Published in final edited form as:

J Proteome Res. 2013 June 7; 12(6): 2365–2374. doi:10.1021/pr3003722.

Methylation of gamma-carboxylated Glu (Gla) allows detection by liquid chromatography-mass spectrometry and the identification of Gla residues in the gamma-glutamyl carboxylase

K. W. Hallgren¹, D. Zhang², M. Kinter³, B. Willard², and K. L. Berkner^{1,*}

¹Departments of Molecular Cardiology and Cleveland Clinic Foundation, Lerner Research Institute, 9500 Euclid Avenue, Cleveland, Ohio 44195

²Mass Spectrometry for Protein Sequencing, Cleveland Clinic Foundation, Lerner Research Institute, 9500 Euclid Avenue, Cleveland, Ohio 44195

³Free Radical Biology and Aging Research Program, Oklahoma Medical Research Foundation, 825 NE 13th Street, Oklahoma City, Oklahoma 73104

Abstract

Gamma-carboxylated Glu (Gla) is a post-translational modification required for the activity of vitamin K-dependent (VKD) proteins that has been difficult to study by mass spectrometry due to the properties of this negatively-charged residue. Gla is generated by a single enzyme, the gamma-glutamyl carboxylase, which has broad biological impact because VKD proteins have diverse functions that include hemostasis, apoptosis, and growth control. The carboxylase also contains Glas, of unknown function, and is an integral membrane protein with poor sequence coverage. To locate these Glas, we first established methods that resulted in high coverage (92%) of uncarboxylated carboxylase. Subsequent analysis of carboxylated carboxylase identified a Gla-peptide (729-758) and a missing region (625-647) that was detected in uncarboxylated carboxylase. We therefore developed an approach to methylate Gla, which efficiently neutralized Gla and improved mass spectrometric analysis. Methylation eliminated CO₂ loss from Gla, increased the ionization of Gla-containing peptide, and appeared to facilitate trypsin digestion. Methylation of a carboxylated carboxylase tryptic digest identified Glas in the 625-647 peptide. These studies provide valuable information for testing the function of carboxylase carboxylation. The methylation approach for studying Gla by mass spectrometry is an important advance that will be broadly applicable to analyzing other VKD proteins.

INTRODUCTION

Vitamin K-dependent (VKD) proteins contain gamma-carboxylated Glu, or Gla (Fig. 1a), which is required for biological activity¹. Multiple Glas are present in VKD proteins (i.e. between 3 and 16 for individual proteins); however, the extent and distribution of Glas required for function are unknown. Currently, 16 VKD proteins have been identified, with diverse functions that include hemostasis, calcium homeostasis, growth control, signal transduction and apoptosis. Proper Gla modification of VKD proteins therefore has broad physiological impact. The analysis of Gla-containing proteins by liquid chromatography-mass spectrometry (LC-MS) has received very little attention because Gla undergoes neutral loss of CO₂ from the gamma-carboxy carbon during collisionally induced dissociation

*To whom correspondence should be addressed: berknek@ccf.org, Phone: (216)445-9760, Fax: (216)444-9263.

(CID)². This property has impeded analysis of VKD proteins, and a method to detect Gla by LC-MS would be of value. For example, new approaches could be used to define the types of undercarboxylated VKD protein forms that are generated during therapy with warfarin, an anti-coagulant that decreases the extent of carboxylation and that is used by millions of people.

The conversion of Glu to Gla is accomplished by the gamma-glutamyl carboxylase, an endoplasmic reticulum resident enzyme that uses the energy of vitamin K oxygenation to drive Glu carboxylation. A single carboxylase modifies all VKD proteins. Carboxylase mutations are associated with a severe bleeding defect¹, and carboxylase null mice die during mid-gestation from a severe hemorrhagic response³. Carboxylase mutations also cause a second disease, i.e. pseudoxanthoma elasticum, characterized by soft tissue calcification^{4,5}. The physiological role of the carboxylase underscores the importance of understanding how this enzyme modifies VKD proteins. However, very few functional residues have been identified. For example, while the enzymatic reaction uses four substrates (O₂, CO₂, vitamin K and Glu), the only carboxylase residues shown to be involved in substrate interaction are those that bind Glu (Fig. 1b)^{6,7}. Interestingly, the carboxylase contains Glas⁸, but their functional significance is unknown and their location cannot be predicted by homology comparisons with other VKD proteins. Defining functional regions is challenging because the carboxylase is an integral membrane protein without a structural determination to guide analysis. In addition, the carboxylase is unique without shared homology to other proteins.

The lack of alternative approaches makes proteomic-based methods particularly attractive for defining carboxylase function; however, analysis is confounded by the hydrophobicity of the carboxylase, which has seven predicted membrane domains (Fig. 1b)⁹. Protein hydrophobicity presents difficulties in obtaining high sequence coverage by mass spectrometry due to limited proteolysis and/or the requirement of detergent for solubility, and the carboxylase is no exception. Two studies have analyzed the carboxylase by using mass spectrometry with sequence confirmation^{10,11}, with only one study indicating sequence coverage (40%). In both studies, the authors noted that the highly hydrophobic nature of the carboxylase was an impediment in achieving better sequence coverage. To date, there has been no systematic effort to improve the coverage.

We undertook a proteomic study to identify the region of the carboxylase that is carboxylated. This study involved a combined effort to obtain extensive coverage of the carboxylase and to develop a method to identify Glas by mass spectrometry (Fig. 1c). Near-comprehensive coverage of uncarboxylated carboxylase was obtained, including the identification of all 40 Glus. We developed a methylation method to neutralize the negative charge on Gla, which facilitated Gla detection by mass spectrometry and which led to the identification of Glas in the carboxylase. The methylation approach is broadly applicable to the analysis of all VKD proteins.

