

## Carboxy-Terminal-Extended Variant of the Human Fibrinogen $\alpha$ Subunit: A Novel Exon Conferring Marked Homology to $\beta$ and $\gamma$ Subunits<sup>†,‡</sup>

Yiping Fu, Lawrence Weissbach,<sup>§</sup> Patricia W. Plant,<sup>||</sup> Carole Oddoux,<sup>⊥</sup> Yan Cao, T. Jake Liang,<sup>#</sup> Samar N. Roy, Colvin M. Redman, and Gerd Grieninger\*

The Lindsley F. Kimball Research Institute of the New York Blood Center, New York, New York 10021

Received September 10, 1992; Revised Manuscript Received October 12, 1992

**ABSTRACT:** Similarities between the N-terminal regions of the three subunits of the clotting protein fibrinogen—( $\alpha\beta\gamma$ )<sub>2</sub>—suggest that they evolved from a common progenitor. However, to date no human  $\alpha$  chain has been found with the strong C-terminal homology shared by the  $\beta$  and  $\gamma$  chains. Here we examine the natural product of a novel fibrinogen  $\alpha$  chain transcript bearing a separate open reading frame that supplies the missing C-terminal homology to the other chains. Additional splicing leads to the use of this extra sequence as a sixth exon elongating the  $\alpha$  chain by 35%. Since the extended  $\alpha$  chain ( $\alpha_E$ ) is assembled into fibrinogen molecules and its synthesis is enhanced by interleukin-6, it suggests participation in both the acute phase response and normal physiology.

Fibrinogen acts as the main factor in the formation of a blood clot by polymerization to a fibrin network and by enabling platelets to aggregate. This multichain molecule—composed of two copies each of the  $\alpha$ ,  $\beta$ , and  $\gamma$  subunits—is encoded by three separate genes which are clustered in an ~50-kb region on chromosome 4q23–q32 (Blomback et al., 1968; Kant et al., 1985). The N-termini of the subunits all contain  $\alpha$ -helical regions which combine in assembled fibrinogen to form a disulfide-interlocked coiled coil (Rao et al., 1991). Evolution of the three subunits by duplication from a common ancestral gene has been advanced on the basis of comparisons, particularly in these N-terminal regions (Doolittle, 1983; Henschen et al., 1983), but there has been no evidence of a domain in the human fibrinogen  $\alpha$  chain that is equivalent to the C-terminal domains of the  $\beta$  and  $\gamma$  chains with stretches of ~250 amino acids sharing marked homology. Recently, two studies, in chicken (Weissbach & Grieninger, 1990) and lamprey (Pan & Doolittle, 1992), have identified nucleotide sequences that could potentially encode this type of domain.<sup>1</sup> In the present study, an extended  $\alpha$  chain ( $\alpha_E$ )<sup>2</sup>—the actual product of a novel human fibrinogen  $\alpha$  chain transcript, with extensive homology to the  $\beta$  and  $\gamma$  chain C-termini—is described and its function explored.

### EXPERIMENTAL PROCEDURES

For Northern analysis, 2  $\mu$ g of poly(A) RNA from adult human liver was run on a 1.2% agarose gel and transferred to a nylon membrane. Single-stranded hybridization probes were generated by PCR with one (antisense) primer as follows: unlabeled dCTP was replaced with 200  $\mu$ Ci of [ $\alpha$ -<sup>32</sup>P]-dCTP and the number of cycles changed to eight with an elongation time of 15 min.

Vectors, containing full-length cDNAs for the three fibrinogen chains (pBC12BI-A $\alpha$ , -B $\beta$ , and - $\gamma$ ), and transfection procedures have been described (Roy et al., 1991). To generate an  $\alpha_E$ -containing vector, a cDNA covering the entire coding region was generated by PCR, then blunt-ended with T4-DNA polymerase, cut with *Bgl*II, and ligated into pBC12BI-A $\alpha$  which had been digested with *Bgl*II/*Sma*I. The insert begins 29 nt upstream of the initiating methionine and ends 8 nt after Gln 847 as verified by sequencing. For calcium phosphate transfection, each cDNA-containing vector was used at 5  $\mu$ g of DNA/mL; a vector containing no insert was used to maintain the DNA concentration at 20  $\mu$ g/mL in all cases.

For generation of anti- $\alpha_E$ , an exon VI-containing cDNA clone was cut first with *Bam*HI and then partially digested with *Hind*III; the longer *Bam*HI/*Hind*III fragment was isolated and ligated into the *Bam*HI/*Hind*III site of pQE-9 (Qiaexpress) which contains six histidines in front of the cloning region. The resulting insert (nt 2664–3583, as verified by sequencing) codes for the 213 C-terminal amino acids of  $\alpha_E$ . The exon VI-containing vector was expressed in *Escherichia coli*, and the recombinant protein, affinity purified via its N-terminal His tag with a nickel-NTA column, was used for antibody production in rabbits.

