PROTEIN C-GLYCOSYLATION

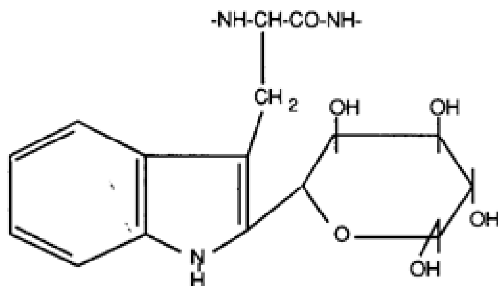
Jan Hofsteenge, Andreas Löffler
Friedrich Miescher-Institut, CH-4002 Basel, Switzerland

Dieter R. Müller, W.J. Richter
Ciba-Geigy Ltd., CH-4002, Basel, Switzerland

Tonny de Beer, Johannes F.G. Vliegenthart
Department of Bio-organic Chemistry, Utrecht University, NL-3584 CH Utrecht, the Netherlands

I. INTRODUCTION

Glycosylation of amino acid side chains of proteins is a common and widespread modification. Two types of carbohydrate-protein linkages have been known for a long time, i.e. the N- and O-glycosidic bonds to Asn, and hydroxy amino acids, respectively (1). Recently, we have reported a third type of linkage, the C-glycosidic attachment of an α-mannopyranosyl moiety to the indolic C-2 atom of Trp-7 in human RNase U_s (2,3):



Both the involvement of the tryptophan side chain and the mode of attachment are unusual. Since this appears to be the first report on such a carbohydrate-protein linkage, we will summarize the evidence for this modification, and describe a number of its properties that may be useful for its detection in other proteins.

Copyright © 1996 by Academic Press, Inc.
All rights of reproduction in any form reserved.

II. MATERIALS AND METHODS

A. Protein chemistry

RNase U_s was isolated from human urine or erythrocytes as described (2). Proteolytic digestion of the enzyme with thermolysin required denaturation by reduction and carboxymethylation (2). Purification of peptides was performed by C_{18} (Vydac, Hispania,CA) reversed-phase HPLC using the standard 0.1% TFA/acetonitrile system. Peptides were demonstrated to be pure by amino acid analysis and Edman degradation using an Applied Biosystems model 470A or 477A sequencer. Protein and peptides were sequenced in the absorptive mode with polybrene, using the standard 03CPTH cycle. Phenylthiohydantoin-amino acids (PTH-amino acids) were separated on-line with a model 120A chromatograph equipped with a Brownlee C_{18} column (2.1x220 mm; Applied Biosystems) operated at 55°C. Buffer A consisted of 7 ml premix buffer per liter of 3% THF in H_2O (Perkin Elmer), 130 µl formic acid, and eluent B was CH_3CN, containing 12% (v/v) isopropanol. The flow rate during separation was 210 µl/min. The gradient used is specified in the "Results" section. PTH-Trp^{Man} was prepared as described previously (2), and its concentration was determined from A_{269} (ε = 19'700 $M^{-1}cm^{-1}$).

Proteolytic digestions of the tetrapeptides with aminopeptidase M (10^{-2} units/nmol, Boehringer, Mannheim, FRG) and carboxypeptidase Y (Boehringer, enzyme/substrate ratio : 1/80, for 4 or 18 h at 37°C) were carried out in 50 mM NH_4HCO_3 and 100 mM Na-citrate, pH 5.0, respectively.

Modification of tryptophan residues with o-nitrophenylsulfenylchloride (NPS) was done in glacial acetic acid for 30 min in the dark. The concentrations of peptide and NPS were 0.2 and 3.3 mM, respectively.

B. Spectroscopy

Absorption spectra were recorded on a Hewlett Packard 8452 A spectrophotometer, and fluorescence emission spectra on a Perkin Elmer LS-3 fluorescence spectrophotometer. Peptide concentrations used to calculate molar extinction coefficients were determined by amino acid analysis. Details of the NMR experiments have been described elsewhere (2,3).

C. Mass spectrometry

ESI-MS and ESI-MS/MS measurements were carried out in the positive ion mode on a PE Sciex API III triple quadrupole mass spectrometer.