EXPERIMENTAL SECTION

Expression and purification of recombinant carboxylase

To express r-carboxylase in mammalian cells, a BamH I fragment containing the carboxylase with a C-terminal FLAG epitope was isolated from r-carboxylase-FLAG/BacPAK¹² and subcloned into the 229 vector¹³ encoding methotrexate resistance. The plasmid was stably-transfected in baby hamster kidney (BHK) cells by selection in media containing 1 μ M methotrexate. Clones were screened using a Gla formation assay that measures [¹⁴C]-CO₂ incorporation into Glu, as before¹⁴, and protein concentration was determined by a BCA assay (Pierce). Carboxylase levels were determined using a BAP-

FLAG standard (Sigma) in a Western with anti-FLAG primary antibody (Sigma) and goat anti-rabbit secondary antibody (GE Healthcare). Carboxylase was purified by adsorption to anti-FLAG agarose (Sigma), followed by washing and elution with FLAG peptide, as previously described for r-carboxylase expressed in insect cells¹². Uncarboxylated- and carboxylated carboxylase forms were generated by culturing cells in the absence or presence of vitamin K (5 µg/ml Phytonadione, Merck), respectively. Amino acid analysis revealed Gla residues only in carboxylase isolated from vitamin K-containing cells, as before⁸.

Identification of carboxylase peptides by in-gel proteolysis

Carboxylase was deglycosylated before proteolysis, and endoglycosidase H (endo H) and PNGase F were first tested for suitability. Carboxylase (1 pmol in a buffer containing 25 mM Tris-HCl pH 7.4, 0.2% CHAPS, 0.3% phosphatidyl choline, and 500 mM NaCl) was denatured for 30 min at 20°C with 0.5% SDS, 40 mM DTT and then incubated at 37°C with either PNGase F (New England Biolabs, 0.2 pmol for 1–48 hr) or endo H (New England Biolabs, 0.2 pmol for 1 hr and adjusted to pH 5.5 with 50 mM citrate). Analysis by a Western, performed as above, indicated more extensive digestion with endo H. For mass spectrometry analysis, carboxylase (200 pmol) was treated with endo H (80 pmol) as above and then electrophoresed on a 10% Tris-HCl Criterion gel (Biorad; 20 pmol carboxylase per lane) and stained with Coomassie (Gel code blue, Pierce). Individual bands were excised, gel pieces were washed twice in 50% ethanol, 5% HAc and then dehydrated with acetonitrile. After reduction with DTT (5 mM) and alkylation with iodoacetamide (15 mM) for 30 min at 20°C, gel pieces were incubated in trypsin (Promega, 4 pmol), Glu C (Roche, 4 pmol) or chymotrypsin (Roche, 5 pmol). The digestions with trypsin and Glu C were overnight, while the chymotrypsin incubation was for 4, 8 or 20 hr, all at 20°C. Peptides were extracted from the gel by two rounds of incubation in 30 µl 50% acetonitrile, 5% formic acid, and the extract was lyophilized to ~10 µl, followed by the addition of 1% HAc (20 µl). The sample (10 µl) was then analyzed by LC-MS. A Finnigan LTQ linear ion trap mass spectrometer system was used, with a self-packed 9 cm × 75 µm Phenomenex Jupiter C18 column and a 2–95% acetonitrile/0.05 M HAc gradient at a flow rate of 0.3 µl/min. In one set of experiments, a Poros R2 column was also used with two different gradients (2–95% or 10–95% acetonitrile/0.05 M HAc) followed by 30 min isocratic elution with 95% acetonitrile/0.05 M HAc. The digests were analyzed in a data-dependent manner by an initial full mass scan (ion scan range 300–2000 Da), which was followed by four CID scans on the most abundant ions identified in the mass scan. Dynamic exclusion was utilized if an ion was identified three times in 15 seconds and was kept on the exclusion list for 20 seconds.

Proteins were identified using the program Mascot to search the CID spectra against the human reference sequence database, which was downloaded on September 8th, 2010 and had 37,391 entries. The fixed and variable modifications were Cys carbamidomethylation and Met oxidation, respectively, and the mass tolerance was 3.0 Da (peptide) and 2.0 Da (fragment ion). The criteria for acceptance were a Mascot peptide ion score greater than 30 and the identification of at least two peptides. Post-translational modifications were identified by searching the CID spectra using the program Sequest, which is bundled into the Bioworks Browser (ThermoFinnigan). This search was performed specifically against a database containing only the carboxylase sequence (accession number NP_000812.2). The fixed modification was Cys carbamidomethylation and variable modifications were oxidized Met, GlcNAc modification of Asn residues, Gla and methylation of Glu, Gla, Asp or the C-terminus. Mass tolerance was the same as in the Mascot search. The criteria for acceptance were Sequest peptide X_{corr} scores >1.5 ($[M+H]^+1$ peptides), >2 ($[M+2H]^+2$ peptides) or >2.5 ($[M+3H]^+3$ peptides). CID spectra for all modified peptides were also manually inspected to confirm identification. Tryptic digests were searched specifically for tryptic peptides while

all other proteolytic digests were searched using a no enzyme filter. In some experiments, selected reaction monitoring (SRM) analyses were also performed to target specific carboxylase peptides. All CID were searched using the programs Mascot and/or Sequest, as described above and as indicated in individual experiments.

Methylation of Glu- and Gla-containing peptides

Methylation was tested as a method to improve the detection of Gla by mass spectrometry. The efficiency of methylation was initially monitored using HPLC and absorbance to detect model peptides, i.e. FLEEL and a Gla (γ)-containing peptide, FL γ EL (both from Anaspec). Duplicate aliquots of 20 nmol peptide were lyophilized to dryness in siliconized tubes, and then incubated for one hr at 20°C in the presence or absence of 200 μ l 2M methanolic HCl. Lyophilization and methylation were repeated one additional time, and all samples were then resuspended in 100 μ l 50% acetonitrile, 5% formic acid. The volume was reduced to approximately 5 μ l by lyophilization, followed by the addition of 50 μ l 1% HAc and analysis by HPLC using a C18 Everest column (Pierce, 150 $\times$ 2.1 mm). Peptides eluted using a 2–90% acetonitrile/0.1% trifluoroacetic acid gradient were detected by absorbance at 210 nm.

