse, S. J. & Doolittle, R. F. (1997) *Nature (London)* **389,** 455–462.
33. Everse, S. J., Spraggon, G., Veerapandian, L., Riley, M. & Doolittle, R. F. (1998) *Biochemistry* **37,** 8637–8642.
34. Everse, S. J., Spraggon, G., Veerapandian, L. & Doolittle, R. F. (1999) *Biochemistry* **38,** 2941–2946.
35. Binnie, C. G. & Lord, S. T. (1993) *Blood* **81,** 3186–3192.
36. Stubbs, M. T., Oschkinat, H., Mayr, I., Huber, R., Angliker, H., Stone, S. R. & Bode, W. (1992) *Eur. J. Biochem.* **206,** 187–195.
37. Koopman, J., Haverkate, F., Lord, S. T., Grimbergen, J. & Mannucci, P. M. (1992) *J. Clin. Invest.* **90,** 238–244.
38. Lord, S. T., Strickland, E. & Jaycock, E. (1996) *Biochemistry* **35,** 2342–2348.
39. Meh, D. A., Siebenlist, K. R. & Mossesson, M. W. (1996) *J. Biol. Chem.* **271,** 23121–23125.
40. Goodsell, D. S. & Olson, A. J. (2000) *Annu. Rev. Biophys. Biomol. Struct.* **29,** 105–153.
41. Brown, J. H., Kim, K.-H., Jun, G., Greenfield, N. J., Dominguez, R., Volkmann, N., Hitchcock-DeGregori, S. E. & Cohen, C. (2001) *Proc. Natl. Acad. Sci. USA*, **98,** 8496–8501. (First Published July 3, 2001; 10.1073/pnas.131219198).
42. Medved, L. V., Ugarova, T., Veklich, Y., Lukinova, N. & Weisel, J. (1990) *J. Mol. Biol.* **216,** 503–509.
43. Voter, W. A., Lucaveche, C., Blaurock, A. E. & Erickson, H. P. (1986) *Biopolymers* **25,** 2359–2373.