III. RESULTS AND DISCUSSION

A. Summary of the approaches to identify Trp^{Man} (2,3)

The chemically determined primary structure of RNase U_s from human urine (4), and the one deduced from the cDNA coding for eosinophil derived neurotoxin (EDN) were identical (5), except for residue 7. The predicted tryptophan residue could not be identified by Edman degradation. The latter was also true for the enzyme isolated from a variety of human tissues or cells (see summary in (2)), suggesting a modification of this residue. A number of different peptides containing Trp-7 were found by ESI-MS to be 162 Da heavier than expected, and to have properties in ESI-MS/MS experiments typical for aromatic C-glycosides (see below).
Since C-glycosidic bonds are resistant to acid hydrolysis, and do not yield a "sugar" ion in MS experiments, the identification of the monosaccharide had to be achieved by NMR.

B. Edman degradation of peptides containing Trp^{Man}

Automated Edman degradation in the absorptive mode of RNase U_s and its N-terminal peptides yielded one major, and two minor PTH-derivatives at position 7 (2). Although it was initially reported that the major derivative eluted shortly after PTH-Tyr, further experiments revealed that most of the time it co-eluted. This could result from changes in the buffer composition, or variations in the batches of column material. To separate PTH-Trp^{Man} from PTH-Tyr, the conditions shown in Fig. 1 were used.

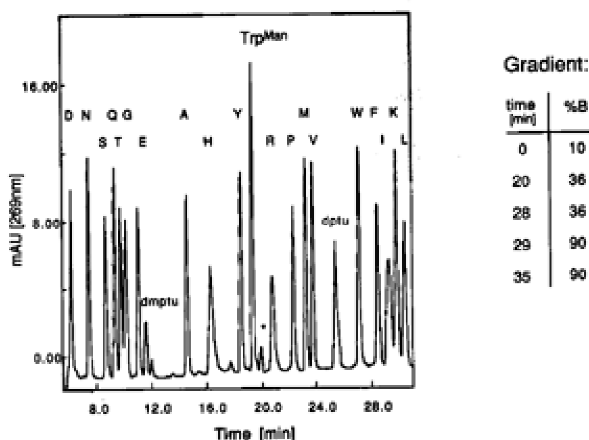


Figure 1 Improved separation of PTH-Trp^{Man} and PTH-Tyr
The major PTH-derivative of Trp^{Man} can be separated from PTH-Tyr by including 12% (v/v) isopropanol in eluent B. The minor PTH-Trp^{Man} derivative, most likely a diastereomer (2), has been indicated with an asterisk.

Another variation during Edman degradation concerns the yield of PTH-Trp^{Man}. During sequencing of F-T-W^{Man}-A the ratio PTH-Trp^{Man}/PTH-Thr varied between 0.2 and 0.9. This variability did not result from the HPLC separation, but rather from differences in the extraction conditions of the anilinothiazolinone. Probably, a more reliable identification of Trp^{Man} could be achieved by solid phase sequencing, and extraction with more polar solvents (6).

C. *Mass spectrometry of peptides containing Trp^{Man}*

Glycosylation of tryptophan can be most sensitively detected by MS. The most salient feature of Trp^{Man}-containing peptides was their increased mass by 162 Da. In addition, multiple losses of 18 Da from various ions (" ↳ " in Fig. 2) were consistently observed in MS/MS experiments of small peptides containing Trp^{Man}. ESI-MS and -MS/MS experiments on (perdeuterio)-acetylated peptides turned out to be very informative: they demonstrated the incorporation of four acyl groups into the mannopyranosyl moiety, and the conspicuous absence of a "sugar ion" (m/z 343), which would be expected for N- or O-glycosidically linked sugars. Finally, the loss of 120 Da (formally four times CH_2=O) observed in the unmodified peptides (Fig. 2) can be taken as evidence for a C-glycosidically linked carbohydrate, as such a loss was also observed with low molecular weight flavone C-glycopyranosides (7).

It is of interest to note that Gäde et al. (8) have identified a hexosylated tryptophan in a neuropeptide from the insect *Carausius morosus*, based on an increased mass of 162 Da. These investigators proposed an N-glycosidic linkage, but a C-glycosidic bond seems to be very well compatible with the FAB-MS/MS spectra (8), in that they also showed the loss of 120 Da from $[M + H]^+$.

