Vitamin K-dependent carboxylase: Possible role of the substrate "propeptide" as an intracellular recognition site

 $(\gamma$ -carboxyglutamic acid/protein C)

J. W. Suttie^{*†}, J. A. Hoskins[‡], J. Engelke^{*}, A. Hopfgartner^{*§}, H. Ehrlich[¶], N. U. Bang[¶], R. M. Belagaje[‡], B. Schoner[‡], and G. L. Long^{‡||}

*Department of Biochemistry, College of Agricultural and Life Sciences, University of Wisconsin-Madison, Madison, WI 53706; ⁴Lilly Laboratory for Clinical Research and Department of Medicine, Indiana University School of Medicine, Indianapolis, IN 46202; and [‡]Division of Molecular Biology, Lilly Research Laboratories, Indianapolis, IN 46285

Communicated by Henry Lardy, September 15, 1986

ABSTRACT The liver microsomal vitamin K-dependent carboxylase catalyzes the posttranslational conversion of specific glutamate residues to γ -carboxyglutamate residues in a limited number of proteins. A number of these proteins have been shown to contain a homologous basic amino acid-rich "propeptide" between the leader sequence and the amino terminus of the mature protein. Plasmids encoding protein C. a vitamin K-dependent protein, containing or lacking a propeptide region were constructed and the protein was expressed in Escherichia coli. The protein products were assayed as substrates in an in vitro vitamin K-dependent carboxylase system. Only proteins containing a propeptide region were substrates for the enzyme. These data support the hypothesis that this sequence of the primary gene product is an important recognition site for this processing enzyme.

The liver microsomal vitamin K-dependent carboxylase catalyzes the posttranslational conversion of glutamate to γ -carboxyglutamate residues in intracellular precursors of a limited number of proteins (1). These proteins include the classical vitamin K-dependent clotting factors (II, VII, IX, and X) as well as two more recently discovered antithrombotic plasma proteins, protein C and protein S (2, 3). The interactions by which this enzyme recognizes its substrate have not been identified. Although studies with low molecular weight peptide substrates have revealed some aspects of the specificity of the glutamate binding site (4), it is likely that the specificity needed to limit carboxylation to a selected group of proteins is a function of other, non-active-site, interactions.

The intracellular, precursor, form of prothrombin (factor II) has been shown (5) to be larger and more basic than the secreted plasma form. Structural information obtained through cDNA sequencing has revealed that the intracellular precursors of prothrombin (6, 7), factor X (8), factor IX (9–11), and protein C (12–14) all contain a basic amino acid-rich "propeptide" between the signal peptide region and the amino terminus of the plasma form of the protein. A homologous amino-terminal extension has also been observed (15, 16) in the precursor form of a vitamin K-dependent bone protein that shows no other sequence homology to the plasma vitamin K-dependent proteins. These observations suggest that this amino-terminal extension, which was observed earlier in proalbumin (17, 18), might be an important recognition site for this carboxylation system.

MATERIALS AND METHODS

Preparation of Bovine Prothrombin Precursor and Abnormal (Des- γ -carboxy) Prothrombin. Bovine prothrombin was

purified as described by Mann (19) and injected into 8-weekold New Zealand White rabbits to obtain an antiserum to bovine prothrombin. A crude IgG fraction of this serum was used to prepare a calcium-independent anti-prothrombin agarose column as described by Swanson and Suttie (20). A 1-month-old calf was given 2.0 mg of dicumarol per kg of body weight at 0, 2, and 4 days and killed on day 6. The liver was homogenized in 2 ml of SI buffer (250 mM sucrose/25 mM imidazole HCl, pH 7.3) per g and centrifuged at 10,000 \times g for 20 min. The supernate was centrifuged at 105,000 \times g for 60 min to obtain a crude microsomal pellet and the pellet was suspended (1 ml/g of liver) in SIK buffer (SI buffer plus 0.5 M KCl) containing 2% Triton X-100 and 1 mM phenylmethylsulfonyl fluoride by using a loose-fitting Dounce homogenizer. The partially solubilized suspension was centrifuged at $105,000 \times g$ for 60 min to remove nonsolubilized material and the supernatant was applied to the anti-prothrombin agarose. After incubation overnight at 4°C, unbound protein was washed off the gel with SIK buffer until the absorbance at 280 nm and the thrombin generated with Echis carinatus venom (20) was zero. The gel was then washed with water and the prothrombin precursor was eluted from the agarose with 8 M urea, pH 3.5. The urea eluate was dialyzed against 0.02 M Tris HCl, pH 7.0/0.1 M NaCl overnight in the cold room and concentrated using a Diaflo XM 50 ultrafiltration membrane. Plasma from the same calf was treated with BaSO₄, and plasma des- γ -carboxyprothrombin was isolated from the supernate in the same manner.

