Crystal structure of a recombinant α_EC domain from human fibrinogen-420

(x-ray structure/blood clotting/protein evolution)

GLEN SPRAGGON*, DIANNE APPLEGATE†, STEPHEN J. EVERSE*, JIAN-ZHONG ZHANG†, LEELA VEERAPANDIAN*, COLVIN REDMAN†, RUSSELL F. DOOLITTLE*‡, AND GERD GRIENINGER†‡

*Center for Molecular Genetics, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0634; and †Lindsley F. Kimball Research Institute of the New York Blood Center, 310 East 67 Street, New York, NY 10021

ABSTRACT The crystal structure of a recombinant α_EC domain from human fibrinogen-420 has been determined at a resolution of 2.1 Å. The protein, which corresponds to the carboxyl domain of the α_E chain, was expressed in and purified from *Pichia pastoris* cells. Felicitously, during crystallization an amino-terminal segment was removed, apparently by a contaminating protease, allowing the 201-residue remaining parent body to crystallize. An x-ray structure was determined by molecular replacement. The electron density was clearly defined, partly as a result of averaging made possible by there being eight molecules in the asymmetric unit related by noncrystallographic symmetry (P1 space group). Virtually all of an asparagine-linked sugar cluster is present. Comparison with structures of the β- and γ-chain carboxyl domains of human fibrinogen revealed that the binding cleft is essentially neutral and should not bind Gly-Pro-Arg or Gly-His-Arg peptides of the sort bound by those other domains. Nonetheless, the cleft is clearly evident, and the possibility of binding a carbohydrate ligand like sialic acid has been considered.

Fibrinogen is a protein found in the blood plasma that is the precursor of the fibrin clot. In all vertebrates examined, the protein is composed of two copies each of α, β, and γ subunits. The β and γ chains are homologous throughout their full lengths, but ordinarily α chains are homologous with β and γ chains over the course of their first 200 residues only, their carboxy-terminal regions having been displaced by some kind of genetic exchange at an early point in vertebrate evolution (1). There are minor forms of fibrinogen, however, in which α chains have carboxyl-terminal domains that are homologous to those found at the carboxyl-termini of β and γ chains. In humans, this minor form is referred to as fibrinogen-420, a reflection of an increase in molecular weight from 340 kDa to 420 kDa as a function of the alternative extended α (α_E) chains (2).

These alternative forms of fibrinogen were first brought to light by the discovery of a bipartite messenger RNA for the α chain in chickens (3). Thus, a cDNA was isolated that encoded the major form of the α chain but also included a second ORF downstream; the second ORF was clearly homologous to the carboxyl domains of β and γ chains. A comparable situation was found to exist in human α-chain transcripts, and it was shown that alternative splicing led to an α_E chain; the larger molecular weight fibrinogen containing these α_E chains was found to be synthesized by hepatoma cells (4). Meanwhile, a similar protein was uncovered in lampreys, except that the major and minor forms were encoded by separate genes (5).

Sequence-based phylogenetic trees showed that the minor form α_EC domains diverged only slightly before the divergence of β and γ chain domains (5).§

The fact that in lamprey the two genes encode different amino-terminal sequences made it possible to measure the minor form in plasma simply by quantifying its unique fibrinopeptide; the minor form accounted for 5–10% of the circulating fibrinogen molecules (7). In human plasma, fibrinogen-420 was found to account for ≈1–2% of all fibrinogen molecules (2). The function of these minor forms of fibrinogen remains unknown, although in humans the relative concentration is significantly elevated in cord blood (8).

In lampreys, incorporation of the alternative form into fibrin is rapid, and the carboxyl domain has been implicated in fibrin crosslinking (9). In contrast, the human equivalent, in the form of a recombinant domain, is not crosslinked by factor XIII (D.A., unpublished data). In lampreys, the intact domain is easily removed from fibrinogen by light trypsin action, but the isolated domain does not bind to Gly-Pro-Arg columns the way γ-chain domains do (9). The corresponding recombinant α_EC domain from humans does not bind to these columns either (D.A. and G.G., unpublished data).

The successful expression of relatively large amounts of the recombinant human protein in the yeast *Pichia pastoris* (D.A. and G.G., unpublished data) made it feasible to consider crystallization. It was hoped that a crystal structure might shed light both on the function of the α_EC domain in particular and the evolution of fibrinogen in general, especially in light of recently published structures of the γC domain (10) and fragment D (11) from human fibrinogen.

