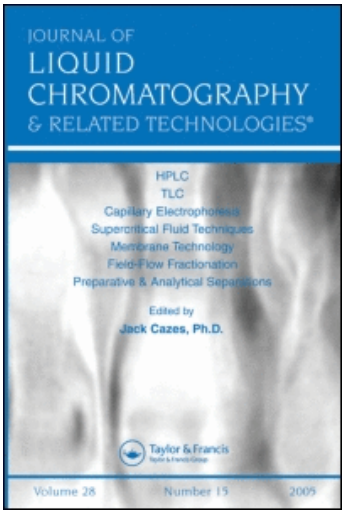


**Journal of Liquid Chromatography & Related Technologies**
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713597273

# Evaluation of Proteolytically Released Carbohydrate-Containing Peptides of Bovine Prothrombin Fragment 1 Using Electrospray Ionization Mass Spectrometry and Capillary Electrophoresis

Keisa R. Birdwell[a]; Todd L. Austell[a]; Roderick S. Black[a]; James W. Jorgenson[a]; Richard G. Hiskey[a]
[a] Department of Chemistry, University of North Carolina, Chapel Hill, CB

**To cite this Article** Birdwell, Keisa R. , Austell, Todd L. , Black, Roderick S. , Jorgenson, James W. and Hiskey, Richard G.(1997) 'Evaluation of Proteolytically Released Carbohydrate-Containing Peptides of Bovine Prothrombin Fragment 1 Using Electrospray Ionization Mass Spectrometry and Capillary Electrophoresis', Journal of Liquid Chromatography & Related Technologies, 20: 7, 987 — 1004
**To link to this Article: DOI:** 10.1080/10826079708010953
**URL:** http://dx.doi.org/10.1080/10826079708010953

J. LIQ. CHROM. & REL. TECHNOL., 20(7), 987–1004 (1997)

# EVALUATION OF PROTEOLYTICALLY RELEASED CARBOHYDRATE-CONTAINING PEPTIDES OF BOVINE PROTHROMBIN FRAGMENT 1 USING ELECTROSPRAY IONIZATION MASS SPECTROMETRY AND CAPILLARY ELECTROPHORESIS

Keisa R. Birdwell, Todd L. Austell, Roderick S. Black, James W. Jorgenson, Richard G. Hiskey*

Department of Chemistry
University of North Carolina at Chapel Hill
CB #3290, Venable Hall
Chapel Hill, North Carolina 27599



## ABSTRACT

Described here is a method used to determine the sialic acid content of the structurally distinct carbohydrates attached at N77 and N101 of bovine prothrombin fragment 1. The protein is proteolytically cleaved with α-chymotrypsin to release two glycopeptides, each containing one of the two glycosylation sites. Electrospray ionization mass spectrometry and capillary electrophoresis are used to determine the structural difference between the carbohydrates of each of these glycopeptides. N77 bears a carbohydrate containing as many as 4 sialic acid (*N*-acetylneuraminic acid) residues, whereas the carbohydrate attached at N101 may carry as many as 5 sialic acid residues.

Partially desialated forms of both of the native carbohydrate chains are observed. For both the N77 and N101-linked carbohydrates, the 1, 2, and 3-sialic acid-containing forms predominate. UV absorbance detection of the peptide-free, underivatized carbohydrate moieties is also demonstrated.

## INTRODUCTION

Glycoprotein activity has been implicated in the modulation of a wide range of functions in living organisms. Much of this functional diversity can be attributed to the carbohydrates attached to these biomolecules. Carbohydrates participate in both physicochemical and biological regulation. Their participation in physical phenomena includes the modification of protein charge, solubility, folding, conformational stability, and proteolytic accessibility.[1,2] In biological processes the carbohydrates of glycoproteins are known to regulate efficacy,[3] to serve as receptors for cellular interactions,[4] to modulate immunological responses,[5] and to function as recognition sites for targeting specific cellular species.[6] Furthermore, carbohydrates can play an important physiological role in the regulation of circulatory lifetimes of glycoproteins such as prothrombin.[7] The gradual biological removal of sialic acid (*N*-acetylneuraminic acid) residues, called desialation, occurs during the "molecular lifetime" of serum glycoproteins, giving rise to an appropriate molecular trafficking phenomenon: only the older, desialated glycoprotein molecules can interact with a hepatic receptor and be consequently removed from circulation for lysosomal degradation.[8] Sialic acids are also known to affect protein conformation and cell-cell adhesion.[7] One study demonstrates that enzymatic removal of sialic acids from blood coagulation protein factor IX results in diminished clotting efficiency.[9]

Attempts to elucidate the structures of glycoprotein carbohydrates are frequently hindered by both the chemical complexity and frequent absence of native chromophoric groups in free carbohydrates. Analysis can be further complicated by heterogeneity resulting from the existence of multiple protein glycoforms. Until the recent study of erythropoietin heterogeneity made by Rush and co-workers,[10] characterization of the carbohydrate portion of glycoproteins was typically performed after cleaving it from the protein backbone. The strategy presented here, which employs chymotryptic glycopeptides, uses very small sample sizes, requires minimal sample handling, and eliminates the need for derivatization.