Bacterial Strains, Plasmids, and Enzymes. Escherichia coli K-12 RV308 (37) (su^- , Δlac -X74, gal IS II::OP308, strA) was the host strain for plasmid constructions and protein expression. The structure of the starting plasmid, pCZ460, containing DNA encoding human protein C is shown in Fig. 1. A more detailed description of its construction is published elsewhere (23). Restriction enzymes were purchased from Bethesda Research Laboratories or New England Biolabs and were used under conditions recommended by the suppliers.

Kinase Treatment, Annealing, and Cloning. Deoxyribonucleotide oligomers were synthesized by using a DNA synthesizer (Applied Biosystems model 380A; Foster City, CA) according to the procedure recommended by the manufacturer. The oligonucleotides were purified by acrylamide gel electrophoresis using standard procedures. Fifty picomoles of each single-stranded deoxyribonucleotide was treated with 10 units of polynucleotide kinase (New England Nuclear) in 50 mM Tris HCl, pH 7.6/10 mM MgCl₂/5 mM dithiothreitol/ 0.1 mM spermidine/0.1 mM EDTA containing 1 μ Ci (1 Ci =

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

[†]To whom reprint requests should be addressed.

Spresent address: Institute for Biochemistry, Schlögelgasse 9, A-8010 Graz, Austria.

¹¹Present address: Department of Biochemistry, University of Vermont College of Medicine, Burlington, VT 05405.

Biochemistry: Suttie et al.



FIG. 1. Structure of pCZ460. The darkened area represents the constitutive *E. coli lpp* promoter, the striped region corresponds to the human protein C gene (beginning with the 10-residue propeptide sequence). The arrows represent the position and direction of transcription for the kanamycin phosphotransferase gene, which confers resistance to kanamycin, and the temperature-sensitive *copB* gene (21), which assists in controlling plasmid copy number. The vector is derived from the plasmid pIMIA (21), which contains the thermoinducible runaway replicon from pKN402 (22).

37 GBq) of $[\alpha^{-32}P]ATP$ for 30 min at 37°C. The linkers were annealed by mixing equal-molar amounts of single-stranded complementary deoxyribonucleotides and heating at 90°C for 10 min, then cooling slowly to 4°C. A 10-fold molar excess of the annealed linkers was ligated to the appropriate restriction fragment (shown in Fig. 2) at 4°C overnight by using T4 DNA ligase (Bethesda Research Laboratories).

Construction of Plasmids pCZ464 and pCZ465 and Isolation of Plasmid DNA. The starting plasmid, pCZ460, was digested to completion with Nde I to produce a linear molecule. As pCZ460 contains two Sst I sites (Fig. 1), it was necessary to digest the DNA with Sst I under conditions that would give only partial digestion. The digested DNA was treated with phenol to stop the enzyme reaction, precipitated in ethanol, and collected by centrifugation. The DNA was ligated as described above. Similar conditions were used for the construction of pCZ465: pCZ460 was digested to completion with Nde I, then treated with Sal I in such a way as to result in only partial digestion (pCZ460 contains four Sal I restriction sites). After extraction with phenol and precipitation with ethanol the DNA was ligated as described above. Ligated DNA was used to transform E. coli RV308 by the standard CaCl₂ procedure (24). Transformed cells were selected by growth on TY medium [10 g Bacto-tryptone (Difco)/5 g Bacto-yeast extract (Difco)/5 g NaCl/15 g agar (Difco) per 1 liter of distilled H₂O] containing 50 µg of kanamycin per ml at 25°C for 2 days. Isolated colonies were inoculated into 1.5 ml of TY broth (with kanamycin) and grown overnight, with shaking, at room temperature. Plasmid DNA was isolated by using the technique of Birnboim and Doly (25). Correct constructions were identified by restriction enzyme analysis and confirmed by DNA sequencing (26) across the junction regions.