Unmethylated- and methylated FLEEL and FL γ EL were also analyzed by LC-MS. Peptides (10 pmol) were identified by data-dependent analysis as above, except that a 2–60% acetonitrile/0.1% formic acid gradient was used. Unmodified forms of the peptides were not detected, indicating that the methylation reactions were complete, and methylation was observed at Glu, Gla and C-terminal carboxylic acids. In a separate experiment, the ionization efficiencies of the peptides were compared by analyzing a mixture of all 4 peptides (2.5 pmol each). The abundance of individual peptides was quantitated by measuring the peak area.

LC-MS analysis of methylated carboxylase tryptic peptides

The effect of methylation on the detection of Gla-containing peptides in carboxylated carboxylase was assessed. Endo H-treated carboxylase (50 pmol) was gel-electrophoresed, carboxymethylated and digested with trypsin (8 pmol) as described above. Peptides extracted from the gel were lyophilized to dryness and then subjected to two rounds of methylation using 200 μ l 2M methanolic HCl and incubation at 20°C for one hour. The samples were lyophilized and resuspended in 50 mM ammonium bicarbonate, pH 8 (100 μ l) to achieve neutral pH, and after a final lyophilization the peptides were brought up in 1% HAc (10 μ l) and analyzed by LC-MS. Duplicate samples were resuspended in 10 μ l 50 mM ammonium bicarbonate, pH 8 and redigested with trypsin (4 pmol) overnight at 20°C, followed by the addition of 1% HAc (30 μ l) and analysis by LC-MS. Protein methylation was confirmed by searching the LC-MS/MS data using the program Sequest and considering carboxylation of Glu residues and methylation of Glu, Gla, Asp and the C-terminus. Analysis of the chromatograms revealed substantially more methylated- than unmethylated peptide, indicating that the methylation reaction occurred with high efficiency.

Analysis by electron transfer dissociation (ETD) was also performed on tryptic digests, as well as a tryptic digest of protein C and a chymotryptic digest of the carboxylated carboxylase, prepared as described above. The instrument used for these experiments was an Orbitrap-Velos with ETD mass spectrometer (Thermo Scientific) equipped with a nano-flow Proxeon Easy n-LC system, which allowed parallel analysis by ETD and CID. Digests (5 μ l) were injected onto a XperTek 218TP C18 (Cobert Associates, St. Louis) 15 cm $\times$ 75 μ m column and separated by a 0–80% acetonitrile/0.1% formic acid gradient at a flow rate of 0.2 μ l/min. Full mass scans (300–2000 Da) were taken in the Orbitrap mass spectrometer, operated at 30,000 resolution (full width at half maximum at m/z 400). In data-dependent

analysis, a full high-resolution mass scan was followed by six data-dependent ETD scans or six CID scans. The ETD reactions were performed for 100, 150, and 200 ms, and ETD and CID fragment ions were analyzed in the LTQ ion trap. The data were searched against the sequence of the carboxylase using the program Sequest, bundled into Proteome Discoverer 1.3 (Thermo Scientific). The searches were carried out considering ETD activation and 2 missed cleavages, with a mass tolerance set to 10 ppm for mass scans and 2.0 Da for the ETD and CID spectra. The fixed and variable modifications, as well as the criteria for acceptance, were as described above.

RESULTS and DISCUSSION

INCREASING SEQUENCE COVERAGE OF THE CARBOXYLASE

Expression of functional r-carboxylase—r-Carboxylase was stably-expressed in BHK cells, and initial analyses were performed on uncarboxylated carboxylase isolated from cells cultured in the absence of vitamin K. The r-carboxylase showed a specific activity that was the same as that of carboxylase isolated from tissue (data not shown), indicating that the r-carboxylase was suitable for subsequent analysis. The carboxylase was affinity-purified using an antibody against a FLAG epitope appended to the C-terminus of the enzyme, and Coomassie staining indicated that the carboxylase was the most prominent band (Fig. 2a).

In-gel digestion with trypsin—The carboxylase has 8 potential sites of glycosylation⁹, and deglycosylation was therefore assessed prior to peptide mapping because the carbohydrates would interfere with peptide detection. The carboxylase was endo H-sensitive, consistent with its location in the endoplasmic reticulum, and an approximate 15 kD decrease in mass was observed (Fig. 2b). PNGase F resulted in only partial deglycosylation, and endo H deglycosylation was therefore subsequently used.

Trypsin was the first protease tested because it has several characteristics desirable for mapping by mass spectrometry. Tryptic peptides of uncarboxylated carboxylase generated as described in the Experimental section were analyzed using a Finnigan LTQ linear ion trap mass spectrometer system. Full scan mass spectra were acquired to determine the peptide mass (Table S1) and CID spectra of ions were successively scanned to determine the peptide sequence. When the CID spectra were searched against the human reference sequence database using the program Mascot, the only peptides identified were those derived from the carboxylase, and the sequence coverage was 48% (Table S1). MALDI-TOF analysis was also performed, and identified a total of 19 peptide ions that mapped to the sequence of the carboxylase (data not shown). All of these peptides were identified in the LC-MS analysis.

