Vol. 171, No. 1, 1990 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
August 31, 1990 Pages 401-406

# CARBOHYDRATE COMPOSITION AND PRESENCE OF A FUCOSE-PROTEIN LINKAGE IN RECOMBINANT HUMAN PRO-UROKINASE

Elzbieta J. Kentzer, Alexander Buko, Gopi Menon and Virender K. Sarin*

PPD Analytical Research Department, Abbott Laboratories,
North Chicago, Illinois 60064

Received July 23, 1990

A new post-translational modification site in the growth factor domain of urinary type plasminogen activator has been identified. A glycopeptide containing the monosaccharide, fucose, covalently linked directly to the peptide backbone has been isolated from the tryptic digest of pro-urokinase expressed in a mouse hybridoma cell line Sp 2/0 Ag 14. The glycopeptide was isolated by semi-preparative reversed phase high performance liquid chromatography. The identity of a fucose containing peptide was confirmed by carbohydrate analysis, amino acid analysis and plasma desorption mass spectrometry (PDMS). A combination of these methodologies showed an equimolar ratio of peptide and fucose in the glycopeptide. This modification is not detected without mass spectrometry because the fucose residue is hydrolyzed under standard acidic conditions of amino acid composition and N-terminal sequence analysis. The site of attachment of fucose to the peptide has been localized towards the N-terminus (within first 23 amino acids) of the protein. Also, the carbohydrate composition of recombinant pro-urokinase is reported. 

Urokinase is a serine protease composed of two polypeptide chains held together by a disulfide bond. It is a potent human plasminogen activator excreted initially in its precursor form (pro-urokinase) in urine by kidney, and by certain tumor cells in culture medium. Pro-urokinase (Pro-UK) is a single chain glycoprotein with a molecular weight of 46,000 - 54,000. It is proteolytically converted by plasmin to two-chain high molecular weight urokinase. This conversion occurs by cleavage of the Lys-Ile peptide bond[1-3]. The primary structure of urokinase and Pro-UK has been determined by using Edman degradation on peptides generated by conventional analytical protein techniques[4,5]. The amino acid sequence of Pro-UK has also been deduced from the nucleotide sequence of the full length human urokinase cDNA[6]. However, very little is known about the carbohydrate content or the structures of glycan moiety/ies attached to the protein. The most common linkages between carbohydrates and proteins in glycopeptides are the following: N-acetylglucosamine-asparagine, N-acetylgalactosamine-serine, N-acetylgalactosamine-threonine, etc.[7]. Carbohydrate chains of glycoprotein can range from monosaccharides to

*To whom reprint requests should be addressed.

0006-291X/90 $1.50

very complex structures containing galactose, mannose, fucose, sialic acid, etc. We describe here an unusual type of carbohydrate-protein linkage found in the Pro-UK molecule expressed in a mouse Sp 2/0 cell line. Examples of fucose present as a branched carbohydrate structure and linked to N-acetylglucosamine have been reported. However, to the best of our knowledge, no precedent for a monosaccharide fucose attached to an amino acid in a distinct protein of biological significance is known[8], but attachment of other monosaccharides to protein structure have been reported. For example, N-acetyl-D-galactosamine linked to threonine at position 3 of Human T-cell growth factor is known[9]. The studies described in this paper report total carbohydrate composition of the recombinant human Pro-UK. We also report an unusual monosaccharide-protein linkage, e.g. fucose linked to an amino acid on the protein backbone.

## METHODS

Recombinant Pro-UK used was produced in a genetically engineered cell line derived from mouse hybridoma cell line Sp 2/0 Ag 14. It was purified to homogeneity as assessed by SDS-PAGE, reversed-phase and size exclusion high performance liquid chromatography. A solution of the purified glycoprotein used for this study had a concentration of 6.87 mg/mL as determined by its absorption at 280 nm. Pro-UK was reduced and alkylated with iodoacetamide[10], then dialyzed in the cold against 0.1 M ammonium bicarbonate buffer. At this stage, the protein precipitated out as a suspension and was used as such for digestion.