### RESULTS AND DISCUSSION

It is widely held that the human fibrinogen  $\alpha$  subunit (67 kDa) is derived from an ~2.5-kb mRNA (Kant et al., 1983; Rixon et al., 1983; Roy et al., 1990). In contrast, the chicken  $\alpha$  subunit (54 kDa) is encoded on a much larger mRNA (3.9 kb) that is bipartite (Weissbach & Grieninger, 1990); i.e., it contains both an  $\alpha$  coding region and an intriguing second open reading frame (termed FASORF) of unknown function. While exploring the substantial homologies between this open

<sup>†</sup> This work was supported in part by the National Institutes of Health through grants to G.G. and C.M.R. as well as a Biomedical Training Fellowship to C.O. and by the American Heart Association through a grant to G.G.

<sup>‡</sup> The complete sequence of the transcript has been deposited in GenBank under Accession No. M58569.

\* To whom correspondence should be addressed (telephone, 212-570-3124; Fax, 212-737-4506).

<sup>§</sup> Present address: Orthopaedic Research Laboratories, Massachusetts General Hospital, Boston, MA 02114.

<sup>||</sup> Present address: Department of Biology, Wingate College, Wingate, NC 28174.

<sup>⊥</sup> Present address: Department of Pharmacology, New York University Medical Center, New York, NY 10016.

<sup>#</sup> Present address: Gastrointestinal Unit, Massachusetts General Hospital, Boston, MA 02114.

<sup>1</sup> There are now indications that in both species these sequences are expressed (Y. Fu and G. Grieninger, unpublished results; R. F. Doolittle, personal communication).

<sup>2</sup> Abbreviations:  $\alpha_E$ , extended variant of the fibrinogen  $\alpha$  subunit; FASORF, fibrinogen  $\alpha$ -subunit second open reading frame; PCR, polymerase chain reaction; RT-PCR, reverse transcriptase PCR.