The most poorly covered region in the carboxylase was between ~100 and 400 amino acids (Table S1). This region is quite hydrophobic (Fig. 1b), and attempts were made to increase the coverage by using alternative columns and/or mobile phase conditions. For example, perfusion chromatography with a Poros R2 column was tested, as were varying acetonitrile gradients followed by 95% isocratic runs. The chromatographic profiles indicated that the most abundant ions observed with these changes were the later eluting peptides (data not shown); however, despite dramatic chromatographic differences, the peptides identified under these varying conditions were very similar (Table S1).

Solution digestion of the carboxylase—Proteolysis in solution has been valuable in identifying membrane proteins in complex mixtures (e.g.¹⁵), and was investigated because in-solution digestion could potentially allow better access of the carboxylase to trypsin. Tests were first performed on radiolabeled carboxylase purified from cells cultured in media containing [³⁵S]-Cys/Met, and promising methods were then applied to analyzing unlabeled

carboxylase by mass spectrometry. Carboxylase purification requires detergent that interferes with HPLC and mass spectral analysis, and we therefore assessed several different precipitation methods, which showed that chloroform/methanol precipitation¹⁶ gave the highest recovery of [³⁵S]-carboxylase (81%, data not shown). Precipitated [³⁵S]-carboxylase was subjected to resuspension and trypsin digestion using a variety of surfactants and/or denaturants (Table S2). The digests were monitored by scintillation counting for recovery of radioactivity and by SDS-PAGE for the quality of the digest. RapiGest gave the best results; however, parallel analysis of an unlabeled tryptic digest by LC-MS indicated poor coverage (20%) that could not be improved by subsequent manipulations (e.g. using RapiGest in combination with other surfactants)(data not shown).

The best coverage obtained by LC-MS was with a carboxylase sample digested with trypsin in the presence of urea and acetonitrile. Data-dependent analysis revealed that most of the peptides were from the carboxylase, and the coverage was 51% (Table S3). Analysis by MALDI did not identify any additional peptides. While a few new peptides were identified that were not observed by in-gel trypsin digestion, much of the carboxylase remained uncovered, and so an alternative approach using other proteases was pursued.

In-gel digestion with chymotrypsin and GluC—Uncarboxylated carboxylase was subjected to in-gel chymotrypsin digestion and subsequent analysis by LC-MS, and Sequest was used to search all of the CID spectra for carboxylase peptides. Chymotrypsin digestion was performed at several different timepoints (i.e. 4, 8 and 20 hr), which together resulted in the identification of 155 peptides that covered 81% of the carboxylase sequence (Table S4). Analysis at different times proved to be useful in generating a diverse mixture that achieved high coverage, as only 36 peptides (23% protein sequence) were identified in all 3 timepoints. Nonspecific cleavage also accounted for the greater complexity of the mixture. Chymotrypsin led to the identification of most of the hydrophobic region (i.e. ~100–400 amino acids) that was poorly covered by trypsin (Fig. 3). GluC was also tested but resulted in only limited coverage (i.e. 27%), and no additional sequences were observed over those detected using chymotrypsin or trypsin.

The combined chymotryptic and tryptic peptide identifications resulted in 92% sequence coverage of the carboxylase (Fig. 3). Extensive sequence coverage has only been obtained for a small number of integral membrane proteins, and so the results showing that chymotrypsin digestion was much more successful than in-solution trypsin digestion provide important information. The extensive coverage of the hydrophobic region in the carboxylase is particularly valuable, as recent studies have indicated that the catalytic residues are within these sequences (Fig. 1)^{12, 17}.

Carbohydrate identification demonstrates the value of improved carboxylase coverage—Six glycosylated residues were identified in chymotryptic and tryptic digests (Table S5), i.e. Asn residues 459, 525, 550, 570, 605 and 627. The chymotryptic digest was from carboxylase treated with endo H, which leaves a GlcNAc on glycosylated Asn residues, and several Asn-containing peptides with an additional 203 Da mass were observed (Table S5). The CID spectra of these peptides confirmed the sites of glycosylation (data not shown). Five of these glycosylated Asn residues were also identified by solution digestion with trypsin (Table S5). The results are informative in revising the view of carboxylase glycosylation. A previous study using mutagenesis concluded that Asn570 is normally not glycosylated but acts as an alternative site of modification when the naturally-glycosylated residues are eliminated¹¹. However, the identification of glycosylated Asn570 in chymotryptic peptides and inability to identify the unglycosylated form of the peptide (Table S5) indicates that this residue is normally glycosylated.

GLA METHYLATION FACILITATES LC-MS DETECTION AND CARBOXYLASE LOCALIZATION

Analysis of carboxylated carboxylase—The coverage obtained by chymotrypsin and trypsin digestion included the detection of all 40 Glu residues (Fig. 3), establishing the basis for next identifying Gla-containing peptides. Carboxylated carboxylase was generated in BHK cells cultured with vitamin K. This source of enzyme was used rather than in vitro carboxylated material because analyses in vitro indicated that the extent of carboxylase carboxylation varied with assay conditions (data not shown). Carboxylated carboxylase isolated from BHK cells showed retarded migration on SDS-PAGE, even after deglycosylation (Fig. S1a, b). This property was due to the addition of the negatively-charged Glas, as pure uncarboxylated carboxylase reacted in vitro also showed a shift in migration (Fig. S1c).

To identify Gla-containing peptide(s), carboxylated carboxylase was analyzed in parallel with the uncarboxylated form. The peptide masses of tryptic peptides (Table S6) and chymotryptic peptides (Table S7) were identified by data-dependent analyses, and CID searches were performed to confirm peak assignment and identification of the peptides. The CID spectra were searched against the sequence of the carboxylase considering the possibility of 44 Da mass modifications of Glu residues due to the addition of a carboxyl group. The CID spectra were also searched for a 44 Da loss because CO₂ is eliminated during peptide fragmentation².