Prothrombin is a blood coagulation zymogen present in blood serum at a concentration of about 100 μg/mL. Prothrombin, together with the other zymogens involved in the enzyme activation cascade of blood coagulation, contains several important post-translational modifications, among which are glycosylation, or carbohydrate attachment. The structures of three structurally distinct carbohydrate chains in bovine prothrombin were determined in the 1970s by chromatographic and electrophoretic analysis of digests produced by hydrazinolysis and exoglycosidase treatment.[11,12] Variation in maximum sialic acid content was established as one of the principal structural differences among these three carbohydrate chains of bovine prothrombin. The positions at which carbohydrates are attached to the protein backbone were later identified from amino acid sequencing; N-linkages to N77, N101 and N376 were determined.[13]



However, these studies did not indicate the specific sialic acid content of the carbohydrates attached at each of the three glycosylation sites. We sought to establish the sequence-specific sialic acid content for each of the carbohydrates of bovine prothrombin fragment 1 (bf1). Bf1 is the 156-residue amino-terminal portion of prothrombin, and hence includes the N77 and N101 glycosylation sites. We report the analysis of the variation in sialic acid content of the N77 and N101-linked carbohydrate moieties of bovine prothrombin.

The primary bioanalytical tools employed in the research described here include electrospray ionization mass spectrometry (ESI-MS) and capillary electrophoresis (CE). ESI-MS is a well-established method for the analysis of biomolecules.[14] ESI-MS provides the "soft" ionization needed for applications in which mass analysis of unfragmented biological macromolecules is desired, and is quite useful in characterizing the heterogeneity of glycoproteins.[15-17] A distribution of multiply-charged molecular ions is commonly observed in an ESI mass spectrum, which is generally deconvoluted to determine molecular weights. CE, the second analytical technique of importance to the research presented here, affords extremely efficient separations of tiny quantities of sample.[18,19]

Previous reports documenting the success of CE in studying carbohydrate structure have been reviewed by Novotny.[20] Successful CE analysis of carbohydrates often requires the use of charged complexing agents to effect separation, and generally also involves derivatization for spectroscopic detection. However, the mixtures analyzed here contain components that differ strictly in the number of negatively charged sialic acid residues they contain, and are readily separated without complexing or ion-pairing additives. Furthermore, no derivatization is required for the UV absorbance detection of these analytes.

## MATERIALS

Sialidase (neuraminidase, from *Clostridium perfringens*), α-chymotrypsin (TLCK treated, from bovine pancreas), trifluoroacetic acid (TFA), and tris(hydroxymethyl) aminomethane (tris) were obtained from Sigma Chemical Co. (St. Louis, MO). Periodic acid and *tert*-butyl alcohol were supplied by Aldrich Chemical Co. (Milwaukee, WI). Sodium chloride and cupric sulfate were purchased from Mallinckrodt (Paris, KY) and EM Science (Gibbstown, NJ), respectively. Sodium hydroxide pellets, hydrochloric acid, acetic acid, and resorcinol were obtained from Fisher Scientific (Pittsburgh, PA). N-glycosidase F (from *Flavobacterium meningosepticum*) was purchased under the trade name N-glycanase from Genzyme (Cambridge, MA). Bovine prothrombin was isolated from blood serum as described by Mann[21] and cleaved with *Echis carinatus* venom (Sigma) to generate bf1 as reported by Pollock et al.[22] Deionized water for the preparation of aqueous solutions was obtained from a Hydro (Research Triangle Park, NC) dual-cartridge purification system; ultra-high purity water, methanol, and acetonitrile for instrumental analyses were obtained from Fisher. All additional reagents used for N-terminal Edman degradation sequencing and amino acid analysis were supplied by Applied Biosystems (Foster City, CA).



## METHODS

### Glycopeptide Isolation

Chymotryptic digestion of 0.5 μmol bf1 (20 mM tris, 100 mM NaCl, 0.02% $NaN_3$, pH 7.4) was begun with a substrate-to-enzyme ratio of 50:1 (w/w). After 14 h of incubation at room temperature, the substrate-to-enzyme ratio was adjusted to 25:1, the temperature was increased to 25°C, and the proteolysis was allowed to proceed for an additional 2 h.

Chromatographic separation and preparatory collection of all proteolytically released peptides was performed on a Spectra Physics SP8800 HPLC (San Jose, CA) using a 10.0 x 250 mm, 5 μm particle size, $C_{18}$ reversed phase column (Rainin Microsorb) with a flow rate of 2.0 mL/min. Mobile phase A consisted of 0.1% (v/v) TFA in water and mobile phase B contained 0.1% TFA in acetonitrile. Chromatography was performed using a linear gradient of 0 to 70% mobile phase B over a total run time of 60 min. UV absorbance at 220 nm was monitored using a Spectra-Physics SP8450 on-line variable wavelength detector (San Jose, CA).

## Glycopeptide Identification

Chromatographed peptides bearing sialic acid-containing carbohydrates were identified using the periodate-resorcinol colorimetric determination described by Jourdian and coworkers.[23] A 0.1 mg portion of each peptide collected was used for the assay. Absorbances were determined using a Gilford model 260 spectrophotometer (Oberlin, OH).

