

C-mannosylation and O-fucosylation of thrombospondin type 1 repeats

Anne Gonzalez de Peredo¹, Dominique Klein¹, Boris Macek² Daniel Hess¹, Jasna Peter-Katalinic² and Jan Hofsteenge^{1,3}

¹Friedrich Miescher Institute for Biomedical Research, a branch of the Novartis Research Foundation, Maulbeerstrasse 66, CH-4058 Basel, Switzerland, ² Institute for Medical Physics and Biophysics, University of Münster, Robert-Koch-Strasse 31, D-48149 Münster, Germany.

³ To whom correspondence should be addressed: Friedrich Miescher-Institut, Maulbeerstrasse 66, CH-4058 Basel, Switzerland. Tel.: +41 61 697 4722; Fax: +41 61 697 3976; e-mail: hofsteenge@fmi.ch

Running title: Protein C-mannosylation and O-fucosylation

Abbreviations:

ESI MSMS: electrospray ionisation tandem-mass spectrometry

HB-GAM: heparin-binding growth-associated molecule

Q-TOF: quadrupole time-of-flight

rF-spondin: recombinant F-spondin

TSP-1: thrombospondin-1

TSR: thrombospondin type 1 repeat

Summary

The final chemical structure of a newly synthesized protein is often only attained after further covalent modification. Ideally, a comprehensive proteome analysis includes this aspect, a task that is complicated by our incomplete knowledge of the range of possible modifications, and often by the lack of suitable analysis methods. Here we present two recently discovered, unusual forms of protein glycosylation, i.e. C-mannosylation and O-fucosylation. Their analysis by a combined mass spectrometric approach is illustrated with peptides from the thrombospondin type 1 repeats (TSRs) of the axonal guidance protein rF-spondin. Nano-ESI-MSMS of isolated peptides showed that 8 out of 10 Trp residues in the TSRs of F-spondin are C-mannosylated. O-fucosylation sites were determined by a recently established nano-ESI Q-TOF MSMS approach. Four out of 5 TSRs carry the disaccharide Hex-dHex-O-Ser/Thr in close proximity to the C-mannosylation sites. In analogy to thrombospondin-1, we assume this to be Glc-Fuc-O-Ser/Thr. Our current knowledge of these glycosylations will be discussed.

Introduction

Ideally, the complete description of a proteome should include the definition of co- and post-translational modifications. The covalent attachment of carbohydrates is a wide spread feature of secreted, cytoplasmic and nuclear proteins (1). Whereas, N-glycosidic attachment of GlcNAc to Asn residues and GalNAc O-linked to Ser/Thr have been known for a relatively long time (2,3), only more recently the existence of a large variety of protein-carbohydrate linkages has been revealed (1). Here we present two of these, including their analysis by a combined mass spectrometric approach.

The covalent attachment of an α -mannopyranosyl residue to the C2 atom of the side chain of Trp was initially described in human RNase 2. The structure of this glycoconjugate (Fig.1), which does not contain a typical glycosidic linkage but rather a C-C bond, has been based on MS and NMR studies (4,5). The reaction is catalysed by a microsomal transferase, which uses dolichyl-phosphate-mannose as the sugar donor and recognizes, in nearly all cases, the sequence Trp-x-x-Trp (6,7). The transferase activity can be detected in organisms ranging from *C. elegans* to man (8). Presently, a total of 49 C-mannosylated tryptophan residues have been identified, derived from 11 different proteins, performing a wide variety of functions [for a review see: (9)]. Evidence for C-mannosylation and for its site of attachment can conveniently be obtained in CID MSMS experiments of isolated peptides. Here we show examples of this kind of analysis and summarize our current knowledge of this modification.

Fucose was long considered to be a terminal sugar of oligosaccharides. This changed, however, when fucose O-linked to Ser or Thr was discovered in the EGF-like modules from proteins that play an important role in fibrinolysis and coagulation [for an overview, see (10,11)]. There exist at least two pathways for the extension of O-linked fucose (12). The first one leads to the formation of the tetrasaccharide NeuAc- α 2-6-Gal- β 1,4-GlcNAc- β 1,3-Fuc- α 1-O-Ser/Thr, as found in factor IX and Notch (13-15). For Notch, it has been shown that

receptor-ligand interactions are modulated by extension of the of O-linked fucosyl residues (16,17). The second pathway yields the disaccharide Glc- β 1,3-Fuc- α 1-O-Ser/Thr, first found as the amino acid glycoside in human urine (18). Recently, we discovered that all three TSRs in TSP-1 carry this disaccharide in close proximity to C-mannosylation sites. Nano-ESI Q-TOF mass spectrometry has proven to be very useful in the determination of O-fucosylation sites (19,20). Here we use the same approach in the analysis of rF-spondin. This protein also contains TSRs, and is secreted by the floor plate during development of the nervous system (21). It functions as a guidance protein for commissural neurons, by attracting their axons to the midline during rostral growth.