Expression of Protein Products and Isolation of Granules. An aliquot of transformed cells from a 50% (vol/vol) glycerol stock was diluted into TY broth containing kanamycin and allowed to grow overnight at room temperature with vigorous shaking. To amplify the synthesis of protein C this stationary-phase culture was diluted 1:50 into TY broth containing kanamycin at 50 μ g/ml and cysteine at 40 mg/ml and shifted to 37°C. The regulation of the plasmid copy number is temperature sensitive. At room temperature only about 10 copies are present per cell; at 37°C the plasmid replication becomes uncontrolled and about 2000 copies of the plasmid accumulate per cell. The *lpp* promoter is constitutive, so the amount of protein produced per cell depends directly upon the number of plasmid molecules present. Certain proteins, including protein C, when overproduced in *E. coli* will accumulate in the cell in the form of granules. After 6–8 hr at 37°C granules were visible within the cell cytoplasm when a phase-contrast microscope (×970) was used. These granules were isolated by using the technique of Schoner *et al.* (27).

Refolding of Human Protein C. The granule suspension, in distilled water, was chilled to 4°C and the pH was quickly raised to 12.5 by dropwise addition of 2.5 M NaOH. The samples were stirred at 4°C for 15 min and dithiothreitol was added to 2.5 mM and the pH was brought to 10.2 by the dropwise addition of 1.2 M HCl. The solubilized protein was stirred rapidly overnight at 4°C in an open tube to allow air oxidation and intramolecular disulfide bond formation. Remaining free cysteine side chains were blocked by the addition of sodium tetrathionate to 10 mM (28), and the sample was dialyzed at 4°C over a period of 24 hr against 20 mM Tris·HCl/20 mM NaCl/4 mM CaCl₂, pH 7.4. Any precipitates formed during the dialysis were removed by slow-speed centrifugation and the soluble protein C was adsorbed and concentrated on a murine monoclonal IgG antibody to the heavy chain of human protein C coupled to Affi-Gel 10 (Bio-Rad) (29). Protein C was eluted from the column by washing with 20 mM Tris·HCl/20 mM NaCl/4 mM EDTA, pH 7.4; the elution was monitored by measuring absorbance at 260 and 280 nm.

Vitamin K-Dependent Carboxylase Assays. A crude microsomal fraction was obtained from vitamin K-deficient rats that had been given 1 mg of vitamin K intracardially 15 min prior to sacrifice as previously described (30). The microsomal pellet was suspended in SIK buffer containing 1.5% Triton X-100 and 2 mM phenylmethylsulfonyl fluoride so that 1 ml of microsomal suspension was equivalent to 0.5 g of liver. Incubation mixtures consisted of 400 μ l of microsomal suspension, 100 μ l of buffer or protein substrate, and 10 μ Ci of NaH¹⁴CO₃ (55 mCi/mmol) in 10 μ l, and reactions were started by the addition of 10 μ l of vitamin KH₂ at 50 μ g/ml in ethanol. Samples were incubated for the times indicated and the reactions were stopped by adding 150 μ l of incubation mixture to 2 ml of cold trichloroacetic acid in a scintillation vial. Free ¹⁴CO₂ was removed by gentle boiling for 2 min before 15 ml of Aquasol (New England Nuclear) was added and radioactivity was determined in a liquid scintillation spectrometer.

RESULTS

When the intracellular bovine prothrombin precursor isolated from a coumarin anticoagulant-treated calf was compared to the des- γ -carboxy form of prothrombin that was circulating in the plasma of the same animal, only the intracellular form was found to be a significant substrate for the rat liver vitamin K-dependent microsomal carboxylase (Table 1). These data suggested that there are some structural differences in these two potential substrates for the enzyme. The basic amino acid-rich propeptide of the intracellular form is lacking in the plasma form, and it is possible that this structural element provides an important interaction needed for efficient carboxylation. To directly test this hypothesis the protein C gene was cloned and expressed in *E. coli*, where the γ -carboxylation of glutamate residues is not known to occur.