EXPERIMENTAL PROCEDURES

Expression and Purification. The recombinant domain, designated rα_EC, was expressed in the yeast *P. pastoris* and purified by chromatography on a mono-Q column (Pharmacia). Full details and characterization are described elsewhere (D.A. and G.G., unpublished data). Briefly put, a segment encoding residues Val-610–Gln-847 of the human α_E chain was inserted into a pPIC9 vector and expressed according to the manufacturer's protocol (Invitrogen). The purified product was homogeneous on SDS gels. Amino-terminal sequencing revealed six residues of the vector upstream of Val-610, increasing the calculated mass of the fusion protein to 27,653.

Additional mass is contributed by a cluster of carbohydrate (2). The final product was dialyzed against 0.02 M Tris buffer, pH 7.5 and shipped to San Diego for crystallization.

Crystallography. Crystals were obtained by vapor diffusion from sitting drops at room temperature. A large range of conditions were explored, including a partial factorial approach with precipitants from a Hampton Research screening kit (12). After numerous attempts, the final conditions used in the well were: 20% polyethylene glycol 3350, 0.15 M CaCl$_2$, 0.1 M imidazole-acetate (pH 5.5), and 0.002 M sodium azide; the concentration of the protein solution was 8 mg/ml 0.02 M Tris, pH 7.5. Drops were set with 5 μl of well solution and 5 μl of protein solution. These drops were streak-seeded from a long series of seedings with low-quality crystals, including needle clusters, that had grown spontaneously. Some of these serial seedings were conducted only after drops were set and pre-incubated for 7–10 days. In the end, a number of good, small crystals and two large crystals were obtained (Fig. 1).

Data were collected at Brookhaven National Synchrotron Light Source beamline X12C. The crystals diffracted to a minimum Bragg spacing of better than 2.1 Å at room temperature (Table 1). Analysis of the data with Denzo and Scalepack (13) showed that the crystals belong to space group P1 with unit cell 71.40 Å, 105.71 Å, 71.35 Å, 104.63°, 108.91°, and 71.53°, although there was an aspect of pseudo-symmetry reminiscent of the space group C2. When the postrefined unit cell dimensions were analyzed by the Collaborative Computing Project No. 4 program TRACER (14), no higher symmetry cell compatible with the data was revealed. If a solvent content of 50% were assumed, this would amount to eight molecules in the unit cell. A native Patterson synthesis of the reduced data showed a large peak (50% of origin) at $x = 0.5$, $y = 0.0$, and $z = 0.0$ (in fractional coordinates). The Patterson suggested that only four unique orientations needed to be derived, the other four molecules being related to the first set by a translation of one-half the unit cell in the x direction. Self-rotation functions produced with POLARRFN (14) with an integration radius of 20 Å on all data between 8.0 and 4.0 Å showed large two-fold peaks around the x, y, and z axis, respectively.

Molecular replacement was carried out with XPLOR (15) with the coordinates of a γC domain taken from a 2.3 Å structure of crosslinked fragments D (16) as a search model. A rotation function followed by PC (Patterson Correlation) refinement (17) on all data between 8.0 and 4.0 Å (integration radius 20 Å) produced four peaks, each related by the self rotation angles. The correlation coefficients of the four peaks were 11.8%, 9.7%, 7.6%, and 7.7%. The highest rotational solution was then used to fix the origin of the P1 cell; the other

Table 1. Data collection and refinement statistics

Data statistics	
Space group	P1
Unit cell dimensions	a = 71.40, b = 105.71, c = 71.35,
	α = 104.63, β = 108.91,
	γ = 71.53
Molecules/asym. unit	8
No. of crystals	3
Resolution, Å	2.1
Observations, N	1,169,328
Unique reflections, N	110,459
Mean redundancy	5.2
Completeness, %	91.2
R_{sym} (I), %	9.8
Refinement statistics	
Resolution range, Å	20.0–2.1
No. of model atoms	12,896
No. water molecules	200
R-value[†]	0.195
Free R-value[‡]	0.255
rmsd from ideals	
Bond length, Å	0.013
Bond angle, °	2.5
Average B value	33.3

*$R_{sym} = (\Sigma|I - \langle I \rangle|)/(\Sigma|I|)$.
[†]Crystallographic R-value $(\Sigma\|F_{(Obs)}| - |F_{(Calc)}\|)/(\Sigma|F_{(Obs)}|))$ with 95% of the native data for refinement.
[‡]Free R-value: R-value based on 5% of the native data withheld from refinement.