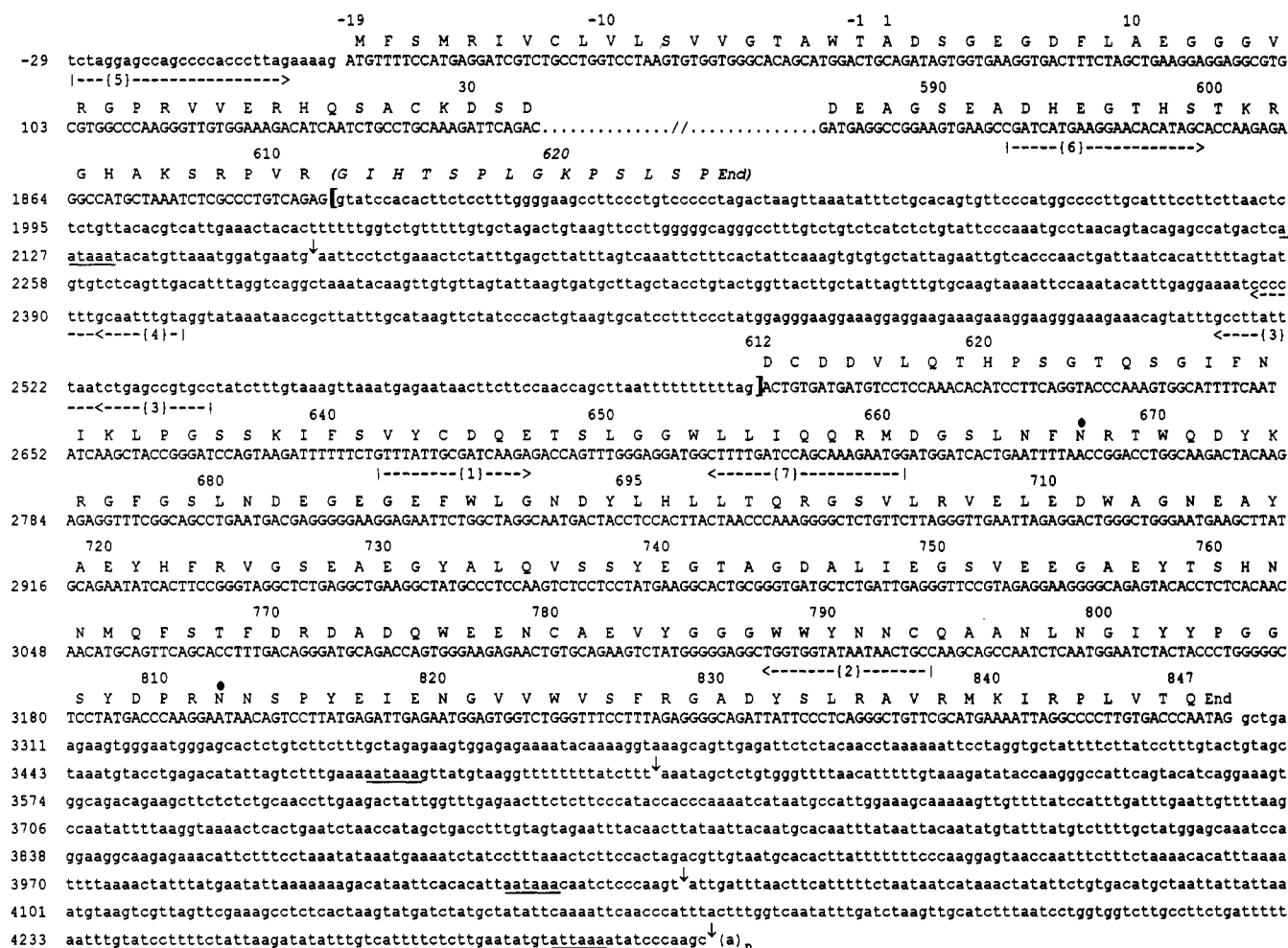


FIGURE 1: Nucleotide and deduced amino acid sequence of the bipartite  $\alpha$ - and FASORF-containing transcript. The midportion of the  $\alpha$  coding region is abbreviated. Nucleotides are numbered beginning with the ATG coding for the initiating methionine of the signal sequence. The amino acids, using single-letter code, are numbered starting at the alanine of the mature protein. Exon sequences (coding regions only) are represented in capital letters; introns and other untranslated sequences, in lower-case letters. Splice sites are marked by square brackets. Potential N-linked glycosylation sites are marked with a diamond. Poly(A) signals are underlined, and poly(A) sites are indicated by vertical arrows. The horizontal, numbered arrows indicate the positions and orientations of selected PCR primers. Whereas the carboxy end of the extended  $\alpha$  chain is encoded by FASORF (last exon), the C-terminal amino acids of the predominant  $\alpha$  chain (italics) are produced by translation into the last intron. The sequence presented is based on the sequencing of several overlapping cDNA clones in combination with direct sequencing of PCR products generated by various primer pairs on human templates, including mRNA (adult and fetal liver, Clontech; Hep-3B2 cells) and cDNA libraries (adult liver and Hep-G2, Stratagene). Standard procedures were used (Berger & Kimmel, 1987; Sambrook et al., 1989). In the primary PCR, human genomic DNA was used as template for the mixed oligonucleotide primers 1 and 2. Primer 1 [GT(ACGT)T(AT)(CT)TG(CT)GA(CT)CA(AG)GA] was biased toward FASORF among the other fibrinogen gene family members due to the presence of a unique glutamine near its 3' end. When the resulting fragment (with 78% identity to chicken FASORF) was used to screen a human adult liver cDNA library in  $\lambda$ gt11 (Clontech), several distinct 3' clones were found that were polyadenylated at the different positions shown. The complete upstream sequence was initially obtained by reverse transcriptase PCR (RT-PCR) with adult liver poly(A) RNA as template using primer 3, followed by nested primer 4 in conjunction with upstream primer 5 based on the known 5' sequence of the  $\alpha$  gene (Chung et al., 1991). Removal of the last intron was first demonstrated by comparison of the two products obtained by RT-PCR of oligo-(dT)-primed adult liver poly(A) RNA in conjunction with different primer pairs such as 6 and 7; splice sites were identified by direct sequencing of the smaller product.

reading frame and the human fibrinogen  $\beta$  and  $\gamma$  chains (Chung et al., 1983a,b), tenascin (Nies et al., 1991), gene X (opposite strand of P450c21) (Morel et al., 1989), the mouse T cell transcript pT49 (Koyama et al., 1987), and the sea cucumber gene FRP-A (Xu & Doolittle, 1990), we encountered the first evidence of a human counterpart. It was in the form of a genomic DNA fragment (corresponding to nt 2685–3141 in Figure 1) generated by PCR using mixed primers that corresponded to well-conserved amino acid clusters apparent in multiple sequence alignments [e.g., see Weissbach and Grienering (1990) and Pan and Doolittle (1992)] of this apparent gene family. The nucleotide sequences of a number of human liver cDNA clones isolated by hybridization to this genomic DNA fragment were consistent with the existence of a long transcript with a bipartite structure like that of chicken  $\alpha$  mRNA. Southern blot analysis of human DNA, using  $\alpha$ - as well as FASORF-specific probes, suggested that

the two coding regions are contiguous and present at only one copy per haploid genome (data not shown).

The complete sequence of the bipartite human transcript was obtained by PCR methodology (Figure 1). The amino acid sequence deduced from its first open reading frame is identical to the consensus sequence for the human fibrinogen  $\alpha$  chain precursor compiled from nucleic acid and protein sequence data,<sup>3</sup> which confirms that all five established  $\alpha$  gene exons are present. Moreover, the beginning of the second frame (FASORF) and its upstream region align with the sequence of the 3'-flanking region of the previously described human  $\alpha$  gene (Chung et al., 1991), indicating an extension of the known transcription unit. This physical connection between human FASORF and the established  $\alpha$  coding region