Three plasmids that differ only in the amino-terminal sequences they encode were constructed for the expression of protein C derivatives. These sequences were based on the

 Table 1.
 Substrate activity of microsomal prothrombin precursor and abnormal plasma prothrombin

	¹⁴ CO ₂ fixed, dpm		
Addition	Exp. A	Exp. B	
None	2230	1300	
Prothrombin precursor	4990	3670	
Abnormal prothrombin	2200	1260	

Incubations were for 55 min and the values are means of duplicate 150-µl aliquots taken from 530-µl incubation mixtures for two separate experiments. Prothrombin precursor or abnormal prothrombin was added to give 10 National Institutes of Health units.

known amino acid sequence deduced from the published cDNA sequences. Plasmid pCZ460 codes for a protein molecule containing a presumed 10-amino acid residue propeptide (Ala-His-Gln-Val-Leu-Arg-Ile-Arg-Lys-Arg) attached to the amino terminus of a sequence that is otherwise identical to that of single-chain protein C (Fig. 2). Similarly, plasmids encoding molecules lacking the propeptide and corresponding to single-chain protein C (pCZ464) and molecules lacking the 41-residue γ -carboxyglutamate region (pCZ465) were constructed (Fig. 2). These plasmids were used to transform E. coli RV308 and were allowed to amplify at 37°C for maximal synthesis of protein C, resulting in the formation of visible intracellular protein granules. When the crude protein obtained from the granules by alkaline extraction was used as a substrate for the carboxylase, only the protein from the cells containing the pCZ460 plasmid was able to act as a substrate for the enzyme (Table 2). Because of the presence of small amounts of endogenous substrates, some increase in fixed ${}^{14}CO_2$ with time above that seen in the absence of vitamin K was observed even in incubation mixtures containing no added protein. However, only protein extracted from granules obtained from cells containing the pCZ460 plasmid was able to stimulate this activity.

A portion of the crude granular protein extract was adsorbed to a monoclonal IgG antibody, and when protein purified in this manner was used as a substrate similar results were obtained (Table 3). Only the material containing the basic amino acid-rich amino-terminal extension (pCZ460) exhibited a carboxylase substrate activity significantly above background.

DISCUSSION

Isolated plasma abnormal (des- γ -carboxy) prothrombin has been found not (31) to be a good substrate for the vitamin K-dependent carboxylase. The data presented here confirm previous reports (32, 33), which demonstrated that a rat liver

 Table 2.
 Substrate activity of granular protein C expressed in

 E. coli
 E. coli

Exp.	Granular F extract a source	Protein added.	¹⁴ CO ₂ fixed, dpm		
		μg	0 min	20 min	60 min
A	Control	0	1200	1530	1750
	pCZ460	59	1120	1960	2820
	pCZ464	61	1010	1270	1670
	pCZ465	43	980	1230	1600
В	Control	0	1120	1650	1870
	pCZ460	22	1160	2100	2630
	pCZ464	23	1090	1450	1640
	pCZ465	16	1153	1500	1540

Values are means of duplicate 150-µl aliquots taken from a single 530-µl incubation mixture containing each extract at the times indicated. Controls incubated for 60 min with no vitamin K added had 1240 dpm in Exp. A and 1230 dpm in Exp. B. Different granular extracts were used in Exps. A and B.



FIG. 2. Structure of the amino terminus of the human protein C mutants, pCZ464 and pCZ465. (A) DNA sequence of the 5' end of the human protein C-encoding region found in pCZ460 along with the corresponding amino acid sequence (shown in one-letter code). The numbers refer to the amino acid sequence, with number 1 corresponding to the first amino acid present in the mature plasma protein light chain. The recognition sequences of the restriction enzymes are underlined. The nine glutamate residues (E) shown in this sequence are modified to γ -carboxyglutamate residues in the mature plasma protein. (B and C) Chemically synthesized linkers used in the construction of the two deletion mutants. bp, Base pairs. The linkers replaced the region between the Nde I and Sst I site (to create pCZ464) or between the Nde I and Sal I site (to create pCZ465) in pCZ460. The resulting DNA and amino acid sequence for each mutant are shown in relationship to the plasmid vector fragment into which it was ligated. The small letters represent nucleotides deleted in the final constructions. The regions of the precursor plasmid (pCZ460) that were incorporated into the synthesized linkers are underlined.

microsomal prothrombin precursor was a better substrate for the carboxylase than isolated bovine abnormal prothrombin.