Surprisingly, no Gla-containing peptides were identified even though the coverage (82% for both forms) was sufficient to identify 39 of 40 Glu residues in the uncarboxylated form (Tables S6, S7). Importantly, the analyses revealed two regions at the C-terminus that were observed only in the uncarboxylated sample (Fig. 4a). These sequences included 2 tryptic peptides (623-647 and 625-647) containing 6 or 5 Glus, respectively, and a chymotryptic peptide (729-758) containing 4 Glu residues. A chymotryptic peptide (642-655) was also detected that was present in approximately 20-fold higher amounts in the uncarboxylated sample.

SRM experiments were used to search for peptides not observed in the data dependent analysis. The peptides investigated included the tryptic peptides 623-647 and 625-647, the chymotryptic peptide 729-758, and two chymotryptic peptides (610-622 and 616-622) that were analyzed because they contain the only Glu (Glu620) not observed in the data dependent analysis. The peptides were searched considering the addition of varying amounts of Glas (e.g. 0–5 for the 625-647 peptide), and the 623-647 and 625-647 peptides were searched for both the Asn and GlcNAc-Asn forms. This analysis revealed a small amount of uncarboxylated, GlcNAc-containing 623-647 and 625-647 in the carboxylated sample. However, the levels were approximately 4 orders of magnitude lower in abundance compared with the uncarboxylated sample (Fig. 4b), and peptides with Glu+44 Da additions were not detected. Glus that are carboxylated in VKD proteins are usually fully modified¹⁸, and the low amount of Glu-containing peptide in carboxylated carboxylase suggests the same stoichiometry in the carboxylase. The results with the 729-758 peptide were similar to that of the 623/625-647 peptides: Glu- but not Gla-containing peptide was observed in the carboxylated carboxylase digest but at an abundance about 3 orders of magnitude lower than in the uncarboxylated carboxylase digest (Fig. 4b). In the case of the 610-622 and 616-622 peptides, the difference in amount of uncarboxylated and carboxylated peptide was about 10-fold (Fig. 4b); however, no peptides with a 44 Da addition were observed. The SRM data were searched for a 44 Da neutral loss, but none were detected, indicating that neutral loss cannot account for the lack of detection of Gla-containing peptides.

Gla detection was recently demonstrated using ETD¹⁹, and we therefore investigated whether this approach could identify carboxylase Glas. Analysis of a trypsin-digested vitamin K-dependent control protein (i.e. protein C) by LC-MS using an Orbitrap Velos mass spectrometer identified Gla containing peptide by ETD (Fig. S2), consistent with the earlier report¹⁹. ETD data-dependent analysis of carboxylase tryptic digests identified fewer peptides in uncarboxylated- and carboxylated carboxylase (Table S8) than observed by CID (Table S6). The main difference was the lack of detection of several small peptides by ETD, consistent with the known preference of ETD for larger peptides with higher charge states²⁰. SRM experiments revealed the 623-647 and 625-647 tryptic peptides in uncarboxylated carboxylase, both by CID analysis (not shown) and by ETD (Table S8). However, neither of these peptides were identified by CID or ETD in the analysis of carboxylated carboxylase, nor were other Gla-containing peptides detected.

Data-dependent and SRM analyses by ETD and CID were also performed on chymotrypsin-digested carboxylated carboxylase. The number of peptides identified by ETD was less than that observed with CID (not shown), most likely due to chymotrypsin generating small peptides with low charge states. Both data-dependent and SRM analysis by CID revealed a 729-758 peptide containing two Glas. The CID spectrum (Fig. 4c) identified residues Glu748 and Glu757 as the sites of modification. The ETD spectrum for 729-758 was uninformative in providing sequence information. The 623-647 sequences were not detected in this analysis, which was not surprising since chymotrypsin did not cover this region even with uncarboxylated carboxylase (Fig. 4a).

Development of a novel method to identify Gla-containing peptide—One possible explanation for the lack of detection of the 623-647 region in carboxylated carboxylase could be that the negative charge on the Gla residue (Fig. 1a) impairs proteolysis. A tryptic peptide C-terminal, but not N-terminal, to 623-647 was observed (Fig. 4a). Impaired trypsin digestion at residues 622 and 624 would result in a large (i.e. 589-647), hydrophobic peptide that might not be detected. We therefore sought a way to neutralize the charge on Gla, and tested the consequent effect on trypsin digestion. Methylation was previously shown to result in Gla detection by Edman degradation and HPLC²¹, and this modification was investigated.

Model peptides (FLEEL and a Gla (γ)-containing peptide, FL γ EL) were first tested and showed that methylation was an effective method for neutralizing the charge on Gla. Modification was efficient, as evidenced by retarded migration during HPLC (Fig. 5a–d). Methylation changed the properties of Gla-containing peptide in ways that improved detection. Methylation eliminated the neutral loss of CO₂ from the gamma-carbon position during peptide fragmentation. Thus, a 44 dalton loss was observed with the unmethylated form of FL γ EL but not FLEEL (Fig. 6a, b), which is most likely due to CO₂ fragmentation, and the abundance of the neutral loss peaks was greater than that of the sequence specific ions. In contrast, neutral loss products were not observed with methylated FL γ EL, and the CID spectra for methylated FL γ EL and FLEEL were very similar (Fig. 6c, d). A second advantage of methylation was increased efficiency of ionization of the FL γ EL peptide. Ionization of unmethylated FL γ EL was about 3-fold lower than that of unmethylated FLEEL, while the ionization efficiencies for the methylated forms were very similar (Fig. 5e).