three solutions were searched over the whole unit cell, producing solutions with correlation coefficients of 18.3%, 14.7%, and 14.5%, at 9.7, 6.5, and 6.5 SD above the mean, respectively. A combination of all four solutions together produced a correlation coefficient of 23.2% and an R factor of 0.552. The four molecules were then translated by one-half a unit cell in

FIG. 1.   Crystal of rα_EC domain used for x-ray diffraction study. The crystal measured $\approx$0.5 mm on its longest edge.
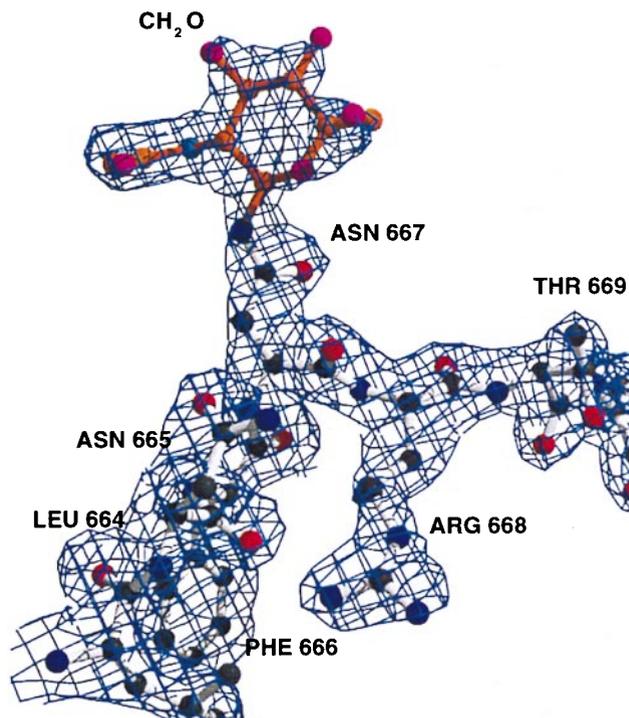
FIG. 2.   Average electron density map showing segment of chain including asparagine α_E667 and attached sugar residues; map calculated from DM coefficients and phases from the unrefined model contoured at 1.5 σ. Residue numbering based on α_E chain from human fibrinogen (4).
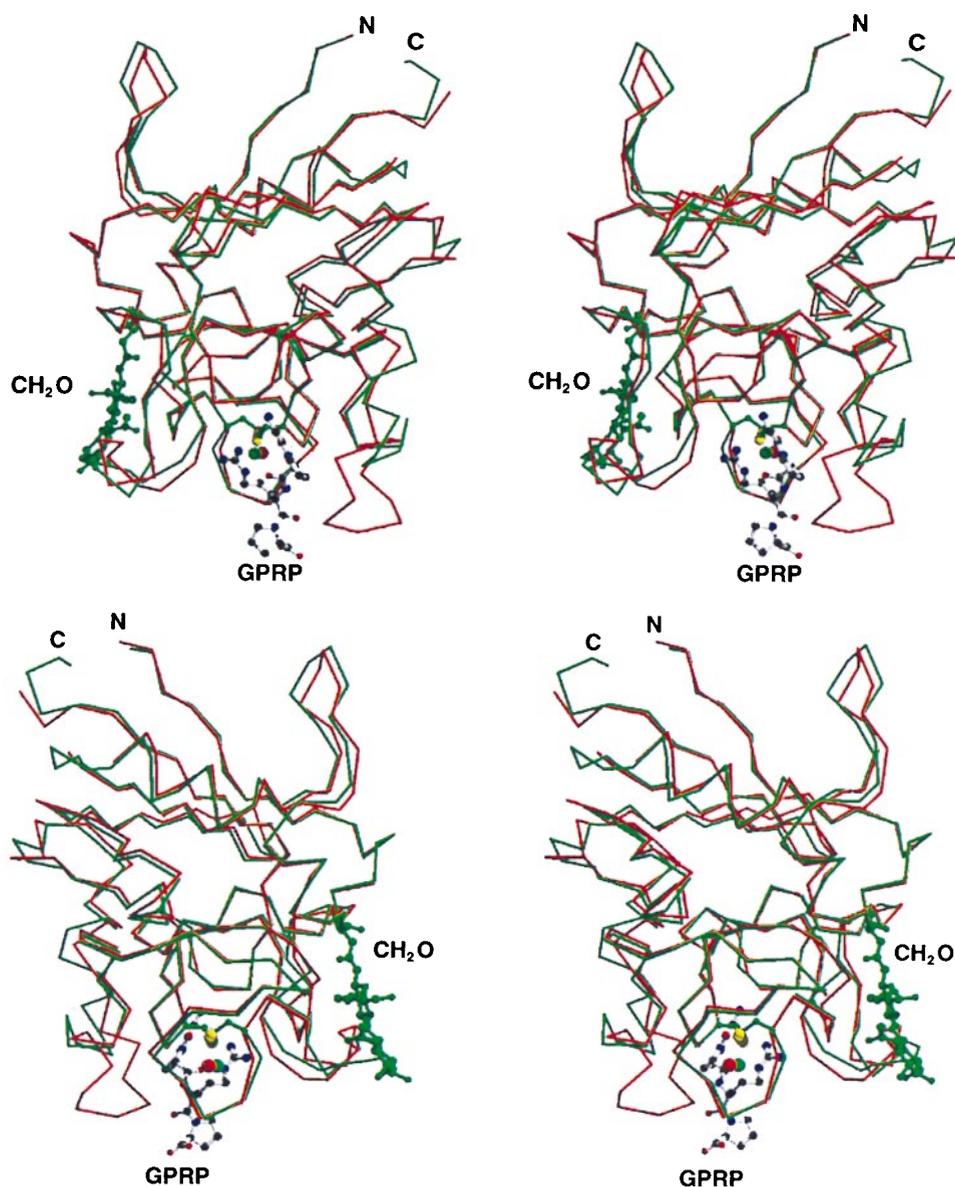
FIG. 3.   Superimposed α-carbon stereo diagrams of α_EC domain (green) and corresponding region of γ chain (red), both from human fibrinogen. The two views are 180° apart. The carbohydrate cluster shown is attached to the Asn at α_E667. The peptide Gly-Pro-Arg-Pro-amide orientation is shown as it binds to the γC domain in fragment double-D (11, 16). Note the equivalent position of calcium atoms (*Upper*, red and green balls behind the ligand). Figure prepared with MOLSCRIPT (31), BOBSCRIPT (32), and RASTER3D (33, 34).

