Li, S.-J., & Marshall, A. G. (1986) *Biochemistry 25*, 3673-3682.
Li, S.-J., Chang, L.-H., Chen, S.-M., & Marshall, A. G. (1984) *Anal. Biochem. 138*, 465-471.
Li, S.-J., Wu, J., & Marshall, A. G. (1987) *Biochemistry 26*, 1578-1585.
Luehrsen, K. R., & Fox, G. E. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 2150-2154.
Luoma, G. A., & Marshall, A. G. (1978a) *Proc. Natl. Acad. Sci. U.S.A. 75*, 4901-4905.
Luoma, G. A., & Marshall, A. G. (1978b) *J. Mol. Biol. 125*, 95-105.
Marshall, A. G., & Wu, J. (1989) *Biol. Magn. Reson.* (in press).
Nishikawa, K., & Takemura, S. (1974) *J. Biochem. (Tokyo) 76*, 935-947.
Osterberg, R., Sjoberg, B., & Garrett, R. A. (1976) *Eur. J. Biochem. 68*, 481-487.
Pieler, T., & Erdmann, V. A. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 4599-4603.
Pieler, T., Digweed, M., & Erdmann, V. A. (1984) in *Gene Expression* (Clark, B. F. C., & Petersen, H. U., Eds.) pp 353-376, Munksgaard, Copenhagen.
Redfield, A. G., Kunz, S. D., & Ralph, E. K. (1975) *J. Magn. Reson. 19*, 114-117.
Reid, B. R. (1981) *Annu. Rev. Biochem. 50*, 969-996.
Reid, B. R., McCollum, L., Ribeiro, N. S., Abbate, J., & Hurd, R. E. (1979) *Biochemistry 18*, 3996-4005.
Rich, A. (1977) *Acc. Chem. Res. 10*, 388-395.
Romby, P., Westhof, E., Toukifimpa, R., Mache, R., Ebel, J.-P., Ehresmann, C., & Ehresmann, B. (1988) *Biochemistry 27*, 4721-4730.
Roy, S., & Redfield, A. G. (1983) *Biochemistry 22*, 1386-1390.
Schimmel, P. R., & Redfield, A. G. (1980) *Annu. Rev. Biophys. Bioeng. 9*, 181-221.
Studnicka, G. M., Eiserling, F. A., & Lake, J. A. (1981) *Nucleic Acids Res. 9*, 1885-1904.
Zhang, P., & Moore, P. B. (1989) *Biochemistry 28*, 4607-4615.

Apolipoprotein(a) Size Heterogeneity Is Related to Variable Number of Repeat Sequences in Its mRNA†

Marlys L. Koschinsky,‡ Ulrike Beisiegel,§ Doris Henne-Bruns,§ Dan L. Eaton,‡ and Richard M. Lawn*,‡

Department of Cardiovascular Research, Genentech, Inc., 460 Point San Bruno Boulevard, South San Francisco, California 94080, and Department of Medicine and Surgery, University of Hamburg, Martinistrasse 52, 2000 Hamburg 20, FRG

Received June 30, 1989; Revised Manuscript Received August 25, 1989

ABSTRACT: Plasma apolipoprotein(a) [apo(a)] shows considerable size heterogeneity, existing as discrete glycoprotein isoform variants that range in apparent molecular mass from approximately 400 to 800 kDa. To study the molecular basis of protein size variability, we have isolated liver RNA from individuals with different apo(a) isoforms, and identified apo(a)-specific transcripts using Northern blot analysis. Transcript sizes were shown to be variable (8.0-12 kb) and in all cases were closely correlated with protein masses (590-850 kDa) as determined from immunoblots. Thus, it is almost certain that apo(a) isoform size variation is due to allelic differences in the number of its tandemly repeated sequences of 114 amino acids that resemble kringle four of plasminogen. The high carbohydrate content of apo(a) makes true molecular weight estimations in SDS-PAGE gels difficult. However, a recombinant form of apo(a) containing 17 kringle repeats (calculated molecular mass of 250 kDa) migrates on SDS-PAGE gels only slightly below apoB-100, with an apparent molecular mass of approximately 500 kDa. Since smaller protein isoforms have been observed in the population, this suggests that plasma apo(a) isoforms contain from less than 17 to greater than 30 tandemly repeated kringle units.