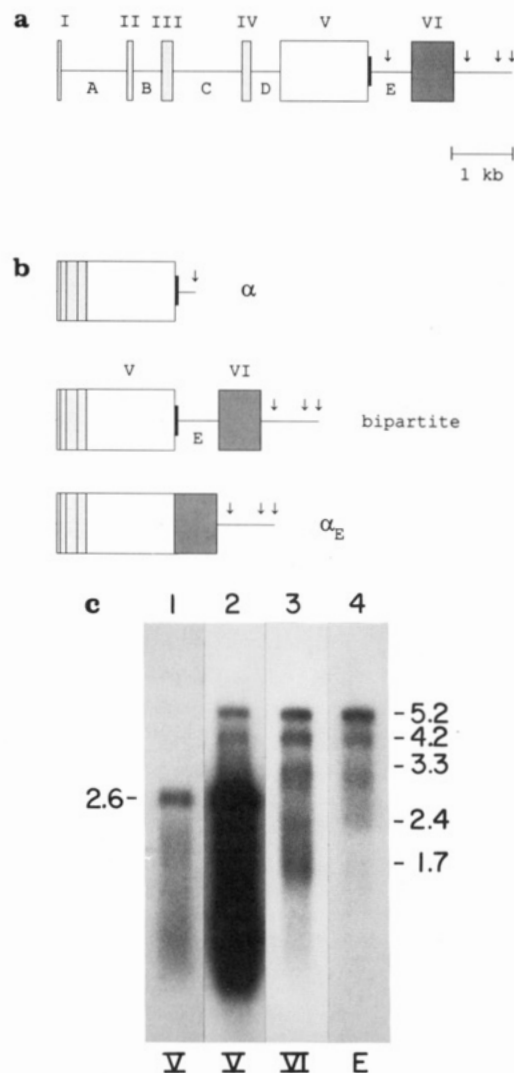
<sup>3</sup> Protein Identification Resource/Dayhoff Database Accession No. A03116.

parallels that of chicken FASORF but not that of the recently described counterpart to FASORF in a lower vertebrate, lamprey, where it appears to have evolved as part of a second fibrinogen  $\alpha$  gene (Pan & Doolittle, 1992).

FASORF, whether human or chicken, lacks an initiating methionine codon, but its upstream nucleotide sequence bears 3'-intron-like features (an AG dinucleotide preceded by a tract of 10 thymidines), suggesting that splicing to an appropriate N-terminal coding region might enable its translation. To test whether some portion of the human  $\alpha$  chain coding region could serve as such an N-terminus, we performed reverse transcriptase PCR (RT-PCR) with human liver mRNA as template using a series of primer pairs: sense in the  $\alpha$ - and antisense in the FASORF-coding regions (e.g., primers 6 and 7 in Figure 1). In each case at least two products were found, one always 700 bp shorter than the other. Sequencing revealed the smaller to be derived from splicing within exon V at a donor site, which is located 14 codons upstream from the C-terminus of the common  $\alpha$  chain, to the anticipated acceptor site at the beginning of FASORF. Sequencing of the cDNA of the spliced transcript proved the presence of all five exons in their entirety in addition to FASORF, which we henceforth denote as exon VI. We refer to the predicted product of the spliced transcript, containing  $\sim 35\%$  more amino acids, as an extended  $\alpha$  chain ( $\alpha_E$ ). In Figure 2a,b the organization of the  $\alpha$  gene and these transcripts are summarized schematically. For clarity, the term "exon" is restricted to the protein coding sequences of this complex gene.

The multiple fibrinogen  $\alpha$  gene transcripts in normal human liver were visualized by Northern blot analysis of poly(A)-containing RNA with hybridization probes specific to exons V and VI and intron E which separates them (Figure 2c). The band at 2.6 kb, hybridizing to only the exon V probe, corresponds to the predominant  $\alpha$  mRNA. Identification of the exon VI-containing transcripts was based on differential hybridization to exon VI and intron E probes and on size predictions, allowing for use of various polyadenylation sites in the 3' untranslated region (particularly at nt 3505 and 4301). The band at 5.2 kb, hybridizing all three probes, is a bipartite transcript; the one at 3.3 kb, which was not recognized by the intron E probe, is a spliced  $\alpha_E$  transcript; and the broad band at 4.2 kb contains both. The identity of two additional transcripts (2.4 and 1.7 kb) is being investigated. The fact that intron E is found on about half of the exon VI-containing species, i.e., the bipartite transcripts, indicates that it is not spliced out as efficiently as are introns A–D.

To date, the strongest similarities evident between the human fibrinogen subunits have been between the carboxy ends of the  $\beta$  and  $\gamma$  chains. With the discovery of the extended  $\alpha$  chain, the common ancestry shared by all three chains is manifest (Figure 3). It is further underscored by the precision with which intron E, demarcating the 5' end of exon VI, aligns with intron positions (all type I splice junctions) at the very beginning of the corresponding homologous regions of the  $\beta$  and  $\gamma$  genes (Chung et al., 1991). As shown in Figure 3a, the sequence of the C-terminal extension of  $\alpha_E$  is as similar to the carboxy ends of the human fibrinogen  $\beta$  and  $\gamma$  chains as the latter two are to each other ( $\sim 40\%$  identity). In light of this symmetry of homology, we propose a revision of earlier evolutionary schemes (Doolittle, 1983; Henschen et al., 1983; Weissbach & Grieninger, 1990) to indicate that the gene duplications leading to development of the three separate subunits occurred at a similar time, more than 700 million years ago (Figure 3b). On the basis of hydrophobicity analyses (Kyte & Doolittle, 1982) (not shown), it is expected that the extended  $\alpha$  C-terminus folds independently into a globular