Experimental Procedures

Proteins

Recombinant F-spondin was expressed in COS-7 cells¹. Plasmid pSecF-spo-His, encoding residues 27-752 of rat F-spondin followed by a myc tag and a His₆ sequence was a generous gift of Dr. A. Klar, the Hebrew University-Hadassah Medical School, Jerusalem (21).

Peptides for the analysis of the C-mannosylation and O-glycosylation sites were obtained from aminoethylated- and carbamidomethylated rF-spondin, respectively. The former was cleaved with endoproteinase Lys-C and trypsin, whereas the latter was digested with trypsin.

Digests were fractionated by C8 reversed phase LC-ESIMS using a Sciex API 300 triple quadrupole mass spectrometer, as previously described (22,23). Fractions were collected and analysed using the same instrument. Peptides have been numbered according to their occurrence in mature F-spondin (21),

The analysis of C-mannosylated peptides by nano ESIMSMS using a Sciex API 300 triple quadrupole mass spectrometer and Edman degradation has been described previously (23).

Determination of O-fucosylation sites was performed by low energy CID tandem-MS in the nanospray mode, using a Q-TOF instrument (Micromass, UK), equipped with a Z-spray atmospheric pressure ionisation source operated at 30°C. The pressure of the collision gas (Ar) in the hexapole collision cell was 2.7×10^{-5} mbar and the collision energy was typically 20 eV. Details of the method have been described. (19).

Results

C-Mannosylation

C-mannosylation occurs only on secreted proteins², which in addition to this modification often also contain N- and/or O-linked oligosaccharides. This, and the fact that only a single

mannosyl residue (mass 162 Da) is added in the process of C-mannosylation, makes it difficult to detect (C²-Man-)Trp in the intact protein. Therefore, the isolation and analysis of relevant peptides is required. Towards that aim, a proteolytic digest of the protein is fractionated by reversed phase LC-MS. Subsequently, the MS data are extracted for the theoretical masses of the peptides containing the recognition sequence Trp-x-x-Trp, assuming the presence of one or more C-linked mannosyl residues. As an example, the detection of peptide KT81 (LVTEWGEWDDC⁵⁴⁵), one of the C-mannosylated peptides from recombinant F-spondin (rF-spondin), is shown in Fig. 2A. The peptide, containing a single C-mannosylated Trp residue (theoretical mass: 1396 Da, av), is identified eluting at 38.3 min. Strong indication for the presence of a (C²-Man-)Trp residue can already be obtained at this stage. Due to the stability of the C-C bond in aromatic C-glycosides, these compounds display a characteristic loss of 120 Da in low energy CID experiments (4,24,25). Probably as a result from in source CID, this phenomenon can also be observed in LC-MS experiments, as illustrated for peptide KT81 in Fig. 2B. This behaviour clearly distinguishes C-mannosylated peptides from ones that contain O-linked mannose, which show a loss of 162 Da under similar LC-MS conditions (Fig.3). Information on the position of the modified Trp residue can be obtained from an MSMS experiment. The results (Fig.4) confirm the identity of the peptide and provide strong evidence for C-hexosylation of Trp-539. Since all hexoses add a mass of 162 Da to the Trp residue, MS does not reveal the identity of the sugar involved. Initially the proof for a C- α -mannosyl residue has been obtained by NMR experiments on peptides from human RNase 2 (4), as well as on the intact protein (26). Furthermore, NMR analysis of peptides from IL-12 β and complement factor C9 also yielded mannose as the sugar present (23,27). In addition, solid phase Edman degradation can be used for the analysis of smaller amounts of the peptides. Using an optimised elution protocol, PTH-(C²-Man-)Trp elutes

shortly after PTH-Tyr (28). The elution position of PTH-(Hex)Trp, with hexoses other than Man, remains to be established.

Following this approach we have detected 48 (C2-Man-Trp) residues in 10 different proteins (Table 1). In addition, it seems probable that also a neuropeptide from *Carausius morosus* is C-glycosylated (29).