Elevated levels of serum Lp(a)[1] are strongly correlated with atherosclerosis in human populations and Lp(a) has been assessed as an independent risk factor in the development of cardiovascular disease (Rhoads et al., 1986; Dahlen et al., 1986; Durrington et al., 1988). Lp(a) closely resembles low-density lipoprotein (LDL) in both lipid composition and the presence of apolipoprotein B-100 (apoB-100). Unlike LDL, Lp(a) contains an additional protein designated apolipoprotein(a) [apo(a)] which is bound to apoB-100 by a disulfide linkage (Gaubatz et al., 1983; Utermann & Weber, 1983; Armstrong et al., 1985; Fless et al., 1985). Apo(a) is a large glycoprotein that is synthesized primarily in the liver (Tomlinson et al., 1989; Kraft et al., 1989) and exhibits considerable size heterogeneity in the human population (Fless et al., 1984; Gaubatz et al., 1987; Utermann et al., 1987, 1988a,b). Marked variability has also been observed with respect to plasma Lp(a) levels, which range from less than 1 over 100 mg/dL in the population. An inverse relationship between isoform size and mean plasma concentration in the population has been observed (Utermann et al., 1987, 1988a). The fact

† This work was supported by Genentech, Inc., and the Deutsche Forschungsgemeinschaft. M.L.K. is partially funded by a postdoctoral fellowship from the Medical Research Council of Canada.

* To whom correspondence should be addressed.
‡ Genentech, Inc.
§ University of Hamburg.

[1] Abbreviations: apo(a), apolipoprotein(a); apoB-100, apolipoprotein B-100; Lp(a), lipoprotein(a); LDL low-density lipoprotein; SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis.

0006-2960/90/0429-640$02.50/0

that the same gene locus controls both Lp(a) phenotype and concentration has been demonstrated by the tight linkeage between Lp(a) levels and isoform sizes and the apo(a) gene, which lies close to the plasminogen locus on human chromosome 6 (Weitkamp et al., 1988; Frank et al., 1988; Drayna et al., 1988).

On the basis of partial amino acid sequence (Eaton et al., 1987) and the subsequent sequence analysis of cloned human apo(a) cDNA (McLean et al., 1987), apo(a) was found to be remarkably similar to human plasminogen. Plasminogen is comprised of a preactivation peptide followed by five kringle domains and a serine protease domain. Human apo(a) consists of multiple tandem repeats of a sequence that resembles kringle 4 of plasminogen, followed by kringle-5-like and protease-like domains (McLean et al., 1987). On this basis it has been hypothesized that different apo(a) protein isoform sizes reflect expansion or contraction of the tandemly repeated kringle region (McLean et al., 1987). In the current study, we describe the relationship between human apo(a) isoform variants and mRNA transcript sizes and address the correlation between protein size variation and human plasma Lp(a) levels.

Experimental Procedures

Source of Human Liver Samples. Eight human liver samples were obtained from patients undergoing liver resection and immediately frozen in liquid nitrogen. The tissue was stored at −70 °C prior to RNA preparation. Lp(a) levels in the patients ranged from 1 to 60 mg/dL as determined by ELISA (W. L. Wong and D. L. Eaton, personal communication).

RNA Isolation and Northern Blot Analysis. One to two grams of liver was homogenized in the presence of guanidinium thiocyanate (MacDonald, 1987), and total RNA was purified by using lithium chloride extraction (Cathala et al., 1983). Poly(A+) RNA was subsequently isolated by affinity chromatography on poly(U)-Sepharose (Ricca et al., 1981). RNA samples were denatured with formamide and separated by electrophoresis in a 1% agarose gel containing 2% formaldehyde according to the method of Maniatis et al. (1982). Following transfer to nitrocellulose, the RNA blot was hybridized to a 2.2-kb restriction fragment (p119Da18.5) derived from the 5′-end of the human apo(a) cDNA clone (McLean et al., 1987). The DNA fragment was labeled with ^{32}P by using random sequence oligonucleotides as primers for Klenow fragment elongation (Feinberg & Vogelstein, 1982). The blot was hybridized for 16 h at 42 °C, in 50% formamide, 5× SSC (SSC: 150 mM NaCl and 15 mM trisodium citrate), 50 mM sodium phosphate, 5× Denhardt's solution, 10% dextran sulfate, and 20 μg/mL carrier DNA. The filter was subsequently washed at 68 °C for 45 min in 0.2× SSC and 0.1% SDS and autoradiographed with intensifying screens at −80 °C. RNA size standards were obtained from Bethesda Research Laboratories.