**FIGURE 2:** Fibrinogen  $\alpha$  gene and its transcripts. (a) Organization of the gene. The term exon is restricted to protein coding sequences only. Exons are depicted as boxes: exons I–IV, light gray; exon V, white; FASORF (exon VI), dark gray. Introns A–E and the 3' untranslated region are drawn as lines. The narrow region at the end of exon V that is alternately translated (to generate the C-terminus of the predominant  $\alpha$  chain) or spliced out (as part of intron E for formation of  $\alpha_E$ ) is depicted half-height in black. Arrows show polyadenylation sites. PCR with human genomic DNA as template was used to extend the structure of the known fibrinogen  $\alpha$  transcription unit, demonstrating that the 3' cDNA sequence of the bipartite transcript, beginning with exon V, is identical to the genomic DNA. (b) Schematic presentation of transcripts. Levels of transcripts were estimated by RT-PCR and Northern blot analysis using different batches of human liver RNA (adult and fetal). Comparable levels were found with RNAs from various sources. Transcripts are presented in order of prevalence: predominant  $\alpha$ ,  $\sim 90\%$ ; bipartite,  $\sim 5\%$ ; extended  $\alpha$  ( $\alpha_E$ ),  $\sim 2\%$ . The same symbols as in (a) are used. (c) Northern blot analysis. The same blot containing human liver RNA is shown hybridized successively to three different probes, each specific for a separate section of the  $\alpha$  gene: lanes 1 and 2 (the latter overexposed to make the higher molecular weight bands visible), exon V probe; lane 3, exon VI probe; lane 4, intron E probe. The following regions were covered by the probes: exon V, most of the exon downstream to nt 1896; exon VI, nt 2608–3288; intron E, its 3' end from nt 2403 to nt 2537. Transcript sizes, estimated relative to RNA size markers, are indicated in kilobases.

region as do the corresponding regions of the  $\beta$  and  $\gamma$  chains.

The presence of  $\alpha_E$  transcripts in normal liver (Figure 2c) suggests a possible role for this variant chain in basic fibrinogen metabolism. To explore the matter further, synthesis, assembly, and secretion of  $\alpha_E$  were studied, as summarized in Figure 4, (a) in COS cells transfected with subunit cDNAs to manipulate fibrinogen subunit composition and (b) in human

$\alpha_E$	DCDDVLQTHPSGT	QSGIFNIKLPGSSKIFSVYCDQETS	SLGGWLLIQQRMDGSLNFRNT
$\beta$	ECEEIIRK	GGETSEMYLIQPDSSVKPYRVYCDMNTENGWTVIQNRQDGSVDGFRK	
$\gamma$	DCQDIANK	GAQSGSLYFIKPLKANQQFLVYCEIDGSGNGWTVFQKRLDGSVDFFKN	
*	<u>D</u> <u>C</u> <u>D</u> <u>I</u> <u>K</u> <u>G</u> <u>Q</u> <u>S</u> <u>G</u> <u>Y</u> <u>I</u> <u>K</u> <u>P</u> <u>S</u> <u>K</u> <u>F</u> <u>V</u> <u>Y</u> <u>C</u> <u>D</u>	<u>T</u> <u>S</u> <u>G</u> <u>G</u> <u>W</u> <u>T</u> <u>V</u> <u>I</u> <u>Q</u> <u>R</u> <u>D</u> <u>G</u> <u>S</u> <u>V</u> <u>D</u> <u>F</u> <u>R</u>	
$\alpha_E$	WDQYKRGFG	SLNDEGEGEFWLGN DYLLHLLTQRG	SVLRVELEDWAG
$\beta$	WDPYKQGFNVATNTDGNKYCGLPGEYWLGN DKISQLTRMG	PTELLIEMEDWKGDKV	
$\gamma$	WQYKEGFG	HLSPGTGTTFWLGNEKIHLISTQSAIPYALRVELEDWNGRTS	
*	<u>W</u> <u>Y</u> <u>K</u> <u>G</u> <u>F</u> <u>G</u>	<u>L</u> <u>G</u> <u>G</u> <u>E</u> <u>F</u> <u>W</u> <u>L</u> <u>G</u> <u>N</u> <u>D</u> <u>K</u> <u>I</u> <u>H</u> <u>L</u> <u>L</u> <u>T</u> <u>G</u> <u>P</u> <u>L</u> <u>R</u> <u>V</u> <u>E</u> <u>L</u> <u>E</u> <u>D</u> <u>W</u> <u>G</u>	
$\alpha_E$	NEAYAEYHFRVSGEAEAGYALQVSSYEGT	AGDALIEGSVEEGAE	YTSHNNMQFSTF
$\beta$	KAHYG	GFTVQNEANKYQISVNKYRG	AGNALMDGASQLMGENRTMTIHNGMFFSTY
$\gamma$	TADYA	MFVKVGEADKYRLTYAYFAGGDAGDAFDGDFDPSDKFFSTSHNGMQFSTW	
*	<u>A</u> <u>Y</u> <u>A</u> <u>F</u> <u>V</u> <u>G</u> <u>E</u> <u>A</u> <u>K</u> <u>Y</u> <u>L</u> <u>V</u> <u>Y</u> <u>G</u> <u>T</u> <u>A</u> <u>G</u> <u>D</u> <u>A</u> <u>L</u> <u>G</u>	<u>E</u> <u>T</u> <u>S</u> <u>H</u> <u>N</u> <u>G</u> <u>M</u> <u>Q</u> <u>F</u> <u>S</u> <u>T</u>	
$\alpha_E$	DRDADQW	EENCAEVYGGGWYNNCQAANLNGIYYPGGSYDPRNNSPYEINGVV	
$\beta$	DRDNDGWLTS	DRKQCSKEDGGGWYNNRCHAANPNRGYRWGGQYTWD	AKHGTDDGVV
$\gamma$	DNDNDKF	EGNCAEQDGSWWMNKCHAGHLNGVYVYGGTYSKAS	TPNGYDNGII
*	<u>D</u> <u>R</u> <u>D</u> <u>N</u> <u>D</u> <u>W</u>	<u>E</u> <u>N</u> <u>C</u> <u>A</u> <u>E</u> <u>D</u> <u>G</u> <u>G</u> <u>W</u> <u>W</u> <u>Y</u> <u>N</u> <u>C</u> <u>H</u> <u>A</u> <u>A</u> <u>N</u> <u>L</u> <u>N</u> <u>G</u> <u>Y</u> <u>Y</u> <u>G</u> <u>G</u> <u>Y</u>	<u>P</u> <u>G</u> <u>D</u> <u>N</u> <u>G</u> <u>V</u> <u>V</u>
$\alpha_E$	WVSFRGADYSLRAVRMKIRPLVTQ		
$\beta$	WMNWKGSWYSMRKMSMKIRPFPPQ		
$\gamma$	WATWKTRWYSMKKTTMKIIPFNRLTIGEGQQHHLGGAKQAGDV		
*	<u>W</u> <u>W</u> <u>K</u> <u>G</u> <u>W</u> <u>Y</u> <u>S</u> <u>M</u> <u>R</u> <u>K</u> <u>M</u> <u>K</u> <u>I</u> <u>R</u> <u>P</u> <u>F</u> <u>Q</u>		