O-Fucosylation

The recent discovery of the O-linked disaccharide Glc-Fuc-O-Ser/Thr in TSP-1, raises the question whether the same modification also occurs in other proteins that contain TSRs. We examined recombinant rF-spondin and properdin, and show results obtained with the former to illustrate the approach. Relevant peptides were produced by tryptic digestion of the carbamidomethylated protein. To simplify the analysis, a secondary digestion with an appropriate protease was performed when needed. Cleavage of peptide T61 with endoproteinase Asp-N yields peptide T61D (DDCS⁵⁴⁵AT⁵⁴⁸CGMGMK) with a monoisotopic mass of 1639.59 Da, which is 308 Da heavier than predicted from the cDNA sequence.

Together with the partial loss of the substituent in low energy LC-MS experiments (data not shown), this strongly suggested the presence of O-linked sugars. Since nano Q-TOF MSMS allows the direct determination of glycosylation sites in O-fucosylated glycopeptides (19), we analysed all potential O-fucosylation sites in rF-spondin by this method. As an example, the results obtained with the doubly charged molecular ion of peptide T61D are shown in Fig.5. The observation of B fragment ions [nomenclature of Domon and Costello, (30)] of the Hex-dHex disaccharide ($m/z = 309.13$), the Hex ($m/z = 163.07$) and dHex ($m/z = 147.09$)³ as well as ions corresponding to their neutral loss of water, provides evidence for the presence of a disaccharide, rather than two monosaccharides. Furthermore, the consecutive loss of the Hex- and dHex- residues, strongly indicates that the disaccharide is attached through the dHex. In

addition to the deglycosylated fragment ions that confirmed the identity of the peptide, ions y7, y9 and b11 were observed with Hex-dHex and dHex still attached (Fig.5). Despite the sometimes low abundance of these ions, their signals displayed good signal-to-noise ratios and high resolution (Fig.5, inset). The results strongly indicated that Thr-548 carries the Hex-dHex disaccharide. Following this approach, the same O-glycosylations were also found in TSR-1, -2 and -4. Table 2 summarizes all O-fucosylation sites that have been found in TSRs to date. It should be noted that whereas for TSP-1 and properdin, the identity of the monosaccharides have been shown to be Glc and Fuc, this has not yet been determined for the peptides from rF-spondin.

Discussion

C-Mannosylation

The results presented here emphasize the abundance of C-mannosylation in TSRs. In fact, all 24 TSRs examined to date, are modified in this way (Table 1). However, other proteins, and possibly other protein families, also contain this glycosylation. For example, IL-12 β is a member of the type 1 cytokine receptor superfamily and C-mannosylated on Trp-319 (27). Interestingly, nearly all members of this superfamily contain the so-called "WSXWS" box (31). The analysis of these proteins will be of interest, because mutations in this box in the erythropoietin receptor, cause its transport to the cell surface to stop at the ER-Golgi interface (32).

The stoichiometry of C-mannosylation shown in Table 1 differs for different Trp residues. In 35 cases full modification was observed, whereas in the remainder, it has been estimated to vary from <10-95% (see the original references for details). Estimating the degree of modification can be difficult. In case the modified and unmodified peptides are obtained in pure form from the initial LC-MS experiment, their ratio can be obtained by integrating their

peaks in the UV trace. If contamination with other peptides occurs, this procedure can not be used. An approximate ratio may be obtained from integration of the peaks in the total ion current trace. However, this will only give very approximate estimates, due to e.g differences in ionization efficiency of modified and unmodified peptides [for a discussion see (23)].

Clearly, more work is needed on this aspect of the analysis of C-mannosylation and covalent modifications in general.

The results obtained with TSRs raise a question with respect to the specificity of C-mannosylation. Both studies in vivo with RNase 2, and in vitro studies using synthetic peptides have revealed that the recognition motif for the transferase is Trp-x-x-Trp, where only the first Trp becomes modified (6,7,22,27). This specificity does, however, not hold for the TSRs. Of the 24 C-mannosylated TSRs, 22 contain the sequence Trp-x-x-Trp-x-x-x, without a Trp or other aromatic residue at position +3 relative to the second Trp. In 7 of these, only the first Trp is modified, in agreement with the RNase 2 results. However, in the other 15, both Trp residues are modified. Similarly, 2 TSRs contain the C-mannosylated sequence ϕ -x-x-Trp (ϕ = Tyr, Phe) (Table 1). A possible explanation for these observations is that features outside of the motif determine whether one or both Trps become modified. Alternatively, more than one C-mannosyltransferase might exist. Cloning of the enzyme(s) catalysing this reaction and the study of their enzymological properties, should resolve this issue.