Determination of Plasma apo(a) Isoform Sizes. Lp(a) was purified from plasma samples (0.3–1.0 mL) by using antibody affinity chromatography. Briefly, samples were loaded onto 1-mL Affi-Gel (Bio-Rad) columns equilibrated in 50 mM Tris-HCl, pH 7.5, and 150 mM NaCl, to which apo(a)-specific monoclonal antibody (W. L. Wong and D. L. Eaton, personal communication) had been coupled. Following extensive washes with 50 mM Tris-HCl, pH 7.5, and 1 M NaCl to eliminate nonspecifically bound proteins, the apo(a) was eluted by using 0.1 M glycine, pH 2.2. After neutralization using 1 M Tris-HCl, pH 7.5, samples were concentrated by Centricon (Amicon) filtration from 2 mL to 50 μL. Samples were brought to 5% sodium dodecyl sulfate and 10 mM dithiothreitol and electrophoresed in 2.5–10% polyacrylamide gradient gels as described previously (Fless et al., 1984). Following overnight transfer to nitrocellulose, the blot was incubated with a polyclonal antibody directed against human Lp(a) (Drayna et al., 1988) and visualized with an alkaline phosphatase conjugated secondary antibody (Bio-Rad). Pharmacia high molecular weight markers (669 000, 440 000, 330 000, and 220 000) as well as human apoB-100 (550 000; calculated peptide mass from sequence of 519 000 plus 8–10% carbohydrate; Chen et al., 1986; Law et al., 1986) were used for isoform size estimations. Size estimates above 669 000 were determined by extrapolation. A truncated apo(a) molecule (calculated peptide mass of 250 000) containing 17 kringle-4-like domains, as well as the kringle-5-like and protease-like domains, was produced in cultured mammalian cells by recombinant DNA techniques (Koschinsky, Tomlinson, Lawn, and Eaton, unpublished results). Purified recombinant protein (1.5 μg) was electrophoresed in a 2.5–10% polyacrylamide gel and its size estimated relative to purified human LDL.

Results

Variation in apo(a) Transcript Sizes. On the basis of Northern blot analysis (Figure 1A), apo(a) transcript sizes in the current sample range from about 8 to 12 kb. Two transcripts were observed in several individuals (lanes 2, 4, 6, and 7) with sizes of 12.2 and 8.2, 11.7 and 8.3, 9.3 and 8.4, and 12.3 and 11.2 kb, respectively. In each of these cases, the smaller transcript was of greater relative intensity. Under the conditions of stringency used for the hybridization, very low cross-reactivity with plasminogen mRNA (3.0 kb) was observed (see Figure 1A).

For the sake of comparison, a sample of the mRNA used for apo(a) cDNA cloning (McLean et al., 1987) was included in the Northern blot analysis (designated BHL in Figure 1A). The size of this transcript was estimated to be 12.0 kb. RNA isolated from tissue culture cells transfected with recombinant apo(a) was also analyzed, and the size of the transcript agrees with the predicted size of approximately 7.6 kb (data not shown).

To allow more accurate apo(a) transcript size estimates, the Northern blot was subsequently hybridized with a DNA probe specific for human apoB. The size of the latter transcript is considered to be 14.1 kb on the basis of the cloned cDNA sequence (Chen et al., 1986; Law et al., 1986). The position of the apoB transcript is indicated in Figure 1A.

Analysis of apo(a) Isoform Sizes. Estimations of apo(a) protein masses were made on the basis of observed mobilities in polyacrylamide gels (Figure 1B). Isoform size variants were found to range from approximately 590 to 850 kDa. Two isoforms were detected in samples 2, 4, and 7 (590 and 850, 600 and 810, 760 and 850 kDa, respectively), and in each instance, the smaller isoform was clearly more abundant. In all cases, apo(a) protein sizes were found to be larger than apoB-100 (see Figure 1B). In the case of samples 2 and 4, the larger isoform was so much abundant than the lower that it was very faint on the original blot and difficult to see in photographic reproduction. Hence, these samples were rerun with more protein and are shown in lanes 2′ and 4′. In lanes with excessive loadings, some breakdown products of apo(a) are visible.