Identification of a second Gla-containing peptide—Glu/Gla methylation was next applied to the carboxylase. Carboxylated- and uncarboxylated forms were deglycosylated, carboxymethylated, and digested in-gel with trypsin, followed by methylation. One aliquot was subjected to LC-MS using the Finnigan LTQ mass spectrometer and another was redigested with trypsin before analysis. A targeted search of the 623-647 and 625-647

peptides considering 0–3 methylated Glas was performed, revealing a 625-647 peptide with a mass consistent with 3 Glas (Table S9) and a CID spectrum (Fig. 7) that confirmed the sequence. Notably, the peptide was only detected after redigestion of the methylated digest with trypsin, suggesting that methylation also impacts trypsin digestion. Additional SRM experiments were performed to search for peptides with 4–6 Glas and to analyze carboxylase digested in solution with trypsin, however no additional Gla-containing peptides were detected (data not shown). Similar results were obtained when a carboxylated carboxylase tryptic digest methylated and redigested with trypsin was analyzed by CID using the Orbitrap Velos mass spectrometer. Carboxylated T625-647 (and T623-647) were not detected by ETD. CID analysis indicated low abundance of the peptide, and ETD may have been less sensitive in detecting the peptide. Lack of detection could also be due to the known preference of ETD for higher charge states than that of the doubly-charged 625-647 peptide.

The CID spectrum of the methylated, carboxylated 625-647 peptide was dominated by cleavage at Pro (Fig. 7), giving a major y_{13} ion with a mass consistent with two Glas between residues 635 and 647. The mass of the y_5 ion (647 Da) is consistent with one Gla at position 637, while the 143 Da mass difference between the y_4 and y_5 ions corresponds to a methylated Glu, suggesting that the other Gla is at position 645. The remaining Gla is at either residue 626 or 630, based on the difference in masses observed between the y_{13} ion and intact peptide ion.

SUMMARY

Structure-function relationships in the carboxylase are poorly understood, including how carboxylase carboxylation regulates activity, and proteomic approaches offer an attractive alternative for investigating this important enzyme. The studies described here to identify Glas in the carboxylase had two major challenges: poor sequence coverage and difficulty in Gla detection by mass spectrometry. We obtained 92% sequence coverage (Fig. 3) that led to the identification of Glas at the C-terminus (Fig. 4c). We also developed a Gla methylation approach (Figs. 5 and 6) that allowed Gla identification (Fig. 7). The Gla-peptide was only observed after the methylated tryptic digest was redigested with trypsin (Table S9), suggesting that Gla methylation improves proteolysis. The high sequence coverage will be valuable in future studies on carboxylase function. For example, conflicting results on the binding site in the carboxylase that mediates VKD protein association have been reported using low-resolution mapping methods^{22, 23}, which can now be resolved by high-resolution mass spectral mapping made possible by the extensive carboxylase coverage.

The identification of Gla residues now allows an investigation into the role of this modification at the C-terminus, a region whose function is unknown. The C-terminus is distal from the active site (Fig. 1), consistent with previous studies suggesting that carboxylase Glas are not required for catalysis⁸. One interesting question is whether the unique location of the carboxylase Gla domain in the endoplasmic reticulum lumen confers a function distinct from that of other VKD proteins, which act at the cell surface or in the extracellular matrix. Future studies that address the role of Glas in the carboxylase should provide important insight into the mechanism by which this enzyme accomplishes VKD protein carboxylation.

The methylation approach developed for the carboxylase should be broadly applicable to analyzing other VKD proteins. The ability of methylation to overcome CO₂ loss and impaired trypsin digestion is a major technological improvement that circumvents previous difficulties in studying these proteins. Impaired trypsin digestion by Gla may be a common

theme, as partial trypsin digestion was observed in another vitamin K-dependent protein (i.e. protein C) previously analyzed by ETD ¹⁹. It will be of interest to determine whether Gla methylation impacts trypsin digestion in other VKD proteins as observed with the carboxylase. The application of mass spectrometry to Gla-containing proteins is currently quite limited, and so the methylation approach described here is a significant advance in studying VKD proteins.

Acknowledgments

We thank Drs. Kurt Runge and Mark Rishavy for their helpful comments on this manuscript. This work was funded by NIH RO1 HL055666 to KLB.

ABBREVIATIONS

VKD	vitamin K-dependent
Gla	gamma-carboxylated Glu
CHAPS	3-[(cholamidopropyl)dimethylammonio]-1-propane-sulfonate
FLEEL	Phe-Leu-Glu-Glu-Leu
FLγEL	Phe-Leu-Gla-Glu-Leu
endo H	endoglycosidase H
LC-MS	liquid chromatography-mass spectrometry
CID	collisionally induced dissociation
ETD	electron transfer dissociation
SRM	selected reaction monitoring

References

1. Berkner KL. Vitamin K-dependent carboxylation. *Vitam Horm.* 2008; 78:131–156. [PubMed: 18374193]
2. Kelleher NL, Zubarev RA, Bush K, Furie B, Furie BC, McLafferty FW, Walsh CT. Localization of labile posttranslational modifications by electron capture dissociation: the case of gamma-carboxyglutamic acid. *Anal Chem.* 1999; 71(19):4250–4253. [PubMed: 10517147]
3. Zhu A, Sun H, Raymond RM Jr, Furie BC, Furie B, Bronstein M, Kaufman RJ, Westrick R, Ginsburg D. Fatal hemorrhage in mice lacking {gamma}-glutamyl carboxylase. *Blood.* 2007; 109(12):5270–5275. [PubMed: 17327402]
4. Li Q, Grange DK, Armstrong NL, Whelan AJ, Hurley MY, Rishavy MA, Hallgren KW, Berkner KL, Schurgers LJ, Jiang Q, Uitto J. Mutations in the GGCX and ABCC6 genes in a family with pseudoxanthoma elasticum-like phenotypes. *J Invest Dermatol.* 2009; 129(3):553–563. [PubMed: 18800149]
5. Vanakker OM, Martin L, Gheduzzi D, Leroy BP, Loeys BL, Guerci VI, Matthys D, Terry SF, Coucke PJ, Pasquali-Ronchetti I, De Paepe A. Pseudoxanthoma elasticum-like phenotype with cutis laxa and multiple coagulation factor deficiency represents a separate genetic entity. *J Invest Dermatol.* 2007; 127(3):581–587. [PubMed: 17110937]
6. Mutucumarana VP, Stafford DW, Stanley TB, Jin DY, Solera J, Brenner B, Azerad R, Wu SM. Expression and characterization of the naturally occurring mutation L394R in human gamma-glutamyl carboxylase. *J Biol Chem.* 2000; 275(42):32572–32577. [PubMed: 10934213]
7. Mutucumarana VP, Acher F, Straight DL, Jin DY, Stafford DW. A conserved region of human vitamin K-dependent carboxylase between residues 393 and 404 is important for its interaction with the glutamate substrate. *J Biol Chem.* 2003; 278(47):46488–46493. [PubMed: 12968027]