the *x* direction to account for the full volume in the unit cell, and rigid body refinement carried out on the resulting eight molecules. This resulted in a correlation coefficient of 35% and an *R* factor of 0.50 on all data between 8.0 and 2.3 Å.

An analysis of packing of the eight molecules revealed that the first 40–50 residues of each γ domain molecule were clashing severely. Accordingly, the first 50 residues were removed from the model, and rigid body refinement was carried out once more; this time the correlation coefficient rose to 43% and the *R* factor dropped to 0.47. An initial model was built based on residues 190–390 of the γC domain noted above. Sequence differences between the γ and α_E domains were represented by alanines, and loops containing deletions or insertions were omitted. Structure factors for the model were calculated using SFALL (14); figures of merit for the phases were calculated by the SIGMAA-weighting technique of Read (18). The resulting phases were improved by density modification incorporating the 8-fold averaging, histogram matching, and solvent flattening with the program DM (14). The resultant map showed unambiguous electron density for

almost all loops, side chains, and glycosylated residues not included in the initial model, indicating the correctness of the molecular replacement solution (Fig. 2). The model was rebuilt (19) with correct sidechains, 198 of the 201 residues in all, and refined with REFMAC (20). Initially, strong noncrystallographic restraints were applied, as well as phase restraints with the starting DM phases. These were gradually removed as the model improved. In the end, the working *R* factor dropped to 0.195 and the free *R* factor to 0.255.