A recombinant apo(a) protein with 17 copies of the repeated kringle-4-like domain (calculated polypeptide mass = 250 kDa) migrated in polyacrylamide gels only slightly below the position of apoB-100, which has a molecular mass of 550 kDa (data not shown). This result underscores the large effect of glycosylation on the migration of apo(a) in gels and influences

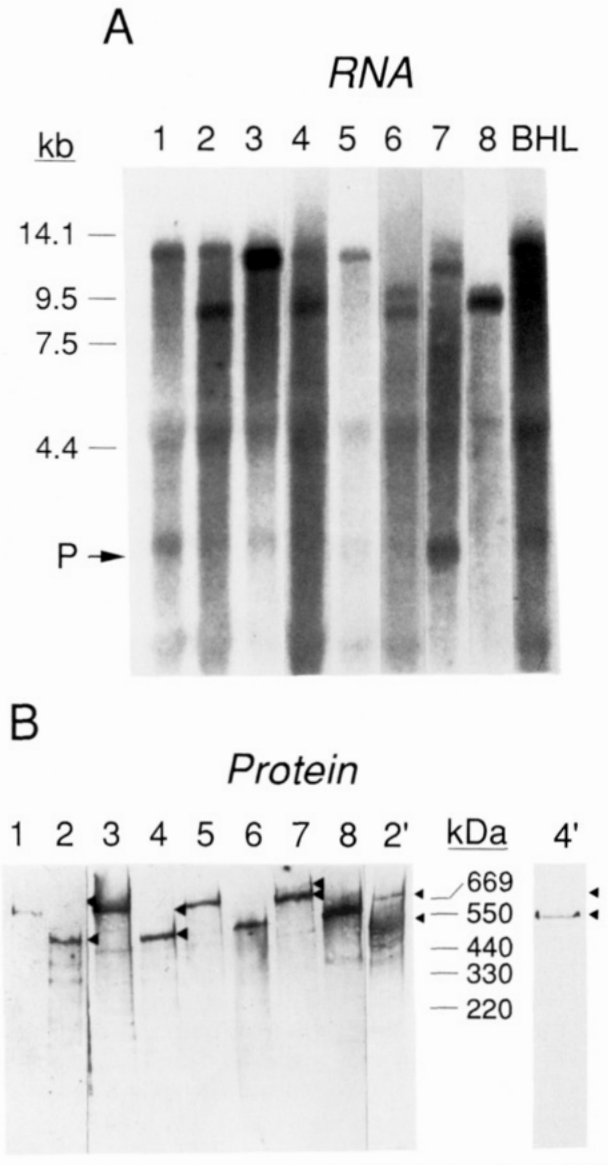

FIGURE 1: (A) Blot hybridization analysis of apo(a) mRNA. Poly(A+) RNA (5–10 μg) isolated from nine unrelated individuals was separated by electrophoresis in a denaturing agarose gel and transferred to nitrocellulose. Lanes 1–8 contain the samples for comparison of RNA and protein size and correspond to lanes 1–8 in (B). The gel also includes RNA (denoted BHL) that was utilized to clone apo(a) cDNA, as described in McLean et al. (1987). The filter was hybridized with a ^{32}P-labeled apo(a) cDNA probe. On the left are indicated the positions of cross-reacting plasminogen mRNA (P; 3 kb) and size standards (Bethesda Research Labs) including the message for apo B (14.1 kb) which was visualized by rehybridizing the filter. All lanes derive from different exposures of a single gel and filter. (B) Immunoblot analysis of apo(a) protein isoforms. Lp(a) was purified from the plasma of eight unrelated individuals and electrophoresed in a 2.5–10% polyacrylamide gel. Following transfer to nitrocellulose, the blot was incubated with an apo(a)-specific antibody. Apo(a) isoforms were then detected by using an alkaline phosphatase conjugated secondary antibody. Arrows point to the position of the two isoforms in samples 2, 4, and 7. Lanes 2, 2′, 4, and 4′ show different loadings of the same sample to underscore the presence of the larger, less abundant isoform. Overloading of some samples reveals degraded forms of apo(a) evidenced by a ladder of sizes that are assumed to be due to cleavage of repeated domains. The positions of apoB-100 (550 kDa) and other size standards (thyroglobulin, ferritin, and their subunits; Pharmacia) are indicated on the right.

the estimate of the range in number of apo(a) kringle repeats in the population.