FIGURE 3: (a) Comparison of the C-terminal regions of the  $\alpha_E$ ,  $\beta$ , and  $\gamma$  chains of human fibrinogen. Alignment of the amino acid sequences is presented using single-letter code. The consensus sequence (\*) is composed of amino acids present in at least two out of the three sequences at the corresponding position; residues present in all three are underlined. The sequences begin with the following amino acids:  $\alpha_E$ , Asp 612;  $\beta$ , Glu 210; and  $\gamma$ , Asp 152 (Chung et al., 1983a,b). (b) Proposed scheme for evolution of the fibrinogen genes. The three subunits are derived by gene duplication from a common ancestor. Shapes are used to symbolize the coding regions for the various domains: wavy line, the  $\alpha$ -helical segment bordered by cysteine brackets (short vertical lines); ellipse, the globular C-terminal domain; rectangle, the long segment of the  $\alpha$  chain—containing oligopeptide repeats—that bears no sequence homology to the  $\beta$  and  $\gamma$  chains (Doolittle et al., 1979). The rectangular section, which also encodes the carboxy end of the common  $\alpha$  chain, may have been a subsequent insertion (as depicted) or part of the ancestral gene that was lost later from the  $\beta$  and  $\gamma$  genes.

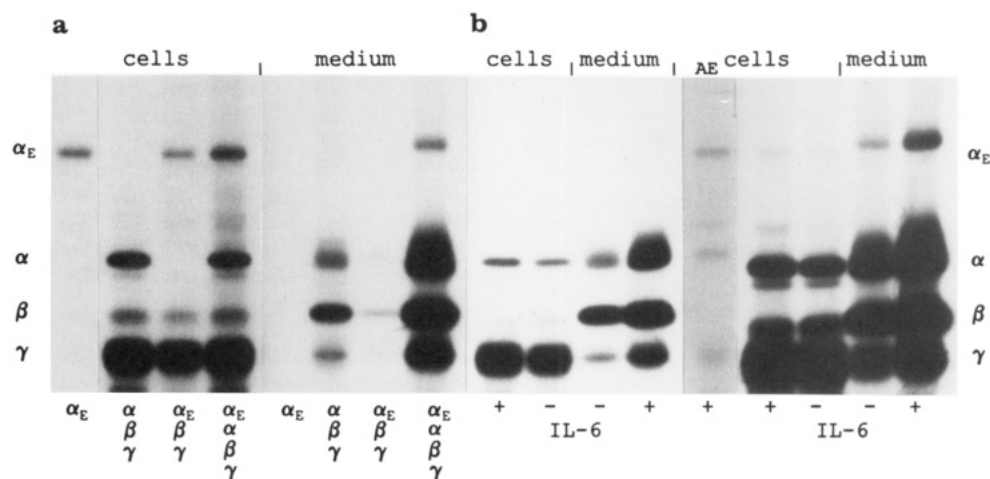


FIGURE 4: Synthesis, assembly, and secretion of the fibrinogen  $\alpha_E$  chain in COS cells and human hepatoma (Hep-G2) cells. (a) Expression of fibrinogen chain cDNAs in COS cells. COS-1 cells were transiently transfected with vectors containing human fibrinogen subunit cDNAs, either alone or in combination, as indicated below each lane. After the cells were incubated for 2 h with [ $^{35}$ S]methionine, the fibrinogen chains were immunoprecipitated from cell lysates and culture medium using a rabbit antibody to whole human fibrinogen (Roy et al., 1991). The chains were separated by SDS-PAGE under reducing conditions and detected by autoradiography. Positions of the individual chains are indicated. (b) Stimulation of fibrinogen chain synthesis by interleukin-6 (IL-6) in Hep-G2 cells. Hep-G2 cells were cultured (Roy et al., 1991) for 24 h in the absence (–) or presence (+) of 20 ng/mL recombinant, baculovirus-derived, human IL-6 (May et al., 1991). Cells were labeled and processed as above for COS cells. Anti-fibrinogen was used in all cases except in the lane marked (AE) where it was replaced by anti- $\alpha_E$  for immunoprecipitation of the lysate of IL-6-stimulated cells. In the right panel, the autoradiogram was overexposed to enhance visibility of the  $\alpha_E$  chain.