8. Berkner KL, Pudota BN. Vitamin K-dependent carboxylation of the carboxylase. *Proc Natl Acad Sci US A*. 1998; 95(2):466–471.
9. Wu SM, Cheung WF, Frazier D, Stafford DW. Cloning and expression of the cDNA for human gamma-glutamyl carboxylase. *Science*. 1991; 254(5038):1634–1636. [PubMed: 1749935]
10. Tie JK, Jin DY, Loiselle DR, Pope RM, Straight DL, Stafford DW. Chemical modification of cysteine residues is a misleading indicator of their status as active site residues in the vitamin K-dependent gamma-glutamyl carboxylation reaction. *J Biol Chem*. 2004; 279(52):54079–54087. [PubMed: 15492002]
11. Tie JK, Zheng MY, Pope RM, Straight DL, Stafford DW. Identification of the N-linked glycosylation sites of vitamin K-dependent carboxylase and effect of glycosylation on carboxylase function. *Biochemistry*. 2006; 45(49):14755–14763. [PubMed: 17144668]
12. Rishavy MA, Berkner KL. Insight into the coupling mechanism of the vitamin K-dependent carboxylase: mutation of histidine 160 disrupts glutamic acid carbanion formation and efficient coupling of vitamin K epoxidation to glutamic acid carboxylation. *Biochemistry*. 2008; 47(37):9836–9846. [PubMed: 18717596]
13. Busby SJ, Mulvihill E, Rao D, Kumar AA, Lioubin P, Heipel M, Sprecher C, Halfpaw L, Prunkard D, Gambee J, Foster DC. Expression of recombinant human plasminogen in mammalian cells is augmented by suppression of plasmin activity. *J Biol Chem*. 1991; 266:15286–15292. [PubMed: 1831201]
14. Berkner KL. Expression of recombinant vitamin K-dependent proteins in mammalian cells: Factors IX and VII. *Methods Enzymol*. 1993; 222:450–477. [PubMed: 8412810]
15. Chen EI, Cociorva D, Norris JL, Yates JR. 3rd, Optimization of mass spectrometry-compatible surfactants for shotgun proteomics. *J Proteome Res*. 2007; 6(7):2529–2538. [PubMed: 17530876]
16. Wessel D, Flugge UI. A method for the quantitative recovery of protein in dilute solution in the presence of detergents and lipids. *Anal Biochem*. 1984; 138(1):141–143. [PubMed: 6731838]
17. Rishavy MA, Hallgren KW, Yakubenko AV, Shtofman RL, Runge KW, Berkner KL. Bronsted analysis reveals Lys218 as the carboxylase active site base that deprotonates vitamin K hydroquinone to initiate vitamin K-dependent protein carboxylation. *Biochemistry*. 2006; 45(44):13239–13248. [PubMed: 17073445]
18. Berkner KL. The vitamin K-dependent carboxylase. *Annu Rev Nutr*. 2005; 25:127–49. [PubMed: 16011462]
19. Ramström M, Sandberg H. Characterization of gamma-carboxylated tryptic peptides by collision-induced dissociation and electron transfer dissociation mass spectrometry. *Eur J Mass Spectrom*. 2011; 17(5):497–506.
20. Good DM, Wirtala M, McAlister GC, Coon JJ. Performance characteristics of electron transfer dissociation mass spectrometry. *Mol Cell Proteomics*. 2007; 6(11):1942–1951. [PubMed: 17673454]
21. Cairns JR, Williamson MK, Price PA. Direct identification of gamma-carboxyglutamic acid in the sequencing of vitamin K-dependent proteins. *Anal Biochem*. 1991; 199(1):93–97. [PubMed: 1807167]
22. Wu SM, Mutucumarana VP, Geromanos S, Stafford DW. The propeptide binding site of the bovine γ -glutamyl carboxylase. *J Biol Chem*. 1997; 272(18):11718–11722. [PubMed: 9115224]
23. Yamada M, Kuliopulos A, Nelson NP, Roth DA, Furie B, Furie BC, Walsh CT. Localization of the factor IX propeptide binding site on recombinant vitamin K dependent carboxylase using benzoylphenylalanine photoaffinity peptide inactivators. *Biochemistry*. 1995; 34(2):481–489. [PubMed: 7819240]

Figure 1. Gla and carboxylase function

a. VKD proteins contain clusters of carboxylated Glu (Gla) that are required for activity, and have been difficult to analyze by mass spectrometry due to loss of CO₂ from Gla during fragmentation. b. The carboxylase generates Gla, and resides in the endoplasmic reticulum membrane with seven predicted transmembrane domains⁹ that are represented here by lines and numbers indicating the amino acid positions of the domains. c. The studies presented here describe tests with different proteases and digestion conditions that obtained sequence coverage (indicated in parentheses) of most of the carboxylase. They also describe the development of a method that resulted in Gla detection by mass spectrometry. Together, these advances led to the identification of the Gla-containing region.