Relationship between apo(a) Transcript Sizes and Isoform Sizes. A direct correlation was observed between apo(a) protein size and mRNA transcript size as determined by gel mobility (see Figure 2). This analysis was based on 10 protein isoform variants (Figure 1B) for which corresponding transcript sizes were determined (Figure 1A).

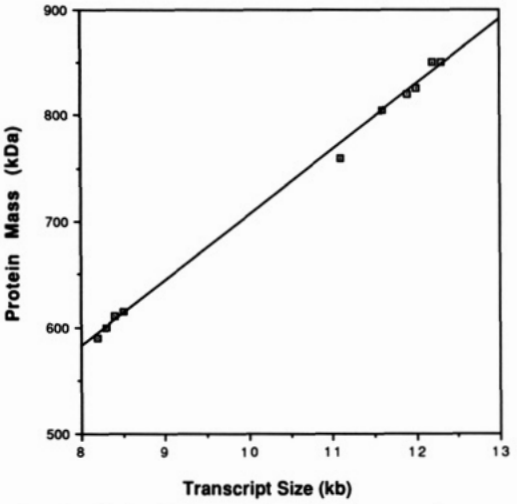

FIGURE 2: Analysis of the relationship between human apo(a) transcript size (kb; *X* axis) and protein mass (kDa; *Y* axis). Ten isoforms are represented for which both determinants could be estimated. Data points are indicated on the graph with boxes. Sizes were determined from the mobility in the gels shown in Figure 1.

DISCUSSION

A number of studies have clearly demonstrated that apo(a) exists in heritable protein size isoforms in the human population (Fless et al., 1984; Gaubatz et al., 1987; Utermann et al., 1987, 1988a,b). However, the molecular basis of this size variation has been unclear. Since apo(a) is a large, highly glycosylated protein (28% carbohydrate by weight; Fless et al., 1986), mechanisms involving posttranslational modification (e.g., variable carbohydrate composition) have been evoked to account for protein size differences. However, on the basis of partial deglycosylation studies, it is apparent that apo(a) size polymorphism does not result entirely from differential glycosylation (Utermann et al., 1987; Fless et al., 1987; W. C. Breckenridge, unpublished results). The discovery of tandemly repeated sequences in apo(a) cDNA suggested that size polymorphism is generated by allelic variation in the number of intragenic repeats, resulting from out of register homologous recombination (McLean et al., 1987). (The repeated domains contain 342 nucleotides, coding for 114 amino acids.) The fact that many of these domains are precisely repeated at the nucleotide level indicates that this variation is a recent evolutionary occurrence and would therefore be evident in the present population.

In the current study, we have shown that human apo(a) protein size variations are the result of different-sized transcripts. A plot of the relationship between transcript size and protein size reveals a positive linear correlation. [While this study was in progress, Hixson et al. (1989) presented similar findings with baboon apo(a) protein and RNA.] Minor deviations in correspondence are probably due to limitations in both protein and RNA size determinations. In one of the four individuals who had two detectable mRNA species, only a single protein could be detected (sample 6). This may be due to low levels of one isoform or to lack of resolution. In the other cases where two isoforms occurs, the large isoform occurs in far lower amounts, necessitating the inclusion of two gel loadings to make them both apparent in Figure 1. The overloaded sample lances contain some visible breakdown products, underscoring some of the difficulties in performing

this study on human samples. Of the limited number of subjects undergoing liver resection available to us, only a subset have high enough levels of apo(a) to make the gel analysis of their RNA and protein trivial.