D. *NMR of peptides containing Trp^{Man}*

^{1}H-NMR spectroscopy of the peptide F-T-W^{Man}-A-Q-W established the residue at position 7 to be a modified tryptophan. All protons of the indole side chain were observed, except for the one at position 2, thus demonstrating the carbohydrate attachment site to be C-2. The combination of ^{1}H- and ^{13}C-NMR data unequivocally showed the presence of an aldohexopyranose, as well as the C-C link to C-2 (2). From vicinal proton-proton coupling constants and ROE intensities it was concluded that the carbohydrate was an α-manno-pyranosyl moiety (3). The monosaccharide did not adopt the usual chair conformation, but appeared to exist in a number of (yet unidentified) different conformations. The availability of NMR data on Trp^{Man} will greatly facilitate the detection of this feature in other peptides. A 1D ^{1}H-NMR spectrum, which can be recorded on amounts as low as 10 nmol of peptide material, should display very similar to identical chemical shifts and coupling constants as observed for Trp^{Man} in F-T-W^{Man}-A-Q-W.

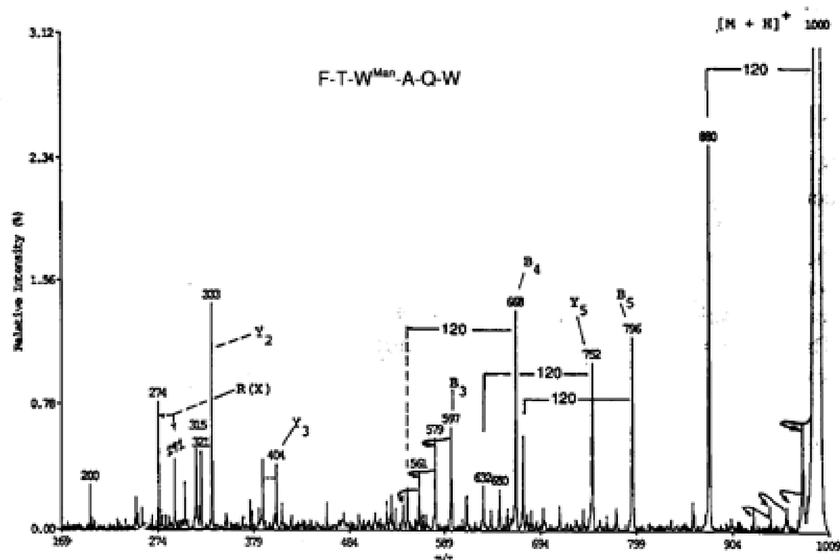


Figure 2 ESI-MS/MS of the hexapeptide F-T-W^{Man}-A-Q-W
Fragments have been labeled using the standard nomenclature. The multiple losses of 18 Da (H_2O) have been indicated by "↪". The loss of 120 Da from fragments containing Trp^{Man} has been indicated, and is typical for aromatic C-glycosidic compounds.

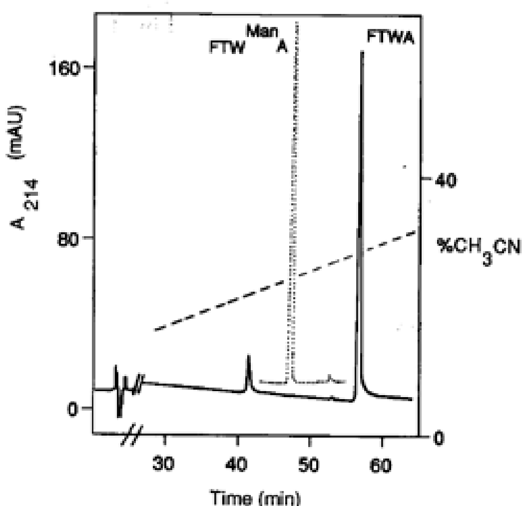


Figure 3 Chromatographic properties of the peptide F-T-W^{Man}-A
Equal amounts of F-T-W-A (full line), or F-T-W^{Man}-A (stippled line) were injected onto a C_{18} reversed phase column equilibrated in 0.1% TFA. The peptides were eluted with a linear gradient of acetonitrile (broken line).

E. Effect of mannosylation of Trp on chromatographic and spectroscopic properties

Fig.3 shows the chromatographic properties of the N-terminal peptide F-T-W^{Man}-A and the synthetic peptide F-T-W-A. The modified peptide eluted considerably earlier from a C_{18} reversed-phase column.