RESULTS AND DISCUSSION

Crystals. The fact that the first 40–50 residues were missing from the electron density map forced us to compare the chemical makeup of the crystals with the starting protein solutions. It was a fortuitous event that led to long incubations at room temperature before streak-seeding because during that period a 37-residue subdomain (plus six residues from the vector) was proteolytically removed from the starting material. Thus, when crystals were removed from drops, washed with
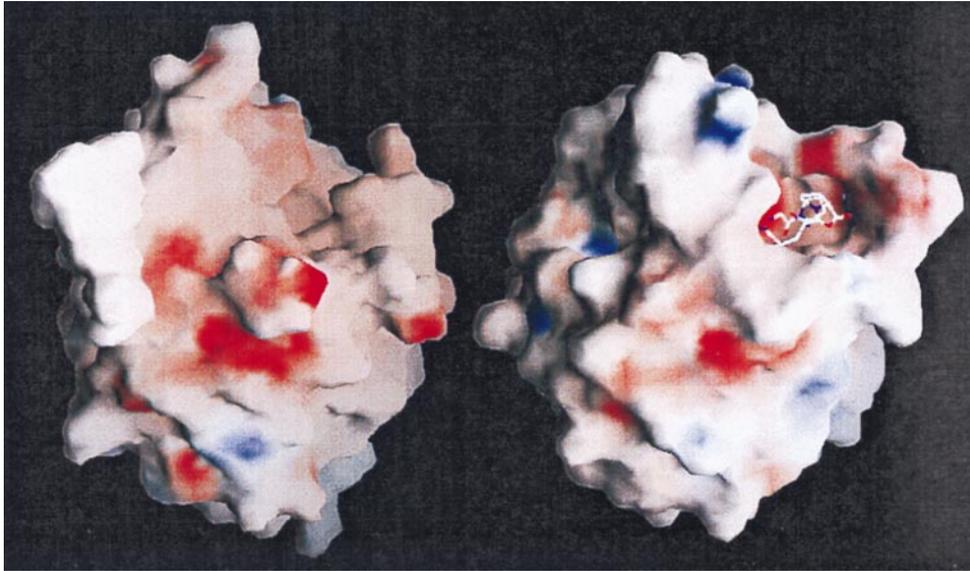
FIG. 4. GRASP (35) depictions of α_EC (*Left*) and γC (*Right*) domains showing equivalent projections with binding clefts. The ball and stick model is liganded Gly-Pro-Arg-Pro-amide. Note lack of negative charge (red) within the binding site of the α_EC chains.

well solution, and examined on SDS gels, the material was significantly smaller (data not shown). Moreover, when harvested crystals themselves were redissolved and examined by amino-terminal sequencing (University of California, San Diego, Sequencing Facility), the sequence began with ETSLG-GWLLI, which corresponds to α_E residues 647–656, indicating

that 37 residues of the construct and the six-amino acid residuum of the vector had been removed by incipient proteolysis.

Structure. As expected, the structure of the α_EC domain closely resembles those of the β and γ domains (Fig. 3). The rms deviation (rsmd) for Cα atoms from the α_EC and γC

```
                340                                                                    405
      FBEH    ..ISVNKYR-GTAGNALMDGASQLMGE---NRTMTIHNGMFFSTYDRDNDGWLTSDPRKQCSKEDGGGWWYN..
      FBER    ..VSVNKYK-GTAGNALMEGASQLVGE---NRTMTIHNGMFFSTYDRDNDGWVTTDPRKQCSKEDGGGWWYN..
  β   FBEC    ..LSVSNYK-GNAGNALMEGASQLYGE---NRTMTIHNGMYFSTYDRDNDGWLTTDPRKQCSKEDGGGWWYN..
      FBEL    ..LWVEDYS-GNAGNALLEGATQLMGD---NRTMTIHNGMQFSTFDRDNDNWNPGDPTKHCSREDAGGWWYN..
                276                                                              337
      FGAH    ..LTYAYFAGGDAGDA-FDGFD--FGDDPSDKFFTSHNGMQFSTWDNDNDKF-----EGNCAEQDGSGWWMN..
      FGAR    ..LTYAYFIGGDAGDA-FDGYD--FGDDPSDKFFTSHNGMHFSTWDNDNDKF-----EGNCAEQDGSGWWMN..
  γ   FGAC    ..LTYAYFIGGERGDA-FDGFN--FGDDPSDKSYTYHNGMRFSTFDNDNDNF-----EGNCAEQDGSGWWMN..
      FGAL    ..LFYSMYLDGDAGNA-FDGFD--FGDDPQDKFYTTHLGMLFSTPERDNDKY-----EGSCAEQDGSGWWMN..
                734                                                              791
      FAH2    ..LQVSSYE-GTAGDALIEGSVEEGAE------YTSHNNMQFSTFDRDADQW-----EENCAEVYGGGWWYN..
      FAR2    ..LQVSSYQ-GTAGDALMEGSVEEGTE------YTSHSNMQFSTFDRDADQW-----EENCAEVYGGGWWYN..
  αE  FAC2    ..LTVSSYE-GTAGDALVAGWLEEGSE------YTSHAQMQFSTFDRDQDHW-----EESCAEVYGGGWWYN..
      FAL2    ..LQVSDYR-GTAGNALVSG VADDPE------LTSHGGMTFSTYDRDTDKW----SDGSCAEWYGGGWWIN..
                                                                            **
      SLUG    ..LQVGGYS-GNAGDA----------------LTFHNGMAFSTNDRDNDAD-----SIDXAKVYHGAWWYK..


                406                                                                461
      FBEH    ..RCHAANPNGRYYWGGQY-TWDMAKHGTDDGVVWMNWKGSWYSMRKMSMKIRPFFPQQ
      FBER    ..RCHAANPNGRYYWGGLY-SWDMSKHGTDDGVVWMNWKGSWYSMRRMSMKIRPVFPQQ
  β   FBEC    ..RCHAANPNGRYYWGGTY-SWDMAKHGTDDGIVWMNWKGSWYSMKKMSMKIKPYFPD
      FBEL    ..RCHAANPNGRYYWGGIY-TKEQADYGTDDGVVWMNWKGSWYSMRQMAMKLRPKWP
                338                                                        411
      FGAH    ..KCHAGHLNGVYYQGGTY-SKASTPNGYDNGIIWATWKTRWYSMKKTTMKIIPFNRLTIGEGQQHHLGGAKQAGDV
      FGAR    ..KCHAGHLNGVYYQGGTY-SKSSTPNGYDNGIIWATWKTRWYSMKETTMKIIPFNRLSIGDGQQHHMGGSKQVGDM
  γ   FGAC    ..RCHAGHLNGPYYIGGVY-SRDTGTNSYDNGIIWATWRDRWYSMKKTTMKIIPFNRLSID-GQQH-SGGLKQVGDS
      FGAL    ..RCHAGHLNGKYYFGGNY-RKTDVEFPYDDGIIWATWHDRWYSLKMTTMKLLPMGRDLSGHGGQQQSKG-NSRGDN
                792                                                        847
      FAH2    ..NCQAANLNGIYYPGGSYDPRNNSPYEIENGVVWVSFRGADYSLRAVRMKIRPLVTQ
      FAR2    ..SCQAANLNGIYYPGGTYDPRNNSPYEIENGVLWVPFRGADYSLWAVRMKIRPLVGQ
  αE  FAC2    ..SCQAANLNGIYYPGGHYDPRYNVPYEIENGVVWIPFRASDYSLKVVRMKIRPLETL
      FAL2    ..ACQAANLNGVYYQGGPYDPREKPPYEVENGVVWATYRGSDYSLKRTAVRFRRVQIPIVE
                                 **********
```