While the mechanism for producing polymorphism in mRNA size is not known, it is likely to occur at the gene level for several reasons. As stated above, the number, length, and identity of repeated kringle-4-like sequences which are a hallmark of the apo(a) gene make them likely targets for addition or subtraction by homologous recombination. One would not suspect that such a mechanism would affect the nonrepeated kringle-5-like and protease-like sequences. Despite the many tandem repeats of identical exons and nearly identical introns within this gene, differential RNA splicing is not evident. All individuals appear to have zero, one, or two detectable mRNA and protein isoforms, consistent with the faithful splicing and expression of at most two alleles of the apo(a) gene. If this hypothesis is correct, apo(a) gene size should directly reflect differences in transcript sizes. We have attempted to demonstrate variations in the size of the apo(a) gene by pulse-field electrophoresis of genomic DNA. However, results from these experiments are difficult to interpret due to the presence in the human genome of four genes with extensive sequence similarity to apo(a) [which include apo(a), plasminogen, and possible pseudogenes; M. Nagashima, R. Lawn, J. Tomlinson, and J. McLean, personal communication], as well as the difficulty of evaluating the contributions of each of the two apo(a) alleles per individual. Data that could complement our approach would be a correlation of the intensity of bands in Southern blots originating from the kringle sequences versus protein size. However, since gene fragment intensity would represent the sum of the two apo(a) alleles per individual and most often only one protein isoform is detectable, a large number of samples would be necessary to make a statistical correlation on this basis. This is beyond the aim of the current study.

We have shown that recombinant apo(a) protein, which contains 17 kringle-4-like domains, migrates on SDS–PAGE close to the position of apoB-100 and has a mass of approximately 500 kDa. On the basis of this observation, it can be inferred that apo(a) protein isoforms which have been reported to be significantly smaller in size than apoB-100 likely contain fewer than 17 kringle repeat units. The mRNA used in the cloning of apo(a) cDNA (BHL; McLean et al., 1987) was also analyzed in this study. The transcript size was shown to be approximately 12 kb. This differs from the previously reported estimate of 14 kb, due to increased resolving power of the gels employed in this study. On the basis of the original size estimate, it was predicted that the cloned apo(a) cDNA species contained 37 plasminogen-like kringle 4 domains (McLean et al., 1987). Our results thus suggest that the published cDNA clone may actually contain less than 37 tandemly repeated kringle units. (Unfortunately, no plasma was available from this individual for protein isoform analysis.) We do not know what number of repeated kringle domains constitutes "functional" apo(a) and the relative contributions of adaptive selection and equilibrium between the addition and subtraction of repeats. Indeed, the function of apo(a) is unknown at the present. On the basis of isoform analysis by ourselves and other investigators, we can estimate that apo(a) variants with as few as ~15 and as many as ~37 kringle repeats can be found in human plasma.

Although not designed to be quantitative, this study affords some observations with respect to the relationship between isoform levels and transcript levels. In the case of heterozygote expression, a consistent correlation was observed between isoform levels and mRNA transcript abundance. For example, the smaller protein isoform in the case of samples 2, 7, and presumably 4 is significantly more abundant than the larger form. This relationship is also visible in the relative amounts of the two transcripts for samples 2, 4, and 7. However, this sample size is small. Exceptions to the direct correlation between transcript levels and plasma protein concentrations were detected in a study of 22 baboons (Hixson et al., 1989). This was interpreted to suggest that apo(a) plasma concentrations reflect both transcriptional and posttranscriptional levels of control (Hixson et al., 1989). We feel that this is likely to be correct in light of the observations that although the plasma concentration of apo(a) bears an inverse relationship to isoform size, there is a large degree of scatter about the mean (Utermann et al., 1988a). Since there is no reason to suspect that the number of kringle repeats in the gene affects promoter strength, it seems reasonable to presume that the observed variation reflects differences at the level of transcription, which are superimposed upon effects occurring at the level of message stability, translation, particle assembly, secretion, or plasma clearance. As an end result, concentrations of apo(a) forms with more kringle repeats are relatively reduced in the circulation. Direct studies of synthesis, assembly, and catabolism are clearly required to fully characterize the regulation of Lp(a) levels in plasma.

ACKNOWLEDGMENTS

In addition to useful discussions, we thank James Tomlinson for his part in the construction, purification, and characterization of recombinant apo(a), Wai-Lee Wong for performing ELISA assays, and James Tomlinson, John McLean, and Mariko Nagashima for use of their unpublished results.