Figure 2. Analysis of purified r-carboxylase

a. Purified carboxylase (Carb) was subjected to SDS-PAGE and staining with Coomassie. b. Aliquots were deglycosylated for 1 hr and then analyzed in a Western, as detailed in the Experimental section. PNGase F treatment for longer times (up to 48 hr) gave similar results to those shown here.

10 20 30 40 50 60 70 80 90
 MAVSAGSART SPSSDKVQKD KAEISGPRQ DSRIGKLLGF EWTDLSSWRR LVTLLNRPTD PASLAVFRFL FGFLMVLDP QERGLSSLDL
 100 110 120 130 140 150 160 170 180
 KYLDGLDVCR FPLLDALRPL PLDWMYLVYT IMFLGALGMM LGLCYRISCY LFLLPYWYVF LLDKTSWNNH SYLYGLLAFQ LTFMDANHYW
 190 200 210 220 230 240 250 260 270
 SVDGLLNAHR RNAHVPLWNY AVLRGQIFIV YFIAGVKKLD ADWVEGY SME YLSRHWLFSP FKLLSEELT SLLVHWGGL LLDLSAGFL
 280 290 300 310 320 330 340 350 360
 FFDVSRISGL FFVSYPHMCN SQLFSIGMFS YVMLASSPLF CSPWPVKLV SYCPQLQQL LPLKAAPQS VSCVYKRSRG KSGQKPLRH
 370 380 390 400 410 420 430 440 450
 QLGAFTLLY LLEQLFLPYS HFLTQGYNNW TNGLYGYSWD MMVHSRSHQH VKITYRDGRT GELGYLNPV FTQSRRWKDH ADMLKQYATC
 460 470 480 490 500 510 520 530 540
 LSRLLPKYNV TEPQIYFDIW VSINDRFQOR IFDPRVDIVQ AAWSPFQRTS WWPQLMDLS PWRAKLQEI KSSLNHTVV FIADFPGLHL
 550 560 570 580 590 600 610 620 630
 ENFVSED LGN TSIQLLQGEV TVELVAEQN QTLREGEKMQ LPAGEYHKVY TTSPSPSCYM YVYVNTTELA LEQDLAYLQE LKEKVENGSE
 640 650 660 670 680 690 700 710 720
 TGPLPPELQF LLEGEVKGGP EPTPLVQTFL RRQRLQEI RRRNTPFHER FFRLLRKLY VFRSFLMTC ISLRNLILGR PSLEQLAQEV
 730 740 750 760
 TYANLRPF EA VGELNPSNTD SSHSNPPESN PDPVHSEFAA ADYKDDDDK

Figure 3. Sequence coverage of the uncarboxylated carboxylase

Sequences identified only by chymotrypsin or by trypsin are highlighted in yellow or blue, respectively, while those identified using either enzyme are highlighted in green. The carboxylase sequence ends at amino acid 758, and the remaining residues comprise the FLAG epitope tag.

Figure 4. Comparison of carboxylated- and uncarboxylated carboxylase C-terminal peptides and identification of a Glu-containing peptide

Trypsin and chymotrypsin digestions of carboxylated (+K) and uncarboxylated (–K) carboxylases were performed as described in Tables S6 and S7, respectively. As described in the text, data-dependent analysis (a) and SRM (b) revealed two regions in the C-terminus that were missing only in the carboxylated form. Inter-run variability was corrected by normalizing the area under each peptide peak ($A(x)$) to the area of the average of 3 reference carboxylase peptides ($A(\text{refpep})$). c. Subsequent CID analysis using an Orbitrap Velos mass spectrometer revealed two Glus in the 729–758 chymotryptic peptide. The mass of the peptide is 3278 Da, and the masses of the y_6 and $y_6\text{-CO}_2$ ions at 759 and 715 Da are consistent with carboxylation at Glu757. The masses of the y_{13} , $y_{13}\text{-CO}_2$ and $y_{13}\text{-2CO}_2$ ions at 770, 748 and 726 Da are consistent with a second carboxylation site at Glu748. The spectrum indicates significant CO_2 loss from both the precursor and sequence specific ions.

a. unmethylated FLEEL **c. unmethylated FL γ EL** **b. methylated FLEEL** **d. methylated FL γ EL** **e. Peptide ionizations** **Figure 5. Methylation of Glu and Gla residues**

a–d. Methylation was tested by incubating FLEEL and a Gla (γ)-containing peptide, FL γ EL, in the presence or absence of methanolic HCl, followed by HPLC and absorbance detection at 210. e. To test whether methylation impacts ionization efficiency, a mixture containing equivalent amounts of methylated- and unmethylated FLEEL and FL γ EL peptides was analyzed by LC-MS and data-dependent analysis. The area under each peak (Area(x)) was normalized to that of unmethylated FLEEL.

Figure 6. Glu methylation eliminates neutral loss

Equivalent amounts of FLEEL peptide and the Glu (γ)-containing peptide FL γ EL were subjected to LC-MS before (a, b) or after (c, d) treatment with methanolic HCl. Almost all of the Glus and Glas were methylated (Me), as indicated by data-dependent analysis and consistent with the results obtained by HPLC (Fig. 5). The CID spectra show that cleavage at the gamma-glutamyl carbon occurs with unmethylated- but not methylated CO_2 (b, d).

Figure 7. CID spectrum of a Gla-containing carboxylase peptide

A 625-647 peptide with a mass (2879.3 Da) indicating 3 Gla residues was identified after in-gel trypsin digestion of carboxylase, followed by methylation and redigestion with trypsin (Table S9). E^{Me}, γ^{2Me} and E* denote methylated Glu, dimethylated Gla and either methylated Glu or Gla, respectively. The spectrum shown is for the [M+2H]⁺² peptide form. The spectrum for the [M+2H]⁺³ form showed a higher signal for the y₄ ion, which is shown in the boxed insert.