FIG. 5. Sequence alignment of regions from various fibrinogen β, γ, and α_E chains and the partial sequence of a lectin from the slug, *Limax flavus*. (26). FBEH, human β; FBER, rat β; FBEC, chicken β; FBEL, lamprey β; FGAH, human γ; FGAR, rat γ; FGAC, chicken γ; FGAL, lamprey γ; FAH2, human α_E; FAR2, rat α_E; FAC2, chicken α_E; and FAL2, lamprey α_E. Only the human sequences are numbered. Those residues in which all 12 nonslug sequences are identical are emboldened, as are two key residues (∗∗) from the binding pocket in which the slug lectin is the same as three of the α_EC domains. The skein at α_E809–818 in which the α_EC differs significantly from β and γ chains is denoted by ∗∗∗∗∗∗∗∗∗∗. The sequence of the remainder of the slug lectin has not yet been reported.

Table 2. Charged amino acids in carboxyl domains of human fibrinogen

Amino acids	γC (res. 152–411)	βC (210–461)	αEC (612–847)
Aspartate	21	18	15
Glutamate	10	13	20
Histidine	7	4	4
Lysine	20	17	4
Arginine	6	13	12
Net charge*	−3.6	−0.2	−18.2

*Net charge at pH 7.3, presuming pK for histidines = 6.5.

domains is 1.26 (192 residues compared), and that of the αEC and the βC domains is 0.96 (193 residues compared). The rmsd for β–γ comparison over the same region is 1.06 (189 residues compared). Thus, in a structural sense, the three domains are just about equally different. Some other expected features included the calcium bound at residues 772–778, as predicted (D.A. and G.G., unpublished data), and the carbohydrate cluster at Asn-667 (6). Also as anticipated (6), there is no evidence for carbohydrate at Asn-812, even though there is a serine two residues further along. The absence of carbohydrate is consistent with a recent report (21) that glycosylation does not occur effectively when the residue after the obligatory serine or threonine is a proline. In the present case, it may be significant, also, that the proline is in the *cis* conformation. The loop region that includes this Asn-Asn-Ser-Pro constellation (res. 812–815) is somewhat disordered.