REFERENCES

Armstrong, V. W., Walli, A. K., & Seidel, D. (1985) *J. Lipid Res. 26*, 1314–1323.

Cathala, G., Savouret, J.-F., Mendez, B., West, B. L., Karin, M., Martial, J. A., & Baxter, J. D. (1983) *DNA 2*, 329–335.

Chen, S. H., Yang, C. Y., Chen, P. F., Setzer, D., Tanimura, M., Li, W. H., Gotto, A. M. & Chan, L. (1986) *J. Biol. Chem. 261*, 12918–12921.

Dahlen, G. H., Guyton, J. R., Attar, M., Farmer, J. A., Kautz, J. A., & Gotto, A. M. (1986) *Circulation 74*, 758–765.

Drayna, D. T., Hegele, R. A., Hass, P. E., Emi, M., Wu, L. L., Eaton, D. L., Lawn, R. M., Williams, R. R., White, R. L., & Lalouel, J. M. (1988) *Genomics 3*, 230–236.

Durrington, P. N., Ishola, M., Hunt, L., Arrol, S., & Bhatnagar, D. (1988) *Lancet No. 1*, 1070–1073.

Eaton, D. L., Fless, G. M., Kohr, W. J., McLean, J. W., Xu, Q. T., Miller, C. G., Lawn, R. M., & Scanu, A. M. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 3224–3228.

Feinberg, A., & Vogelstein, B. (1982) *Anal. Biochem. 132*, 6.

Fless, G. M., Rolih, C. A., & Scanu, S. M. (1984) *J. Biol. Chem. 259*, 11470–11478.

Fless, G. M., ZumMallen, M. E., & Scanu, A. M. (1985) *J. Lipid Res. 26*, 1224–1229.

Fless, G. M., ZumMallen, M. E., & Scanu, A. M. (1986) *J. Biol. Chem. 261*, 8712–8718.

Fless, G. M., ZumMallen, M. E., & Scanu, A. M. (1987) *Proceedings of Workshop on Lipoprotein Heterogeniety* (Lippel, K., Ed.) NIH Publication 87-2646, pp 141–153, U.S. GPO, Washington, DC.

Frank, S. L., Klisak, K., Sparkes, R. S., Mohandas, T., Tomlinson, J. E., McLean, J. W., Lawn, R. M., & Lusis,

A. J. (1988) *Hum. Genet. 79*, 352-356.
Gaubatz, J. W., Heideman, C., Gotto, A. M., Jr., Morrisett, J. D., & Dahlen, G. H. (1983) *J. Biol. Chem. 258*, 4582-4589.
Gaubatz, J. W., Chari, M. V., Nava, M. L., Guyton, J. R., & Morrisett, J. D. (1987) *J. Lipid Res. 28*, 69-79.
Hixson, J. E., Rainwater, D. L., Britten, M. L., & Manis, G. S. (1989) *J. Biol. Chem. 264*, 6013-6016.
Kaufman, R. J., & Sharp, P. A. (1982) *Mol. Cell. Biol. 2*, 1304-1319.
Kraft, H. G., Menzel, H. J., Hoppichler, F., Vogel, W., & Utermann, G. (1989) *J. Clin. Invest. 83*, 137-142.
Law, S. W., Grant, S. M., Higuchi, K., Hospattankar, A., Lackner, K., Lee, N., & Brewer, H. B., Jr. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 8142-8146.
MacDonald, R. J., Swift, G. H., Pryzybyla, A. W., & Chirgwin, J. M. (1987) *Methods Enzymol. 152*, 219-227.
Maniatis, T., Fritsch, E. F., & Sambrook, J. (1982) *Molecular Cloning. A Laboratory Manual*, Cold Spring Harbor Laboratories, Cold Spring Harbor, NY.
McLean, J. W., Tomlinson, J. E., Kuang, W.-J., Eaton, D. L., Chen, E. Y., Fless, G. M., Scanu, A. M., & Lawn, R. M. (1987) *Nature 330*, 132-137.
Rhoads, G. G., Dahlen, G., Berg, K., Morton, N. E., & Dannenberg, A. L. (1986) *JAMA, J. Am. Med. Assoc. 256*, 2540-2544.
Ricca, G. A., Hamilton, R. W., McLean, J. W., Conn, A., Kalinyak, J. E., & Taylor, J. M. (1981) *J. Biol. Chem. 256*, 10362-10368.
Segrest, J. P., & Jackson, R. L. (1972) *Methods Enzymol. 28*, 54-63.
Tomlinson, J. E., McLean, J. W. M., & Lawn, R. M. (1989) *J. Biol. Chem. 264*, 5957-5965.
Utermann, G., & Weber, W. (1963) *FEBS Lett. 154*, 357-361.
Utermann, G., Menzel, H. J., Kraft, H. G., Duba, H. C., Kemmler, H. G., & Seitz, C. (1987) *J. Clin. Invest. 80*, 458-465.
Utermann, G., Kraft, H. G., Menzel, H. J., Hopferwieser, T., & Seitz, C. (1988a) *Hum. Genet. 78*, 41-46.
Utermann, G., Duba, C., & Menzel, H. J. (1988b) *Hum. Genet. 48*, 47-50.
Weitkamp, L. R., Guttormsen, S. A., & Schultz, J. S. (1988) *Hum. Genet. 79*, 80-82.