### Tryptic Digestion

For each mg of reduced and alkylated Pro-UK suspension, 10 $\mu$L of TPCK treated trypsin (Sigma) solution (1 mg/mL in 10 mM HCl) was added. It was incubated at 25°C for 27 hours, with the addition of fresh 10 $\mu$L of trypsin after 8 and 24 hours. The digestion was stopped by freezing at -20°C. The resultant peptides were separated by reversed-phase chromatography on a Vydac $C_{18}$ column using 0.1% aq.trifluoroacetic acid (TFA)/acetonitrile ($CH_3CN$) gradient. A tryptic map typical of Pro-UK at 215 nm is shown in Figure 1. Ninety-five fractions were collected manually in separate 1.5 mL Eppendorf tubes.

### Carbohydrate Analysis

Carbohydrate compositions of the native Pro-UK and desired tryptic digest peptides were determined by carrying out hydrolysis with 4 M TFA at 125°C for 45 minutes. The hydrolysate was evaporated to dryness and then dissolved in water before analyzing on a Dionex BIOLC using an amperometric detector[11].

### Amino Acid Composition and N-terminal Sequence Analysis

Vapor-phase hydrolysis of Pro-UK or desired tryptic peptides was performed using 12N HCl/TFA (2:1, v/v), 0.3% phenol at 160°C for 60 minutes in vacuo. Acid hydrolysates were then analyzed on a Beckman 6300 amino acid analyzer. N-terminal sequence analysis was carried out using standard Edman degradation methodology on an Applied Biosystems 477A gas-phase sequencer system, according to the manufacturer's protocol.

### Plasma Desorption Mass Spectrometry

Mass spectra were acquired using a BIO-ION 20 mass spectrometer (Uppsala, Sweden)[12]. The theoretical mass of the centroid for the protonated molecular ion $(M+H)^+$, for the tryptic fragments were obtained from a Macintosh computer program MacProMass, version 1.0 obtained from Beckman Research Institute of City of Hope, Duarte, CA (Terry D. Lee and Sunil Vemuri).

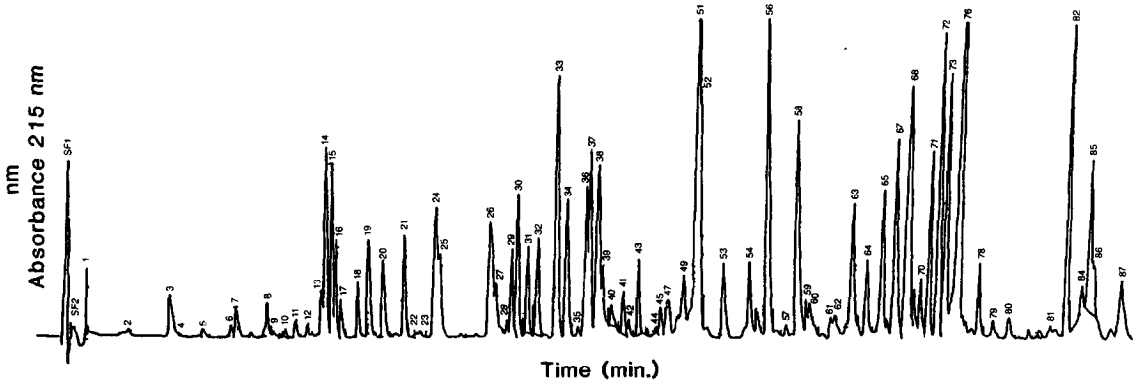


Figure 1. TFA-based tryptic map of Pro-UK. This separation was performed on a 5-um Vydac $C_{18}$ column (3.9 mm x 25 cm). Mobile Phase A consisted of aqueous 0.1% TFA and Mobile Phase B was 60% acetonitrile with 0.1% TFA in water. Labelled peaks represent fractions collected and analyzed.

## RESULTS AND DISCUSSION

Pro-UK is a glycoprotein of 411 amino acids that include 24 cysteine, 22 arginine, and 27 lysine residues. It contains a potential N-glycosylation site in the B chain at residues 301-303. However, the carbohydrate composition of this glycoprotein is not known. Results of carbohydrate analysis are shown in Figure 2 and Table 1. The composition in Table 1 is represented both in weight