## Glycopeptide Digestion

Sialic acids were removed from glycopeptides using 0.1-0.2 units of sialidase per 1.0 mg peptide (20 mM tris, 100 mM NaCl, 0.02% $NaN_3$, pH 6.5). Digests were typically performed at room temperature on 10-20 μg (about 2-4 nmol) of peptide. Aliquots removed during digestion were immediately analyzed by CE.

Glycopeptides were deglycosylated using 0.015 units of N-glycosidase F per 1.0 mg peptide (20 mM tris, 8 mM sodium phosphate, 105 mM NaCl, 5% glycerol, 0.02% $NaN_3$, pH 7.4). Digests were typically performed on a 10 μg sample of peptide. Following overnight incubation at 37°C, the samples were analyzed by CE.



## Amino Acid Sequencing and Amino Acid Analysis

All peptides studied were identified by N-terminal Edman degradation sequencing using an Applied Biosystems Sequenator model 475SA (Foster City, CA). After each cycle, the phenylthiohydantoin (PTH) amino acid derivatives were identified using reversed phase HPLC. Amino acid analysis was performed according to the manufacturer's instructions using an Applied Biosystems analyzer equipped with a Model 420A derivatizer, 120A Amino Acid PTH analyzer, and 920 Data Analysis Module.

## Electrospray Mass Spectrometry

Mass spectra were collected using a Vestec Model 201A quadrupole mass spectrometer (PerSeptive Biosystems/Vestec Mass Spectrometry Products, Houston, TX) equipped with an ESI source. All samples were prepared by dissolving lyophilized material in 1:1 (v/v) methanol/4% acetic acid to a concentration of approximately 1 mg/mL. Samples were continuously infused into the ESI source at 2 to 4 μL/min using a syringe pump (Model 341B, Sage

Pumps, Boston, MA). Spectrometer operating conditions were as follows: ES needle voltage, 1.7-2.0 kV; nozzle potential, 276 V; repeller potential, 20 V; ES probe temperature, 46-58 ± 5°C; block temperature, ≈ 259°C; lens temperature, ≈ 116°C. Needle-to-nozzle and nozzle-to-skimmer distances were adjusted as necessary to obtain sufficient signals. Electrospray mass spectra (1000-2000 Da) were collected in SIM mode with a Vector/2 data system (Teknivent Corp., Maryland Heights, MO). Multiple ESI-MS scans were averaged to improve signal-to-noise ratios.

### Capillary Electrophoresis

Electrophoresis instrumentation was constructed as described by Jorgenson and Lukacs[18] using a Spellman (Plainview, NY) high voltage power supply. Separations were conducted at room temperature with a 100 cm (75 cm effective length), 50 μm i.d., 360 μm o.d., unmodified fused silica, polyimide-coated capillary (Polymicro Technologies, Phoenix, AZ). A run buffer containing 50 mM NaCl and 10 mM tris (pH 8.4) was employed with a separation voltage of 22 kV. Samples were introduced at the anode end of the capillary using the electromigration method (12 kV for 5 sec). Detection was accomplished with a Linear Model UVIS 200 on-capillary spectrophotometric detector (Linear, Reno, NE). Electropherograms were acquired using an Apple Macintosh computer with a Rainin MacIntegrator (Ridgefield, NJ) data collection package. Data was filtered with a 5-point binomial filter[24] using the Igor Pro software package (WaveMetrics, Lake Oswego, OR).



## RESULTS AND DISCUSSION

### ESI-MS of Bf1

ESI-MS analysis of native bf1 reveals several populations of protein species that differ in weight by multiples of 300-330 Da (Figure 1). These differences are attributed to variation in the number of sialic acid residues associated with the carbohydrate. To ensure that no significant artifactual desialation of bf1 occurred during the HPLC separations (mobile phase pH ≈ 3) performed in these studies, colorimetric determinations of sialic acid content were performed on both chromatographed and unchromatographed protein (data not shown). No difference in sialic acid content was apparent between the two samples.

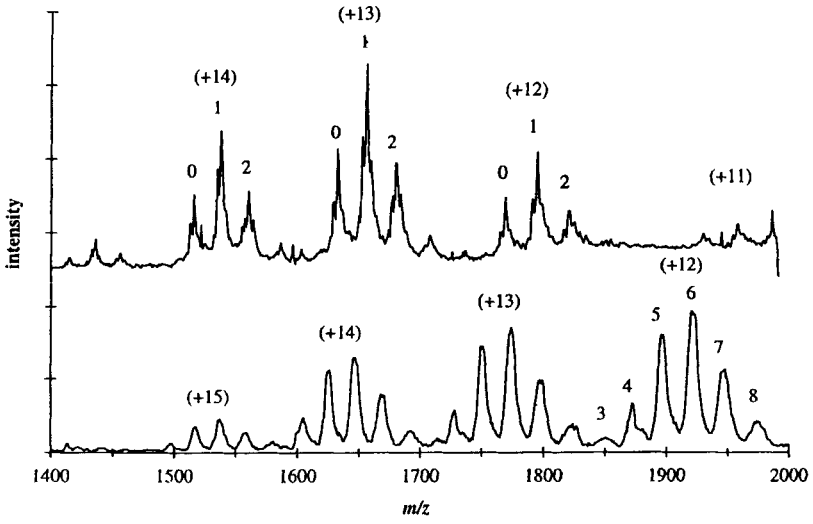


**Figure 1.** ESI mass spectra of native (lower trace) and sialidase-treated (upper trace) bf1. The number of sialic acids contained by each species is indicated; the charge $z$ on ions in each group is indicated in parenthesis. See Table 1 for the molecular weights obtained by computer deconvolution of the spectrum. Details of sample preparation and mass spectrometer operating conditions are presented in the Methods section.