The most interesting feature of the structure is doubtless the binding cleft. As expected on the basis of sequence comparisons, the cleft cannot possibly bind positively charged peptides of the sort exemplified by Gly-Pro-Arg or Gly-His-Arg. Nonetheless, the cleft remains well defined (Fig. 4) and, as discussed below, is likely functional.

Evolution. It has long been supposed that the vertebrate fibrinogen molecule was assembled during evolution from two principal components found in other animal proteins: a three-stranded α-helical-coiled coil and a globular carboxyl-terminal domain. From the beginning, it was postulated that the prototypic molecule was a homotrimer or homohexamer (1). That the major type of α chain is homologous to the β and γ chains over the course of its first third only was attributed to the original carboxyl-terminal domain of the chain having been lost in some kind of genetic and evolutionary exchange (1), doubtless a part of the specialization and differentiation that accompanied the evolution of fibrin formation. All other known proteins containing three-stranded coiled coils are homotrimers or homohexamers (22). One natural advantage of homopolymers is the amplification that accrues with weak binding for repetitive ligands of the sort that occur on many microbial surfaces (23). As such, the prototypic fibrinogen may have evolved from entities involved in host defense. Whatever the case, all vertebrate fibrinogens are now dimers of heterotrimers, the three constituent chains of which have differentiated greatly from their common ancestor.

Previous sequence analysis had made it clear that the αEC domain is not only homologous to the β and γ carboxyl-terminal domains of vertebrate fibrinogen and other proteins (4) but that it also is more closely related to those two fibrinogen domains than are any other known fibrinogen-related structures (5). In a phylogenetic sense, then, the αE represents a good candidate for a component of the prototypic fibrinogen. Arguments based on function make an even more compelling case that αE chains may represent the prototype α-chain.

Function. That the ancient α-chain carboxyl domain has survived throughout the vertebrate radiation is testimony to its being essential, and yet the function of the extended fibrinogen α chain and the role of its carboxyl-terminal domain remain

mysterious. Among the possibilities, however, are binding to particular cell types or other macromolecules. It is significant that the majority of other animal proteins that contain fibrinogen-related carboxyl domains are involved in cell adhesion (4, 24). This observation is wholly consistent with suggestions that fibrinogen evolved from a lectin-like molecule (10, 22, 24). Indeed, there is circumstantial evidence that the potential binding cleft on the αEC domain may be suitable for binding some kind of sugar. In this regard, of all the fibrinogen-related domains reported, only the βC and γC domains are known to bind peptides. The key residues in ligand binding for the β and γ chains include β397–398 (11) and γ329–330 (25), which are Glu-Asp and Gln-Asp, respectively. The corresponding residues in αEC are a Val and Tyr at αE783–784. Exactly the same residues occur in a sialic acid-binding lectin found in an invertebrate, the slug *Limax flavus* (26). Coincidentally, also, the αEC domains have a gap at the very same place as does the slug lectin (Fig. 5), a deletion which involves a significant shortening of one of the loops adjacent to the binding pocket.

Given these similarities, we would cautiously suggest that the most likely role of the αEC domain involves carbohydrate binding of some sort. Whether such binding would involve other factors or cells involved in clotting or the extracellular matrix remains to be determined. It will be of great interest to find if rαEC binds sialic acid or other sugars.

The αEC domain also contains a 10-residue segment (αE809–818), the sequence of which is unchanged in all those mammals examined and which appears unique among the fibrinogen-related domains (Fig. 5). The segment forms a well defined loop that could easily be involved in binding to other macromolecules. The αEC domain is significantly more electronegative than the βC and γC domains (Table 2), a feature that may bear on its attachments. The electronegative character is mainly attributable to a dearth of lysine sidechains, rather than to a build-up of negative charges.

Finally, it should be mentioned that the carbohydrate at Asn αE667 is located at a position where several other fibrinogen-related domains have potential glycosylation sites (24). In addition to all the other known αEC domains in mammals and chicken and the equivalent domain in lamprey, putative carbohydrate-binding sites occur at this position in the *Drosophila* scabrous protein (27), in the lamprey γ chain (28), and in the mammalian fibrinogen-related protein known as PT49 (29, 30).