Table 3. Substrate activity of protein C expressed in *E. coli* and then purified with antibody

Protein	Protein added.	¹⁴ CO ₂ fixed, dpm		Specific activity.
source	μg	Total	Net	dpm/mg
Control	0	10,240	_	
pCZ460	40	17,320	7080	177,000
pCZ464	60	10,640	400	6,670

Values are means of assays of duplicate aliquots from a 60-min incubation and are calculated on the basis of the total volume of the incubation mixture. A control incubated for 60 min with no vitamin K added had 6910 dpm.

Similar results were seen in this study when both preparations were obtained from the same species. As there are differences in carbohydrate content of these two proteins (5), and since it is possible that the plasma abnormal prothrombin is partially carboxylated, there may be more than one structural difference in these two potential carboxylase substrates. The E. coli-expressed proteins therefore represent a clearer test of the importance of a basic amino acid-rich terminal region for efficient carboxylation. Although the total vitamin K-dependent incorporation of ¹⁴CO₂ in these incubations was low, the data clearly indicate that only the protein preparation containing this region (pCZ460) was able to act as a substrate for the enzyme. Because of the extensive disulfide bond interactions in protein C (34) it is unlikely that more than a minor fraction of the granule extract protein is in the native protein C conformation after alkaline extraction, reduction, and reoxidation. In spite of this, a fraction of the crude granule material could be antibody purified, and substrate activity was again shown to be dependent on a basic amino terminal extension.

Direct protein sequencing of the vitamin K-dependent bone γ -carboxyglutamate protein intracellular precursor (15) has now shown that its propeptide contains 26 residues. The characterization of two different mutations of factor IX (35, 36) indicates that the propeptide protein is 18 residues long, and it has been suggested (36) that the protein C propeptide should consist of 21 residues. The partial propeptide sequence utilized in this study did, however, appear to impart substrate activity to a molecule that contained the appropriate glutamate sites for carboxylation but was not carboxylated. These data provide no indication of whether the influence of this region is to alter the conformation of the γ -carboxyglutamate region of the substrate to increase its interaction with the enzyme or if there is a non-substratebinding site on the enzyme that interacts with this basic region of the molecule, resulting in an increase in inherent catalytic activity or stability. These in vitro probes of the importance of this region of the primary gene product to the intracellular processing of these proteins cannot, of course, provide a clear explanation of the in vivo necessity of this portion of the molecule. These questions will best be answered by appropriate mutagenesis and expression of altered protein in a mammalian cell system.

Note Added in Proof. Amino acid sequencing of the antibody-purified material from the cells containing the pCZ460 plasmid confirmed the presence of the anticipated propeptide region.

This research was supported by the College of Agricultural and Life Sciences, University of Wisconsin–Madison, in part by Grant AM-14881 from the National Institutes of Health, and in part by Lilly Research Laboratories.