## Table 1

### Glycoforms for Native and Desialated Bf1 Detected by ESI-MS

| No. of Sialic Acids | Charge Species | MW (From ESI-MS) | MW (Calculated)[a] |
|---|---|---|---|
| 0 | +14, +13, +12, +11 | 21,240.32 (±0.94) | 21,344.48 |
| 1 | +14, +13, +12, +11 | 21,547.21 (±2.68) | 21,619.74 |
| 2 | +14, +13, +12, +11 | 21,853.55 (±1.08) | 21,895.00 |
| 3 | Barely Detectable | Not Calculated | 22,170.26 |
| 4 | +15, +14, +13, +12 | 22,453.50(±1.88) | 22,445.51 |
| 5 | +15, +14, +13, +12 | 22,749.76 (±2.72) | 22,720.77 |
| 6 | +15, +14, +13, +12 | 23,042.72 (±2.60) | 22,996.03 |
| 7 | +15, +14, +13, +12 | 23,355.37 (±10.75) | 23,271.29 |
| 8 | +15, +14, +13, +12 | 23,686.62 (±4.49) | 23,546.55 |

[a]Theoretical molecular weight computed using the amino acid sequence and carbohydrate composition of the assigned bf1 glycoform.

A mass spectrum was also acquired after treatment of bf1 with sialidase, an enzyme which cleaves sialic acids from glycoproteins. Examination of the untreated and sialidase-treated bf1 spectra shown in Figure 1 reveals a total of nine major peaks (species containing 0 through 8 sialic acids), corresponding to a maximum content of 8 sialic acid residues. As seen in Figure 1, the relative intensities for the bf1 peak groups increase with *m/z*, indicating that the predominant charge species may lie above the *m/z* range limit (2000) of the mass spectrometer. However, deconvolution and subsequent molecular weight determinations can be made based on the signals from molecular ions within the *m/z* range of the instrument. The molecular weights obtained by deconvolution of the mass spectra are presented in Table 1, along with the theoretical molecular weight of each species. Differences in the theoretical and experimental weights are attributed to additional microheterogeneity of the protein (see below).



## General Strategy

Peptides are generated by proteolysis of bf1 with α-chymotrypsin and are isolated using semipreparative HPLC. Glycopeptides are distinguished from non-glycosylated peptides using a periodate-resorcinol colorimetric determination.[23] Amino acid sequencing and amino acid analysis are used to determine the amino acid sequence of each glycopeptide. Given this sequence and the common core composition of the carbohydrate chains, a theoretical molecular weight, which does not account for sialic acid residues, is calculated for each glycopeptide. The number of sialic acids contained by each glycoform is then determined by ESI-MS analysis of the glycopeptides. A complimentary series of experiments in which the isolated glycopeptides are analyzed by CE before and after treatment with several enzymes is also conducted. The ESI-MS and CE profiles of the sialic acid content are then used to establish the maximum number of sialic acid residues contained by each carbohydrate.

## Isolation and Identification of Chymotryptic Glycopeptides

Reversed phase semi-preparatory HPLC was used to isolate each glycopeptide from an α-chymotryptic digest of bf1. Peptides were collected preparatively, separated into aliquots of 0.15-0.3 mg each, and lyophilized. The two glycopeptides each elute as singlet peaks despite the differences in sialic acid content of their carbohydrate moieties. This is demonstrated by the fact that only three sialic acid-containing bands, corresponding to the two glycopeptides and a small amount of undigested protein, were detected during the entire HPLC run. Amino acid analysis and N-terminal Edman degradation

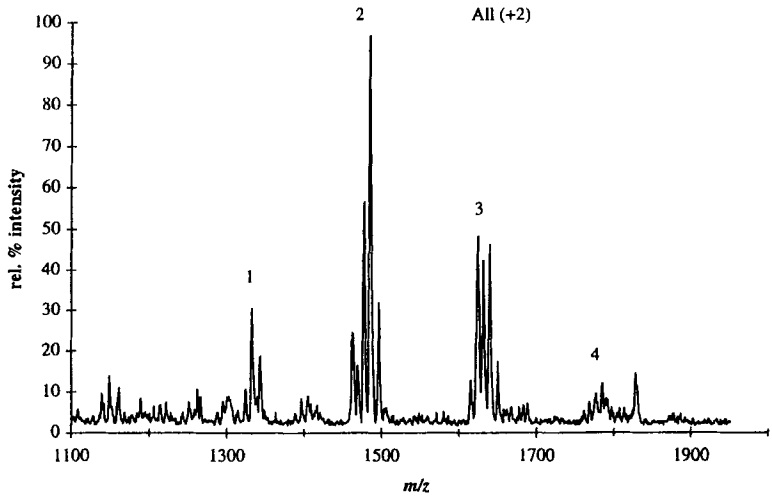


**Figure 2.** ESI-MS of the N77 glycopeptide of bfl. The number of sialic acids present on each species is indicated; all ions have a charge state $z$ of +2. Table 2 shows the estimated weights of glycoforms obtained after deconvolution of this spectrum. See Methods section for experimental conditions.