hepatoma cells stimulated by interleukin-6 (IL-6) to mimic the acute phase response of the liver (Sehgal et al., 1989). Fibrinogen was immunoprecipitated from cells as well as medium with antibody to whole human fibrinogen, and the reduced subunits were separated by SDS-PAGE. Since only assembled fibrinogen is secreted (Amrani et al., 1983; Kudryk et al., 1982; Plant & Griener, 1986)—and this holds true also for COS cells transfected with vectors containing fibrinogen  $\alpha$ ,  $\beta$ , and  $\gamma$  cDNAs (Roy et al., 1991)—subunits detected in the medium by this procedure reflect the constituents of the oligomeric molecule.

When COS cells were transfected with a vector containing  $\alpha_E$  cDNA alone, they expressed a protein, immunoprecipitable with anti-fibrinogen, of  $\sim 95$  kDa (Figure 4a), close to the predicted molecular size of mature  $\alpha_E$  (92 843 Da) with

its 847 amino acids. This protein was not detectable in the medium, as expected for an unassembled fibrinogen chain. When these cells were cotransfected with vectors containing the fibrinogen  $\beta$  and  $\gamma$  cDNAs, all three chains were synthesized but not secreted into the medium. When the vector containing the common  $\alpha$  chain cDNA was included, however, fibrinogen molecules, some containing the  $\alpha_E$  chain, were secreted. The apparent increase in molecular mass of  $\alpha_E$  (to  $\sim 110$  kDa) upon export to the medium indicates a significant degree of processing, possibly involving the potential glycosylation sites in the extended C-terminus (Figure 1).

The obligatory presence of the short  $\alpha$  chain for  $\alpha_E$  incorporation may reflect steric prohibition against the joining of two elongated chains in the same fibrinogen molecule. Alternatively, it may indicate a specific active role for the

shorter  $\alpha$  chain in the intracellular assembly of fibrinogen, one which could be related to its proline-terminated carboxy end (shown in italics in Figure 1). The fact that this segment of 14 amino acids, excluded from  $\alpha_E$ , is also missing from the  $\alpha$  chain as it has been characterized in circulating plasma fibrinogen (Doolittle et al., 1979; Henschen et al., 1980; Kant et al., 1983; Rixon et al., 1983) strengthens such a notion.

When human hepatoma cells are exposed to IL-6 in culture, fibrinogen secretion is increased (May et al., 1988). In Figure 4b (left panel) it can be seen that IL-6-treated Hep-G2 cells increase synthesis and secretion of fibrinogen  $\alpha$ ,  $\beta$ , and  $\gamma$  chains to a similar extent. Overexposure of the autoradiogram (right panel) revealed, in addition, comparable stimulation of a protein (~95 kDa in cell lysate, ~110 kDa in medium) immunoprecipitable with anti-fibrinogen. This protein was positively identified as  $\alpha_E$  by specific recognition with an antibody to its C-terminal domain; no significant cross-reactivity of the antibody with either the  $\beta$  or  $\gamma$  subunit was detected (Figure 4b, right panel, lane marked AE). Thus  $\alpha_E$  appears to be a component of both normal and acute phase fibrinogen expression.

The further potential significance of the novel C-terminus of  $\alpha_E$  may derive from its similarity to carboxy domains in a number of non-fibrinogen proteins from fruit fly to man (Baker et al., 1990; Koyama et al., 1987; Morel et al., 1989; Nies et al., 1991; Norenberg et al., 1992; Xu & Doolittle, 1990). Where functions are known, these proteins are constituents of the extracellular matrix and have adhesive properties. Whether the globular domain of  $\alpha_E$  contributes in a subtle way to the primary function of fibrinogen (clot formation) or, following the example of other differentially used exons (Chan et al., 1991; Descombes & Schibler, 1991; Early et al., 1980), promotes an alternative function [e.g., in the less well studied associations of fibrinogen with wound healing and with malignant tissue epithelia (Dvorak, 1986)], the mere fact that it is preserved in the genome signals a vital role in fibrinogen metabolism.

## ACKNOWLEDGMENT

For their critical reading of the manuscript, we thank P. B. Sehgal (whom we also acknowledge for the gift of recombinant IL-6) and J. Adamson. We are grateful to K. M. Hertzberg for editorial assistance.