Proteolytic Formation and Properties of a Fragment of Protein C Containing the γ-Carboxyglutamic Acid Rich Domain and the EGF-like Region†

Ann-Kristin Öhlin,*,‡ Ingemar Björk,§ and Johan Stenflo‡

Department of Clinical Chemistry, University of Lund, Malmö General Hospital, S-214 01 Malmö, Sweden, and Department of Veterinary Medical Chemistry, Swedish University of Agricultural Sciences, The Biomedical Center, Box 575, S-751 23, Uppsala, Sweden

Received July 14, 1989

ABSTRACT: The function of the epidermal growth factor (EGF) like domains in the vitamin K dependent plasma proteins is largely unknown. In order to elucidate the function of these domains in protein C, we have devised a method to isolate the EGF-like region from the light chain connected to the NH_2-terminal region, containing the γ-carboxyglutamic acid (Gla) residues. This was accomplished by tryptic cleavage of protein C that had been reversibly modified with citraconic anhydride to prevent cleavage at the lysine residue (in position 43) that is located between the two regions. The isolated fragment consists of residues 1-143 from the light chain of protein C connected by a disulfide bond to residues 108-131 from the heavy chain. Upon Ca^{2+} binding to the isolated Gla-EGF fragment from bovine protein C, the tryptophan fluorescence emission was quenched in a manner indicating binding to at least two classes of binding sites. These were presumably the Gla-independent Ca^{2+}-binding site located in the EGF-like region and the lower affinity sites in the Gla region. A comparison with the tryptophan fluorescence quenching that occurred upon Ca^{2+} binding to the separately isolated EGF-like and Gla regions suggested that the EGF-like region influenced the structure and Ca^{2+} binding of the Gla region. The isolated Gla-EGF fragment functioned as an inhibitor of the anticoagulant effect of activated protein C in a clotting assay, whereas no inhibition was observed with either the Gla region or the EGF-like region.

Both coagulant and anticoagulant systems are required to maintain normal hemostasis. The protein C system can be regarded as an anticoagulant counterpart to the blood clotting cascade, and it regulates the activation rates of factor X and prothrombin (Clouse & Comp, 1986; Stenflo, 1988; Esmon, 1989). The key protein of this system, protein C, is a vitamin K dependent serine protease zymogen that is activated by thrombin bound to the endothelial cell membrane protein thrombomodulin (Esmon et al., 1982). Activated protein C, in conjunction with its cofactor, protein S, rapidly inactivates factors Va and VIIIa by limited proteolysis (Walker et al., 1979; Vehar & Davie, 1980; Marlar et al., 1982; Suzuki et al., 1983). The role of protein C as an anticoagulant in vivo is shown by the fact that a deficiency of the protein leads to an increased risk of developing thrombosis (Griffin et al., 1981).

Protein C, like the other vitamin K dependent plasma proteins, is composed of domains with discrete structure and

† This investigation was supported by grants from the Swedish Medical Research Council (Projects 4487 and 4212), Albert Pålsson's Foundation, Kock's Foundation, and Österlund's Foundation.
* To whom correspondence should be addressed.
‡ Malmö General Hospital.
§ The Biomedical Center.

0006-2960/90/0429-644$02.50/0