## Table 2

### Glycoforms for N77 Glycopeptide of Bf1 Detected by ESI-MS

| No. of Sialic Acids | Charge Species | MW (From ESI-MS)[a] | MW (Calculated)[b] |
|---|---|---|---|
| 1 | +2 ($m/z \approx 1340$) | 2680 | 2,654.62 |
| 2 | +2 ($m/z \approx 1480$) | 2960 | 2,945.88 |
| 3 | +2 ($m/z \approx 1630$) | 3260 | 3,226.06 |
| 4 (minor) | +2 ($m/z \approx 1780$) | 3560 | 3,516.30 |

[a]Peak locations determined by Vector/2 centroid peak finder deconvolution software.

[b]Theoretical molecular weight computed from the known amino acid sequence and carbohydrate composition of the assigned structure.

sequencing were used to determine the primary amino acid sequence of each of the glycopeptides. R75-W81 bears the N77-linked carbohydrate, and R91-F114 bears the N101-linked carbohydrate. It is worthwhile to note that α-chymotrypsin is not expected to catalyze cleavage on the C-terminal side of W90 to generate the observed N101 glycopeptide; however, amino acid analysis, sequencing, and mass spectrometric analysis all support the assignment of the structure reported here. Sequencing and CE analyses also demonstrated the presence of a contaminating tripeptide, A43-Y45, in the N77 glycopeptide fraction. The removal of this contaminant was effected by washing with three volumes of water in a Centricon-3 (3000 Da cutoff; Amicon, Beverly, MA) and was monitored using CE.

## ESI-MS Analysis of Glycopeptides



The deconvoluted mass spectrum of the N77 glycopeptide (Figure 2; Table 2) shows three major peaks corresponding to glycoforms containing 1, 2, and 3 sialic acids. This spectrum was generated using the signal from the only detectable charge state, $z = +2$. The molecular weight corresponding to the heaviest of the major species in this spectrum, 3260 Da, agrees well (within about 1%) with the calculated theoretical molecular weight of 3226 Da for the 3-sialic acid-containing glycopeptide. An extremely low signal from a species of molecular weight 3560 Da suggests that a glycoform containing 4 sialic acids is also present.

Glycoforms containing zero to four sialic acids are observed in the mass spectrum of the N101 glycopeptide (Figure 3; Table 3). The deconvoluted mass spectrum was obtained using signals from the ionized species in the three charge states ($z = +3, +4, +5$) that fall within the $m/z$ range of the spectrometer. The theoretical molecular weight of the 4-sialic acid-containing glycopeptide is 5607.57 Da, in excellent agreement (within 1%) with the experimentally observed value of 5643.51 Da for the heaviest glycoform. The consistent incremental molecular weight difference of approximately 300 Da, corresponding to one sialic acid residue, can be seen in Tables 2 and 3. Examination of individual "peaks" in both mass spectra (Figures 2 and 3) reveals fine structure that results from additional structural microheterogeneity of the protein. Such heterogeneity could be attributed, for example, to a variety of substitutions to which sialic acids are subject, including hydroxymethylation and hydroxyesterification.[7] Note that the 5-sialic acid-containing species observed by CE (see below) is not readily identified in the mass spectrum, probably owed to insufficient signal.

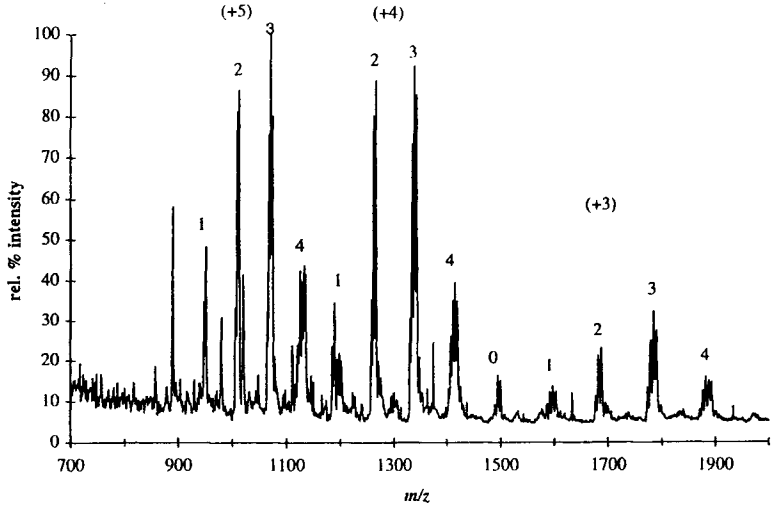


**Figure 3**. ESI-MS of the N101 glycopeptide of bfl. Peak numbers indicate the number of attached sialic acids, and the charge *z* for each group of species is indicated in parenthesis. Molecular weights obtained from the deconvoluted mass spectrum are shown in Table 3. Additional experimental conditions are given in the Methods section.