Based on antibody inhibition studies, it has been proposed that the activity of F-spondin in axonal guidance is located in the TSRs (33). Whether the Trp-x-x-Trp motifs are involved is not known. By far the most studies on this motif have been performed with TSP-1. Synthetic peptides containing C-mannosylation motifs were shown to bind to and/or inhibit the interactions of TSP-1 with fibronectin, TGF- β and heparin (34,35). It should be noted, however, that in all cases non-C-mannosylated peptides were used [for a review see (34)]. The

two Trp residues were shown to be important for the interaction with heparin. Interestingly, the neurite extending activity of F-spondin can be inhibited by this glycosaminoglycan.

O-Fucosylation

The gas phase instability of O-linked fucose poses a challenging problem for the determination of the exact attachment site in peptides and proteins by MS methods. Due to the high sensitivity and excellent signal-to noise ratio of Q-TOF instruments, it is often possible to successfully characterize this modification in peptides. The sugar-containing b- and y-ions, needed for the identification, are always detected at 10-100 fold lower levels than the corresponding unglycosylated ones. Using low pmol amounts of peptide, low collision energy and long acquisition times (see 'Experimental Procedures'), glycosylated ions can often unambiguously be identified with nano ESI Q-TOF MSMS.

Using this approach, we have identified 3 O-fucosylation sites in TSP-1, 4 in properdin and 4 in rF-spondin (note that in this case the identity of the sugars remains to be determined).

Whereas TSR2-4 in rF-spondin are fully modified at the sites indicated in Table 2, TSR1 appears to be modified for less than 10%. In properdin, O-fucosylation was complete in TSR2-4, but only approximately 50% in TSR1.

Examination of the TSR superfamily reveals that 63 out of 88 TSRs contain a Ser or Thr at the position corresponding to the O-fucosylation sites reported here (20). This suggests that the modification occurs in many more proteins. In all cases the O-fucosylation and C-mannosylation sites are in close proximity, yet it seems unlikely that one requires the other. Protein chemical studies on recombinant TSR3 from TSP-1 revealed the presence of (C2-Man-)Trp, Glc-Fuc-O-Thr and Fuc-O-Thr, which occurred in all six possible combinations, strongly suggesting that the two glycosylations occur independently (20).

Three forms of O-fucosylation are presently known, attachment of : (i) a single fucosyl residue to the insect proteinase inhibitor PMP-C and EGF modules (10,36), (ii) NeuAc-Gal-GlcNAc-Fuc- to EGF modules (37) and (iii) Glc-Fuc- to TSR modules (20);(this paper). In addition to the functional studies on Notch mentioned in the introduction, also those on EGF-like domains in other proteins suggest that O-fucosylation has a function. Urokinase-type plasminogen activator is mitogenic for a number of cells in culture, and this activity depends on its EGF-like domain that carries a fucosyl residue on Thr-18. The defucosylated domain binds as tightly to cells as the fucosylated one, but it is not able to produce the mitogenic signal (38-40). The effect of O-linked fucose on protein structure has been examined in PMP-C (41) and the EGF-like domain derived from blood coagulation factor VII (42). In both cases little difference in structure was found between the protein with and without fucose. However, the dynamic fluctuations of PMP-C were decreased and its stability increased, upon O-fucosylation. The overall structures of PMP-C and the EGF-like module from factor VII are different. However, both are all- β proteins with a central β -sheet and three disulfide bridges. TSR modules are likely to have a β -sheet structure as well, based on their homology with the all- β proteins HB-GAM and midkine (43). If this proposition is correct, the O-fucosylation site in TSRs is located in a loop between the first two β -sheets. What is the function of O-fucosylation in TSRs, is not known yet. However, in a number of TSR proteins the sequence in which O-fucosylation takes place has been implicated in cell binding. For instance, the cell surface receptor CD36 that mediates the inhibition of neovascularization by TSP-1 in vivo, has been proposed to bind to sequences CSVTCG [for a review see (34)]. It is important to note, that unmodified synthetic peptides were used in these studies.