- 1. Suttie, J. W. (1985) Annu. Rev. Biochem. 54, 459-477.
- Jackson, C. M. & Nemerson, Y. (1980) Annu. Rev. Biochem. 49, 765-811.
- Suttie, J. W. (1983) in *Plasma Protein Secretion by the Liver*, eds. Glaumann, H., Peters, T., Jr., & Redman, C. (Academic, London), pp. 375-403.
- Rich, D. H., Lehrman, S. R., Kawai, M., Goodman, H. L. & Suttie, J. W. (1981) J. Med. Chem. 24, 706-711.
- Swanson, J. C. & Suttie, J. W. (1985) Biochemistry 24, 3890– 3897.
- Degen, S. J. F., MacGillivray, R. T. A. & Davie, E. W. (1983) Biochemistry 22, 2087-2097.
- 7. MacGillivray, R. T. A. & Davie, E. W. (1984) Biochemistry 23, 1626-1634.
- Fung, M. R., Campbell, R. M. & MacGillivray, R. T. A. (1984) Nucleic Acids Res. 12, 4481-4492.
- Anson, D. S., Choo, K. H., Rees, D. J. G., Giannelli, F., Gould, K., Huddleston, J. A. & Brownlee, G. G. (1984) *EMBO J.* 3, 1053–1064.
- 10. Kurachi, K. & Davie, E. W. (1982) Proc. Natl. Acad. Sci. USA 79, 6461-6464.
- Jayne, M., de la Salle, H., Schamber, F., Balland, A., Kohli, V., Findeli, A., Tolstoshev, P. & Lecocq, J.-P. (1983) Nucleic Acids Res. 11, 2325-2335.
- Long, G. L., Belagaje, R. M. & MacGillivray, R. T. A. (1984) Proc. Natl. Acad. Sci. USA 81, 5653-5656.
- Beckmann, R. J., Schmidt, R. J., Santerre, R. F., Plutzky, J., Crabtree, G. R. & Long, G. L. (1985) Nucleic Acids Res. 13, 5233-5247.
- Foster, D. C., Yoshitake, S. & Davie, E. W. (1985) Biochemistry 82, 4673-4677.
- Pan, L. C., Williamson, M. K. & Price, P. A. (1985) J. Biol. Chem. 260, 13398-13401.
- Pan, L. C. & Price, P. A. (1985) Proc. Natl. Acad. Sci. USA 82, 6109-6113.
- 17. Russell, J. H. & Geller, D. M. (1973) Biochem. Biophys. Res. Commun. 55, 239-245.
- Judah, J. D., Gamble, M. & Steadman, J. H. (1973) Biochem. J. 134, 1083-1091.
- 19. Mann, K. G. (1976) Methods Enzymol. 45, 123-156.
- Swanson, J. C. & Suttie, J. W. (1982) Biochemistry 21, 6011– 6018.
- Masui, Y., Coleman, J. & Inouye, M. (1983) in Experimental Manipulation of Gene Expression, ed. Iouye, M. (Academic, New York), pp. 15-32.
- 22. Uhlin, B. E., Molin, S., Gustafsson, P. & Nordstrom, K. (1979) Gene 6, 91-106.
- 23. Schoner, B., Belagaje, R. & Schoner, R. (1987) Methods Enzymol., in press.
- Wensink, P. C., Finegan, D. J., Donalson, J. E. & Hogness, D. S. (1974) Cell 3, 315–325.
- 25. Birnboim, H. C. & Doly, J. (1979) Nucleic Acids Res. 7, 1513-1523.
- 26. Maxam, A. M. & Gilbert, W. (1980) Methods Enzymol. 65, 499-560.
- Schoner, R. G., Ellis, L. F. & Schoner, B. E. (1985) Bio/ Technology 3, 151–154.
- 28. Means, G. E. & Feeney, R. E. (1971) Chemical Modification of Proteins (Holden-Day, San Francisco, CA), p. 159.
- Comp, P. C., Nixon, R. R., Cooper, M. R. & Esmon, C. T. (1984) J. Clin. Invest. 74, 2082–2088.
- Shah, D. V. & Suttie, J. W. (1983) Proc. Soc. Exp. Biol. Med. 173, 148-152.
- Soute, B. A. M., Vermeer, C., De Metz, M., Hemker, H. C. & Lijnen, H. R. (1981) Biochim. Biophys. Acta 676, 101-107.
- 32. Shah, D. V., Swanson, J. C. & Suttie, J. W. (1983) Arch. Biochem. Biophys. 222, 216-221.
- Evans, M. R., Sung, M. W. P. & Esnouf, M. P. (1984) Biochem. Soc. Trans. 12, 1051–1052.
- Fernlund, P., Stenflo, J. & Tufvesson, A. (1978) Proc. Natl. Acad. Sci. USA 75, 5889–5892.
- Diuguid, D. L., Rabiet, M. J., Liebman, H. A., Furie, B. C. & Furie, B. (1985) Blood 66, 333a (abstr.).
- 36. Bentley, A. K., Rees, D. J. G., Rizza, C. & Brownlee, G. G. (1986) Cell 45, 343-348.
- 37. Maurer, R., Meyer, B. J. & Ptashne, M. (1980) J. Mol. Biol. 139, 147-161.