## Table 3

### Glycoforms for N101 Glycopeptide of Bf1 Detected by ESI-MS

| No. of Sialic Acids | Charge Species | MW (From ESI-MS)[a] | MW (Calculated)[b] |
|---|---|---|---|
| 0 | +3 | Not Calculated | 4,506.54 |
| 1 | +5, +4, +3 | 4,745.63 (±0.34) | 4,781.79 |
| 2 | +5, +4, +3 | 5,047.08 (±0.14) | 5,057.05 |
| 3 | +5, +4, +3 | 5,354.10 (±8.49) | 5,332.31 |
| 4 | +5, +4, +3 | 5,643.51 (±4.66) | 5,607.57 |

[a]The molecular weight of the most abundant species in each peak group is reported.

[b]Theoretical molecular weight computed from the known amino acid sequence and carbohydrate composition of the assigned structure.

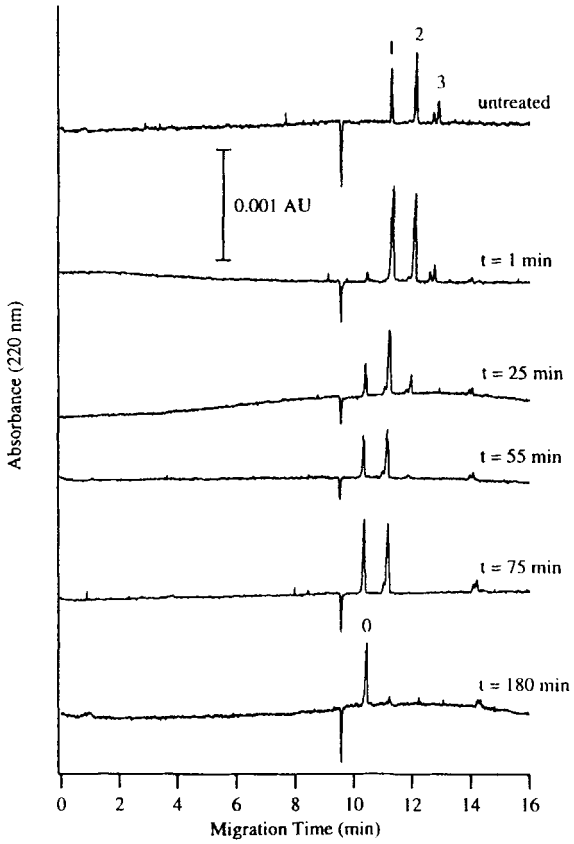


**Figure 4.** Electropherograms obtained during sialidase treatment of the N77 glycopeptide after the indicated incubation time. The number of sialic acids contained by each glycoform is indicated. Sample preparation and instrumental conditions are included in the Methods section.

## CE Analysis of Glycopeptides

Multiple peaks were observed in CE electropherograms of both glycopeptides (Figures 4 and 5), representing variation in the number of sialic acid residues contained by the carbohydrate chains. The strong dependence of electrophoretic mobility on sialic acid content is well-established and occurs as a result of the negative charge character of the sialic acid residue at moderate to high pH (CE run buffer, pH 8.4).

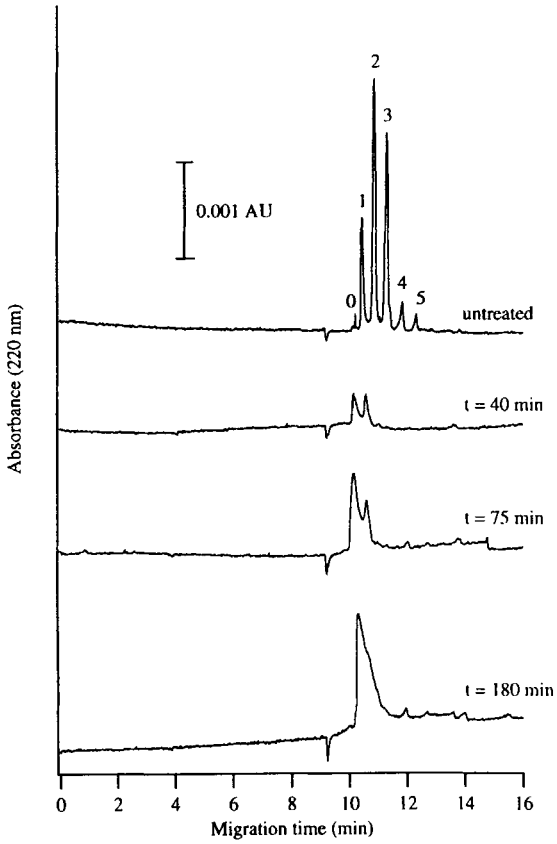


**Figure 5.** Electropherograms obtained during sialidase-catalyzed desialation of the N101 glycopeptide. CE runs were performed at the end of the incubation period indicated on each trace, as described in the Methods section. The number of sialic acids contained by each glycoform is indicated.

To provide further evidence that the observed heterogeneity arises from a distribution of sialic acid-containing species, enzymatic digestions of both glycopeptides were conducted and analyzed by CE. These experiments employed sialidase and N-glycosidase F, which catalyze sialic acid removal and deglycosylation respectively.