Also the spectroscopic properties of the indole moiety were influenced by the mannosyl residue. The shape of the absorption spectrum of F-T-W^{Man}-A was typical for that of an indole, but its maximum was shifted 2 nm toward longer wavelength and was 1.6-times higher than that of F-T-W-A (Fig.4A). The maximum fluorescence emission increased 2.8-fold, and was shifted 2 nm toward shorter wavelength (Fig. 4B).

F. Protein chemical properties of C-mannosylated peptides

The mannopyranosyl moiety attached to the indole influenced several protein chemical properties of the peptides. Tryptophan can readily be modified at the C-2 position with sulfenylchloride derivatives (9). Treatment of F-T-W-A with o-nitrophenylsulfenylchloride resulted in the conversion of 98% of the peptide into the expected product with a higher retention time on HPLC (Fig.5A), and the characteristic absorbance maximum at 365 nm (data not shown). In contrast, no such product was found with F-T-W^{Man}-A, and 91% of the peptide appeared in an unaltered form (Fig. 5B).

Proteolytic digestion of peptides was affected by the presence of the mannosyl moiety. The peptide comprising residue 1-12 of RNase U_s was not cleaved at position 7 by chymotrypsin, an enzyme with a preference for cleaving after tryptophan (data not shown). Unexpectedly, also the action of exopeptidases was affected by the modification. Cleavage of F-T-W^{Man}-A with Aminopeptidase M resulted in a limited digestion, yielding T-W^{Man}-A as the end product (Fig.6A). In contrast, digestion of the synthetic peptide F-T-W-A yielded free amino acids (Fig. 6B). Apparently, the mannosyl substituent rendered the T-W bond resistant to cleavage. Similarly, the action of carboxypepidases A and Y on F-T-W^{Man}-A was also blocked by the modification (data not shown).

IV. Concluding remarks

At present, two examples of carbohydrate attached to the indole of Trp in a polypeptide have been reported, namely the neuropeptide from *C. morosus* (7), and human RNase U_s (2). Although a C-glycosidic linkage has only been established for the latter, it seems likely that the same occurs in the neuropeptide as well (see above). It remains to be seen how general this kind of modification is, but its occurrence in man and insects suggests a wide phylogenetic distribution. Other questions to be answered are whether other monosaccharides are C-glycosidically linked to Trp, and how the biosynthesis of C-glycosylated Trp takes place. It may be of interest to point out that only

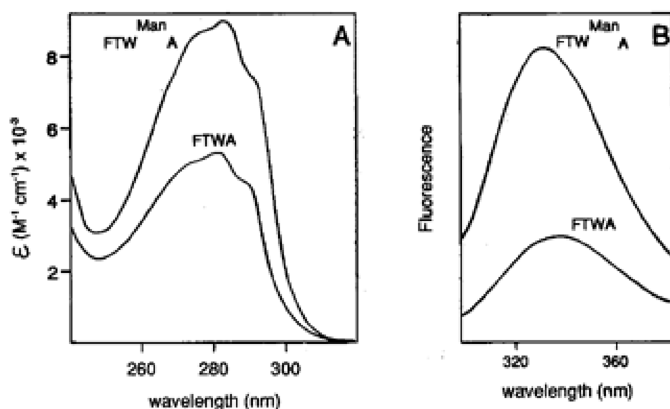


Figure 4 Spectroscopic properties of the peptide F-T-W^{Man}-A
A. Absorbance spectra of the peptides F-T-W-A and F-T-W^{Man}-A.
B. Emission fluorescence spectra of the same peptides; excitation was at 285 nm.

Figure 5 Reaction of the peptide F-T-W^{Man}-A with o-nitrophenyl sulfenylchloride (NPS)
The peptides were treated with NPS as described in "Experimental". The reaction mixture was analysed by reversed phase HPLC. In order to identify the reaction product, the eluate was also monitored at 365 nm (data not shown).
A: F-T-W-A; B: F-T-W^{Man}-A.