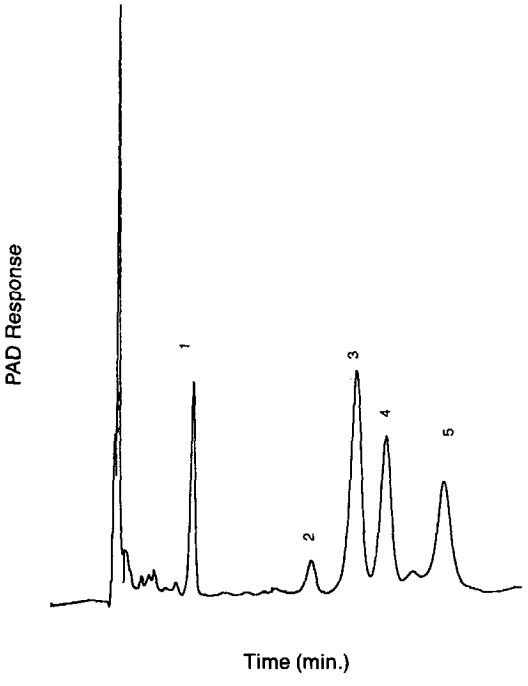


Figure 2. Chromatographic profile of the carbohydrate analysis of recombinant Pro-UK. Conditions for hydrolysis were as described in methods. Detection was by PAD, 1000 nA full scale. The identity of monosaccharides is as follows: 1. Fucose, 2. Galactosamine, 3. Glucosamine, 4. Galactose, 5. Mannose.

Table 1. Carbohydrate Composition of Pro-UK[1]

| Carbohydrate | Wt. % | Molar Ratio[2] |
|---|---|---|
| Fucose | 0.69 | 2.0 |
| Glacatosamine | 0.20 | 0.4 |
| Glucosamine | 2.17 | 4.8 |
| Galactose | 1.15 | 3.1 |
| Mannose | 1.03 | 2.8 |

[1]Composition was calculated against USP galactose standard.

[2]Molar ratio was calculated using M.W. of 48,000 for Pro-UK.

percent and molar ratio. It must be emphasized that the digestion method used here for the carbohydrate analysis is a compromise to obtain the best over-all recovery of both amino and neutral sugars in a single hydrolytic step. Under this condition, one may obtain reduced recovery of N-acetylglucosamine. It is also known that a lower recovery of mannose may be obtained if it is present in the carbohydrate structure as a disubstituted residue[11].

All of the isolated tryptic digest peptides have been analyzed using a combination of amino acid composition, N-terminal sequence analysis and PDMS. The tryptic peptides were found to be consistent with the primary sequence derived from the cDNA sequence of Pro-UK. The peptide found in fraction 38 (Figure 1) of the digest had an amino acid composition and N-terminal sequence consistent with the N-terminus tryptic peptide. However, the molecular ion observed (m/z 2737), was higher than that predicted by the cDNA sequence (m/z 2591). This mass difference of 146 amu suggested that a monosaccharide, fucose or another pentose, is attached to the N-terminus peptide. In order to confirm the identity of the prosthetic group, an aliquot of fraction 38 was analyzed for its carbohydrate composition as described in the method section. Results of the carbohydrate analysis are described in Figure 3. It showed the presence of only fucose in equimolar proportion to the peptide. The identity of fucose was further confirmed by spiking the hydrolysate with an authentic sample of fucose (Figure 3C). Attempts were made to localize the amino acid to which fucose is attached by carrying out Edman degradation all the way to the C-terminus residue of the tryptic peptide found in fraction 38 (Table 2). However, no unusual PTH-amino acid was observed in any cycle. Since fucose would be bound to an amino

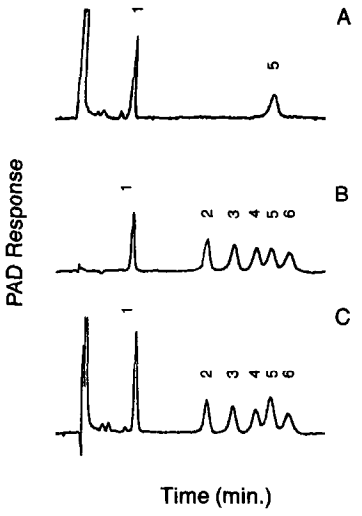


**Figure 3**. Chromatographic profile of the carbohydrate analysis of the tryptic fraction 38. A - Hydrolysate of fraction 38; B - mixture of standard carbohydrates; C - mixture of standard carbohydrates spiked with the fraction 38 hydrolysate. Conditions for hydrolysis were as described in methods. Detection was by PAD 30 nA full scale. The identity of monosaccharides is as follows: 1. Fucose, 2. Galactosamine, 3. Glucosamine, 4. Galactose, 5. Glucose, 6. Mannose.

acid via O-linkage, it would be expected that this bond would be labile to TFA at elevated temperature during the cleavage or conversion steps of standard Edman degradation cycle.