The electropherogram of the undigested N77 glycopeptide (Figure 4) contains three peaks, representing the species containing 1, 2, and 3 sialic acids. Adjacent peaks represent glycoforms that differ in composition by one negatively charged sialic acid residue, with the population containing three sialic acids possessing the longest migration time (12.8 min). The presence of

approximately 5% glycerol, added to increase enzyme stability, is responsible for a brief baseline drop that marks the electroosmotic flow (EOF). Specific peak assignments are substantiated by the data acquired during the course of sialidase digestion as shown in Figure 4. The gradual emergence of a fully desialated (0-sialic acid-containing) glycoform is observed; exhaustive treatment (180 min) converts the entire population to this species. Minor peaks (e. g., unlabeled peak at 12.7 min) are likely due to further heterogeneity of the protein which may include isomeric variation. CE data indicates that the N77 carbohydrate bears a maximum of 3 sialic acids. The 4-sialic acid-containing population, identified as a minor component by mass spectrometry, was not detected by CE.

In the electropherogram of the N101 glycopeptide (Figure 5), glycoforms containing zero to five sialic acids are observed. The 1, 2, and 3-sialic acid containing species are clearly the predominate forms, as indicated by their larger peak areas. The gradual disappearance of the sialated variants during digestion with sialidase supports these peak assignments. All glycoforms are ultimately desialated to generate a 0-sialic acid-containing population, which has a migration time of 11 min.



The peak shapes observed in the sialidase-treated samples are irregular; the source of this unexpected behavior, which is also observed after N-glycosidase F treatment (see below) of the N101 glycopeptide, is presently unknown. One potential cause would be a peak tailing effect originating from the increased analyte/capillary wall interaction that could be exaggerated by removal of negatively charged moieties from this large glycopeptide.

In a separate set of experiments, each glycopeptide was exhaustively incubated with N-glycosidase F, which effects cleavage of the *N*-linked carbohydrate from the peptide backbone, leaving an aspartic acid residue. The CE electropherogram of each of these digests (Figure 6) shows a peak corresponding to the deglycosylated peptide as well as multiple peaks corresponding to populations of free carbohydrate glycoforms. The free carbohydrate peaks are barely detected, but exhibit an absorbance at 220 nm, apparently because of several UV-active amide bonds.

The data presented here show that the carbohydrate of bovine prothrombin linked to N77 bears a maximum of 4 sialic acids, and the carbohydrate moiety attached at N101 may contain as many as 5 sialic acids. The most abundant forms of both the N77 and N101 carbohydrates are those containing 1, 2, or 3 sialic acid residues.

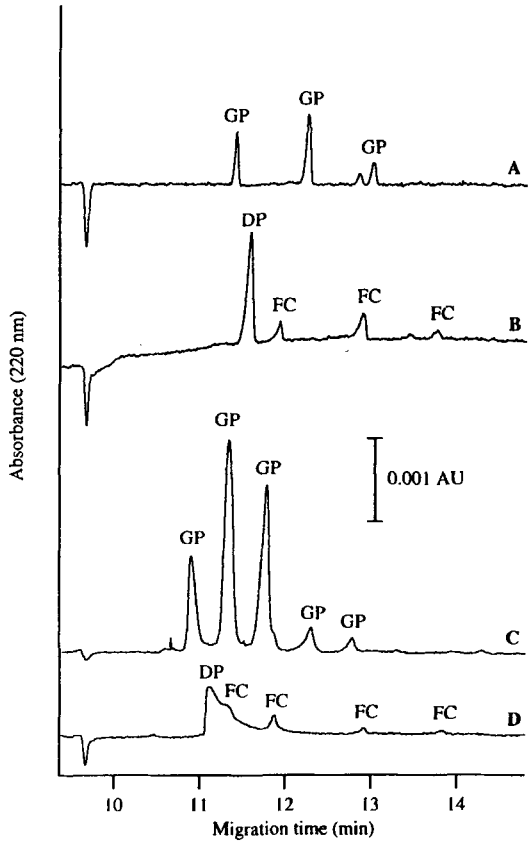


**Figure 6.** Electropherograms obtained before and after N-glycosidase F-catalyzed removal of the carbohydrate chains of both glycopeptides. A and B show N77 before and after N-glycosidase F treatment respectively; C and D show N101 before and after treatment respectively. Peaks are labeled accordingly: GP, glycopeptide; DP, deglycosylated peptide; FC, free carbohydrate. Additional experimental conditions are given in the Methods section.

Evaluation of structural variation in the carbohydrates of serum proteins such as prothrombin could greatly speed diagnosis and ultimately be used to improve treatment strategies for coagulation disorders thought to be related to glycosylation.[25-28] Many of the coagulation proteins are, in fact, similar to prothrombin in structure[29,30] and could be effectively studied using methods developed for the characterization of prothrombin. Analytical schemes employing sample-conservative techniques such as ESI-MS and CE are especially attractive for projects aimed at the characterization of other coagulation proteins, which are present in blood serum at concentrations as low

as 1/200th of that of prothrombin. The approach set fourth here should facilitate identification of the carbohydrate moieties attached at specific glycosylation sites in proteins.