1. Doolittle, R. F., Watt, K. W. K., Cottrell, B. A., Strong, D. D. & Riley, M. (1979) *Nature (London)* **280,** 464–468.
2. Fu, Y. & Grieninger, G. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 2625–2628.
3. Weissbach, L. & Grieninger, G. (1990) *Proc. Natl. Acad. Sci. USA* **87,** 5198–5202.
4. Fu, Y., Weissbach, L., Plant, P. W., Oddoux, C., Cao, Y., Liang, T. J., Roy, S. N., Redman, C. M. & Grieninger, G. (1992) *Biochemistry* **31,** 11968–11972.
5. Pan, Y. & Doolittle, R. F. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 2066–2070.
6. Fu, Y., Cao, Y., Hertzberg, K. M. & Grieninger, G. (1995) *Genomics* **30,** 71–76.
7. Doolittle, R. F., Riley, R. & Pan, Y. (1992) *Thromb. Res.* **68,** 489–493.
8. Grieninger, G., Lu, X., Cao, Y., Fu, Y., Kudryk, B., Galanakis, D. K. & Hertzberg, K. M. (1997) *Blood* **90,** 2609–2614.
9. Shipwash, E., Pan, Y. & Doolittle, R. F. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 968–972.
10. Yee, V. C., Pratt, K. P., Cote, H. C., LeTrong, I., Chung, D. W., Davie, E. W., Stenkamp, R. E. & Teller, D. C. (1997) *Structure (London)* **5,** 125–138.

11. Spraggon, G., Everse, S. J. & Doolittle, R. F. (1997) *Nature (London)* **389,** 455–462.
12. Jancarik, J. & Kim, S. H. (1991) *J. Appl. Crystallogr.* **24,** 409–411.
13. Otwinowski, Z. & Minor, W. (1997) *Methods Enzymol.* **276,** 307–326.
14. Collaborative Computing Project Number 4 (1994) *The CCP4 Suite: Programs for Protein Crystallography*, Version 3.1, *Acta Crystallogr. D* **50**, 760–763.
15. Brunger, A. T. (1992) X-PLOR A System for X-Ray Crystallography and NMR (Yale Univ. Press, New Haven, CT), Version 3.1.
16. Everse, S. J., Spraggon, G., Veerapandian, L., Riley, M. & Doolittle, R. F. (1998) *Biochemistry* **37,** 8637–8642.
17. Brunger, A. T. (1990) *Acta Crystallogr. A* **46,** 46–57.
18. Read, R. J. (1986) *Acta Crystallogr. A* **42,** 140–149.
19. Jones, T. A., Zou, J.-Y, Cowan, S. W. & Kjeldgaard, M. (1991) *Acta Crystallogr. A* **47,** 110–119.
20. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997) *Acta Crystallogr. D* **53,** 240–255.
21. Mellquist, J. L., Kasturi, L., Spitalnik, S. L. & Shakin-Eshleman, S. H. (1998) *Biochemistry* **37,** 6833–6837.
22. Doolittle, R. F., Spraggon, G. & Everse, S. J. (1997) in *Plasminogen Related Growth Factors*, eds. Bock, G. R. & Goode, J. A. (Wiley, New York), pp. 4–23.
23. Weis, W. I. & Drickamer, K. (1996) *Annu. Rev. Biochem.* **65,** 441–473.
24. Doolittle, R. F. (1992) *Protein Sci.* **1,** 1563–1577.
25. Pratt, K. P., Cote, H. C. F., Chung, D. W., Stenkamp, R. E. & Davie, E. W. (1997) *Proc. Natl. Acad. Sci. USA* **94,** 7176–7181.
26. Knibbs, R. N., Osborne, S. E., Glick, G. D. & Goldstein, I. J. (1993) *J. Biol. Chem.* **268,** 18524–18531.
27. Baker, N. E., Mlodzik, M. & Rubin, G. M. (1990) *Science* **250,** 1370–1377.
28. Strong, D. D., Moore, M., Cottrell, B. A., Bohonus, V. L., Pontes, M., Evans, B., Riley, M. & Doolittle, R. F. (1985) *Biochemistry* **24,** 92–101.
29. Koyama, T., Hall, L. R., Haser, W. G., Tonegawa, S. & Saito, H. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 1609–1613.
30. Ruegg, C. & Pytela, R. (1995) *Gene* **160,** 257–262.
31. Kraulis, P. J. (1991) *J. Appl. Crystallogr.* **24,** 946–950.
32. Esnouf, R. M. (1997) *J. Mol. Graphics* **15,** 133–138.
33. Bacon, D. J. & Anderson, W. F. (1988) *J. Mol. Graphics* **6,** 219–220.
34. Merritt, E. A. & Murphy, M. E. P. (1994) *Acta Crystallogr. D* **50,** 869–873.
35. Nicholls, A., Sharp, K. & Honig, B. (1991) *Proteins Struct. Funct. Genet.* **11,** 281–296.