Table 2. Automated Sequence Analysis of Fraction 38 (~ 600 pmol)

| cycle | residue | pmol |
|---|---|---|
| 1 | Ser | 176 |
| 2 | Asn | 119 |
| 3 | Glu | 121 |
| 4 | Leu | 244 |
| 5 | His | 57 |
| 6 | Gln | 169 |
| 7 | Val | 181 |
| 8 | Pro | 220 |
| 9 | Ser | 164 |
| 10 | Asn | 159 |
| 11 | --- | --- |
| 12 | Asp | 78 |
| 13 | --- | --- |
| 14 | Leu | 98 |
| 15 | Asn | 37 |
| 16 | Gly | 55 |
| 17 | Gly | 62 |
| 18 | Thr | 50 |
| 19 | --- | --- |
| 20 | Val | 21 |
| 21 | Ser | 32 |
| 22 | Asn | 28 |
| 23 | Lys | 8 |

The exact location of the modified amino acid is being actively pursued. Preliminary studies from our laboratory suggest that fucose is linked to threonine at position 18 of the sequence. The attachment of a single fucose to a peptide backbone is an unusual finding and would constitute a new monosaccharide-amino acid linkage in polypeptides. Example of fucose linked to a single amino acid, threonine, has been reported[13]. However, in this case fucose is present as a disaccharide and is attached to a single amino acid, not a polypeptide. The biological significance of this linkage in a glycoprotein is not known. Studies are ongoing to determine if this unique linkage is a manifestation of the mouse hybridoma cell line Sp 2/0 used here or is a more general phenomenon. In any case, a unique carbohydrate-amino acid linkage in a glycoprotein is reported here, the significance of which has yet to be established.

**Acknowledgment**: The authors would like to thank Sally Dorwin for providing us with amino acid composition and sequencing.

## REFERENCES

1. Husain, S.S., Gurewich, V., and Lipinski, B. (1983) Arch. Biochm. and Biophy. 220, 31-38.
2. Dano, K., Andreasen, P.A., Gronhadh-Hansen, J., Kristensen, P., Nielsen, L.S., and Striver, L. (1985) In Advance in Cancer Research, Vol. 44, 139-266.
3. Gurewich, V. (1988) Seminars in Thrombosis and Hemostasis 14, 110-115.
4. Steffens, G.J., Gunzler, W.A., Otting, F., Frankus, E., and Flohe, L. (1982) Hoppe-Seyler's Z. Physiol. Chem. 363, 1043-1058.
5. Gunzler, W.A., Steffens, G.J., Otting, F., Kim, S.A., Frankus, E., and Flohe, L. (1982) Hoppe-Seyler's Z. Physiol. Chem. 363, 1155-1165.
6. Holmes, W.E., Pennica, D., Blaber, M., Rey, M.W., Gunzler, W.A., Steffens, G.J., and Heyneker, H.L. (1985) Biotechnology 3, 923-929.
7. Wold, F. (1981) Annual Review of Biochemistry 50, 783-814.
8. Morton, P.A., and Steiner, S.M. (1985) Biochem. J. 225, 59-65.
9. Robb, R.J., Kutny, R.M., Panico, M., Morris, H., DeGrado, W.F., and Chowdry, V. (1983) Biochem. Biophys. Res. Comm. 116 (3), 1049.
10. Methods in Molecular Biology, Vol. 1, Proteins (Eds. Walker, J.M.), Humana Press, N. Jersey (1984) 33-36.
11. Hardy, M.R., Townsend, R.R., and Lee, R.C. (1988) Anal. Biochemistry 170, 1-9.
12. Cotter, R.J. (1988) Anal. Chemistry 60, 781A.
13. Hallgren, P., Lundbald, A., and Svensson, S. (1975) J. of Biol. Chem. 250, 14, 5312-5314.