## ACKNOWLEDGMENTS

KRB and RSB wish to acknowledge Ms. Paula Berkowitz and Mr. Russ Henry for technical assistance. This work was supported by grants HL-20161 (RGH), HL-06350 (RGH), and GM-39515 (JWJ) from the National Institutes of Health and grant CHE-9215320 (JWJ) from the National Science Foundation.

## REFERENCES

1. J. C. Paulson, Trends Biochem. Sci., **14,** 272-276 (1989).

2. H. Lis, N. Sharon, Eur. J. Biochem., **218,** 1-27 (1993).

3. P. M. Rudd, H. C. Joao, E. Coghill, P. Fiten, M. R. Saunders, G. Opdenakker, R. A. Dwek, Biochemistry, **33,** 17-22 (1994).

4. Y. Bourne, H. van Tilbeurgh, C. Cambillau, Curr. Opin. Struct. Biol., **3,** 681-686 (1993).

5. T. Feizi, R. A. Childs, Biochem. J., **245,** 1-11 (1987).

6. A. Stockell Hartree, A. G. C. Renwick, Biochem. J., **287,** 665-679 (1992).

7. R. Schauer, Trends Biochem. Sci., **10,** 357-360 (1985).

8. G. Ashwell, J. Harford, Ann. Rev. Biochem., **51,** 531-554 (1982).

9. S. I. Chavin, S. M. Weidner, J. Biol. Chem., **259,** 3387-3390 (1984).

10. R. S. Rush, P. L. Derby, T. W. Strickland, M. F. Rohde, Anal. Chem., **65,** 1834-1842 (1993).

11. G. L. Nelsestuen, J. W. Suttie, J. Biol. Chem., **247,** 6096-6102 (1972).

12. T. Mizuochi, K. Yamashita, K. Fujikawa, W. Kisiel, A. Kobata, J. Biol. Chem., **254,** 6419-6425 (1979).

13. S. Magnusson, T. E. Petersen, L. Sottrup-Jensen, H. Claeys, "Complete Primary Structure of Prothrombin: Isolation, Structure and Reactivity of Ten Carboxylated Glutamic Acid Residues and Regulation of Prothrombin Activation by Thrombin," in **Proteases and Biological Control,** E. Reich, D. B. Rifkin, E. Shaw, eds., Cold Sprg. Harbor Lab., 1975, pp. 123-149.

14. R. D. Smith, J. A. Loo, C. G. Edmonds, C. J. Barinaga, H. R. Udseth, Anal. Chem., **62,** 882-899 (1990).

15. A. W. Guzetta, L. J. Basa, W. S. Hancock, B. A. Keyt, W. F. Bennett, Anal. Chem., **65,** 2953-2962 (1993).

16. S. A. Carr, M. J. Huddleson, M. F. Bean, Protein Sci., **2,** 183-196 (1993).

17. R. S. Rush, P. L. Derby, D. M. Smith, C. Merry, G. Rogers, M. F. Rohde, V. Katta, Anal. Chem., **67,** 1442-1452 (1995).

18. J. W. Jorgenson, K. D. Lukacs, Science, **222,** 266-272 (1983).

19. J. W. Jorgenson, Methods (San Diego), **4,** 179-190 (1992).

20. M. V. Novotny, J. Sudor, Electrophoresis, **14,** 373-389 (1993).

21. K. G. Mann, Methods in Enzymol., **45,** 123-156 (1977).

22. J. S. Pollock, A. J. Shepard, D. J. Weber, D. L. Olson, D. G. Klapper, L. G. Pedersen, R. G. Hiskey, J. Biol. Chem., **263,** 14216-14233 (1988).

23. G. W. Jourdian, L. Dean, J. Biol. Chem., **246,** 430-436 (1970).

24. P. Marchand, L. Marmet, Rev. Sci. Instrum., **54,** 1034-1041 (1983).

25. H. R. Gralnick, H. Givelber, E. Abrams, N. Eng. J. Med., **299,** 221-226 (1978).

26. J. M. Sodetz, S. V. Pizzo, P. A. McKee, J. Biol. Chem., **252,** 5538-5546 (1977).

27. H. Maekawa, K. Yamazumi, S.-I. Muramatsu, M. Kaneko, H. Hirata, N. Takahashi, N. B. de Bosch, Z. Carvajal, A. Ojeda, C. L. Arocha-Pinango, M. Matsuda, J. Biol. Chem., **266,** 11575-11581 (1990).

28. B. Polack, J.-M. Freyssinet, Electrophoresis, **7**, 413-416 (1986).

29. D. A. Walz, D. Hewett-Emmett, M.-C. Guillin, "Amino Acid Sequences and Molecular Homology of the Vitamin K-Dependent Clotting Factors," in **Prothrombin and Other Vitamin K Proteins,** W. H. Seegers, D. A. Walz, eds., CRC Press, Boca Raton, Florida, 1986, pp. 125-155.

30. A. Kobata, T. Mizuochi, "Current Status of Carbohydrate Constituents and Prospects," in **Prothrombin and Other Vitamin K Proteins,** W. H. Seegers, D. A. Walz, eds., CRC Press, Boca Raton, Florida, 1986, pp. 81- 94.

Received July 24, 1996
Accepted August 26, 1996
Manuscript 4224