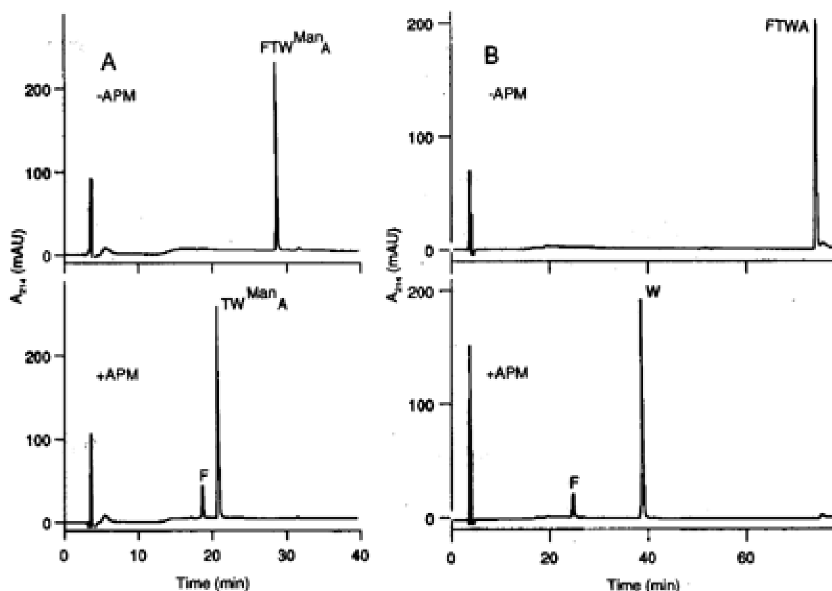


Figure 6 Digestion of F-T-W^{Man}-A with aminopeptidase M (APM)
The peptides were treated with APM for 2 h at 37°C, and the digestion products were separated by reversed phase HPLC. The products were identified by amino acid analysis and MS. A: F-T-W^{Man}-A; B: F-T-W-A

modified Trp-7 has been observed in RNase U_s, and that an unmodified tryptophan occurs in close proximity (position 10). This suggests a certain degree of specificity of the C-glycosylation reaction. In man this modification is not restricted to proteins isolated from urine. Protein chemical and MS analysis of RNase U_s from human erythrocytes demonstrated the presence of Trp^{Man}-7 and unmodified Trp-10 (Löffler et al., manuscript in preparation).

It seems likely that the first indication for the presence of Trp^{Man} in a polypeptide will be obtained by Edman degradation and MS. It should therefore be stressed that under certain conditions PTH-Trp^{Man} and PTH-Tyr coelute. Moreover, the yield of PTH-Trp^{Man} is lower than that of most other amino acids, and varies. It seems prudent to reexamine peptides with an apparently low yield of PTH-Tyr and a 185 Da higher mass (the difference between Tyr and Trp^{Man}) than expected from the protein sequence.

ACKNOWLEDGMENTS

We would like to thank Renate Matthies for help with the protein chemistry, Drs. J. Krieg and D. Hess for reading the manuscript, and the latter also for preparing Figure 1.

REFERENCES

1. Lis, H. and Sharon, N. (1993) Eur. J. Biochem. **218** , 1-27.
2. Hofsteenge, J., Müller, D.R., de Beer, T., Löffler, A., Richter, W.J. and Vliegenhart, J.F.G. (1994) Biochemistry **33**, 13524-13530.
3. de Beer, T., Vliegenthart, J.F.G., Löffler, A., and Hofsteenge, J. Biochemistry, in press.
4. Beintema, J.J., Hofsteenge, J., Iwama, M., Morita, T., Ohgi, K., Irie, M., Sugiyama, R.H., Schieven, G.L., Dekker, C.A. and Glitz, D.G. (1988) Biochem. **27**, 4530-4538.
5. Hamman, K.J., Ten, R.M., Loegering, D.A., Jenkins, R.B., Heise, M.T., Schad, C.R., Pease, L.R., Gleich, G.J. and Barker, R.L. (1990) Genomics **7**, 535-546.
6. Gooley, A.A., Classon, B.J., Marschalek, R. and Williams, K.L. (1991) Biochem. Biophys. Res. Commun. **178**, 1194-1201.
7. Becchi, M, and Fraisse, D. (1989) Biomed Environ Mass Spectrom **18**, 122-130.
8. Gäde, G., Kellner, R., Rinehart, K.L. and Proefke, M.L. (1992) Biochem. Biophys. Res. Comm. **189**, 1303-1309.
9. Fontana, A. and Scoffone, E. (1972) Methods Enzymology **25**, 482-494