## REFERENCES

- Amrani, D. L., Plant, P. W., Pindyck, J., Mosesson, M. W., & Grieninger, G. (1983) *Biochim. Biophys. Acta* 743, 394–400.
- Baker, N. E., Mlodzik, M., & Rubin, G. M. (1990) *Science* 250, 1370–1377.
- Berger, S. L., & Kimmel, A. R., Eds. (1987) *Methods in Enzymology*, Vol. 152, Academic Press, New York.
- Blomback, B., Blomback, M., Henschen, A., Hessel, B., Iwanaga, S., & Woods, K. R. (1968) *Nature* 218, 130–134.
- Chan, A. M.-L., Rubin, J. S., Bottaro, D. P., Hirschfield, D. W., Chedid, M., & Aaronson, S. A. (1991) *Science* 254, 1382–1385.
- Chung, D. W., Chan, W.-Y., & Davie, E. W. (1983a) *Biochemistry* 22, 3250–3256.
- Chung, D. W., Que, B. G., Rixon, M. W., Mace, M., Jr., & Davie, E. W. (1983b) *Biochemistry* 22, 3244–3250.
- Chung, D. W., Harris, J. E., & Davie, E. W. (1991) *Adv. Exp. Med. Biol.* 281, 39–48.
- Descombes, P., & Schibler, U. (1991) *Cell* 67, 569–579.
- Doolittle, R. F. (1983) *Ann. N.Y. Acad. Sci.* 408, 13–27.
- Doolittle, R. F., Watt, K. W. K., Cottrell, B. A., Strong, D. D., & Riley, M. (1979) *Nature* 280, 464–468.
- Dvorak, H. F. (1986) *N. Engl. J. Med.* 315, 1650–1659.
- Early, P., Rogers, J., David, M., Calame, K., Bond, M., Wall, R., & Hood, L. (1980) *Cell* 20, 313–319.
- Henschen, A., Lottspeich, F., Southan, C., & Topfer-Petersen, E. (1980) in *Protides of the Biological Fluids, Proceedings of the 28th Colloquium* (Peeters, H., Ed.) pp 51–56, Pergamon Press, Oxford.
- Henschen, A., Lottspeich, F., Kehl, M., & Southan, C. (1983) *Ann. N.Y. Acad. Sci.* 408, 28–43.
- Kant, J. A., Lord, S. T., & Crabtree, G. R. (1983) *Proc. Natl. Acad. Sci. U.S.A.* 80, 3953–3957.
- Kant, J. A., Fornace, A. J., Jr., Saxe, D., Simon, M. I., McBride, O. W., & Crabtree, G. R. (1985) *Proc. Natl. Acad. Sci. U.S.A.* 82, 2344–2348.
- Koyama, T., Hall, L. R., Haser, W. G., Tonegawa, S., & Saito, H. (1987) *Proc. Natl. Acad. Sci. U.S.A.* 84, 1609–1613.
- Kudryk, B., Okada, M., Redman, C. M., & Blomback, B. (1982) *Eur. J. Biochem.* 125, 673–682.
- Kyte, J., & Doolittle, R. F. (1982) *J. Mol. Biol.* 157, 105–132.
- May, L. T., Ghayeb, J., Santhanam, U., Tatter, S. B., Sthoeger, Z., Helfgott, D. C., Chiorazzi, N., Grieninger, G., & Sehgal, P. B. (1988) *J. Biol. Chem.* 263, 7760–7766.
- May, L. T., Shaw, J. E., Khanna, A. K., Zabriskie, J. B., & Sehgal, P. B. (1991) *Cytokine* 3, 204–211.
- Morel, Y., Bristow, J., Gitelman, S. E., & Miller, W. L. (1989) *Proc. Natl. Acad. Sci. U.S.A.* 86, 6582–6586.
- Nies, D. E., Hemesath, T. J., Kim, J.-H., Gulcher, J. R., & Stefansson, K. (1991) *J. Biol. Chem.* 266, 2818–2823.
- Norenberg, U., Wille, H., Wolff, J. M., Frank, R., & Rathjen, F. G. (1992) *Neuron* 8, 849–863.
- Pan, Y., & Doolittle, R. F. (1992) *Proc. Natl. Acad. Sci. U.S.A.* 89, 2066–2070.
- Plant, P. W., & Grieninger, G. (1986) *J. Biol. Chem.* 261, 2331–2336.
- Rao, S. P. S., Poojary, M. D., Elliott, B. W., Jr., Melanson, L. A., Oriel, B., & Cohen, C. (1991) *J. Mol. Biol.* 222, 89–98.
- Rixon, M. W., Chan, W.-Y., Davie, E. W., & Chung, D. W. (1983) *Biochemistry* 22, 3237–3244.
- Roy, S. N., Mukhopadhyay, G., & Redman, C. M. (1990) *J. Biol. Chem.* 265, 6389–6393.
- Roy, S. N., Procyk, R., Kudryk, B. J., & Redman, C. M. (1991) *J. Biol. Chem.* 266, 4758–4763.
- Sambrook, J., Fritsch, E. F., & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
- Sehgal, P. B., Grieninger, G., & Tosato, G., Eds. (1989) *Ann. N.Y. Acad. Sci.* 557, 1–583.
- Weissbach, L., & Grieninger, G. (1990) *Proc. Natl. Acad. Sci. U.S.A.* 87, 5198–5202.
- Xu, X., & Doolittle, R. F. (1990) *Proc. Natl. Acad. Sci. U.S.A.* 87, 2097–2101.