The Golgi is the only secretory compartment containing a transporter for GDP-Fuc (44). It seems likely that O-fucosylation of TSR modules, like that of EGF-modules (10), takes place there. Leukocyte adhesion deficiency II has recently been attributed to an aberrant form of the

GDP-Fuc transporter, causing a general lack of fucosylation of glycoconjugates (45,46). The leukocyte-associated defects in this disease can be understood by the lack of α 1,3-linked fucose in Lewis x antigen- based selectin ligands (47). However, the non-fucosylated glycoconjugates that cause the defects in mental development are unknown. It is, therefore, interesting to point out that the TSR proteins TSP-1, F-spondin, Unc-5 and the semaphorins 5A and B all play a role in neuronal development (34). Because O-fucosylation has now been found in the first two of these proteins, and several putative sites are also present in the others, analysis of their O-fucosylation status and functional properties in the context of the disease seems warranted.

Acknowledgements

We would like to thank Dr. Avihu Klar, Hebrew University, Jerusalem for providing the rF-spondin expression vector and Renate Matthies for help with protein sequencing.

References

1. Varki, A., Cummings, C., Esko, J. D., Freeze, H., Hart, G. W., Marth, J. D. (1999) Essentials of Glycobiology, 1st Ed, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York
2. Gottschalk, A. (1966) Historical introduction in Glycoproteins (Gottschalk, A., ed) pp. 1-19, Elsevier Publishing Company, Amsterdam
3. Gottschalk, A. (1966) Carbohydrate-peptide linkages in glycoproteins and methods for their elucidation in Glycoproteins (Gottschalk, A., ed) pp. 273-295, Elsevier Publishing Company, Amsterdam
4. Hofsteenge, J., Müller, D. R., de Beer, T., Löffler, A., Richter, W. J., Vliegthart, J. F. G. (1994) New type of linkage between a carbohydrate and a protein: C- glycosylation of a specific tryptophan residue in human RNase Us. *Biochemistry*. 33, 13524-13530
5. de Beer, T., Vliegthart, J. F. G., Löffler, A., Hofsteenge, J. (1995) The hexopyranosyl residue that is C-glycosidically linked to the side chain of tryptophan-7 in human RNase Us is α -mannopyranose. *Biochemistry*. 34, 11785-11789
6. Krieg, J., Hartmann, S., Vicentini, A., Gläsner, W., Hess, D., Hofsteenge, J. (1998) The recognition signal for C-mannosylation of Trp-7 in RNase 2 consists of the sequence Trp-x-x-Trp. *Mol.Biol.Cell*. 9, 301-309
7. Doucey, M.-A., Hess, D., Cacan, R., Hofsteenge, J. (1998) Protein C-mannosylation is enzyme-catalysed and uses dolichyl-phosphate-mannose as a precursor. *Mol.Biol.Cell*. 9, 291-300
8. Doucey, M.-A. (1998) La C-mannosylation du tryptophane: un nouveau type de glycosylation des protéines, Université Louis Pasteur de Strassbourg
9. Furmanek, A., Hofsteenge, J. (2000) Protein C-mannosylation: facts and questions. *Acta Biochim Pol*. 47, 781-789

10. Harris, R. J., Spellman, M. W. (1993) O-linked fucose and other post-translational modification unique to EGF modules. *Glycobiology*. 3, 219-224
11. Vliegthart, J. F. G., Casset, F. (1998) Novel forms of protein glycosylation. *Current Opinion in Structural Biology*. 8, 565-571
12. Moloney, D. J., Lin, A. I., Haltiwanger, R. S. (1997) The O-linked fucose glycosylation pathway. *J. Biol. Chem.* 272, 19046-19050
13. Harris, R. J., van Halbeek, H., Glushka, J., Basa, L. J., Ling, V. T., Smith, K. J., Spellman, M. W. (1993) Identification and structural analysis of the tetrasaccharide NeuAc $\alpha(2\rightarrow6)$ Gal $\beta(1\rightarrow4)$ GlcNAc $\beta(1\rightarrow3)$ Fuc $\alpha 1\rightarrow$ O-linked to serine 61 of human factor IX. *Biochemistry*. 32, 6539-6547
14. Nishimura, H., Takao, T., Hase, S., Shimonishi, Y., Iwanaga, S. (1992) Human factor IX has a tetrasaccharide O-glycosidically linked to Serine 61 through the fucose residue. *J. Biol. Chem.* 267, 17520-17525
15. Moloney, D. J., Shair, L. H., Lu, F. M., Xia, J., Locke, R., Matta, K. L., Haltiwanger, R. S. (2000) Mammalian Notch1 is modified with two unusual forms of O-linked glycosylation found on epidermal growth factor-like modules. *J Biol Chem.* 275, 9604-9611
16. Moloney, D. J., Panin, V. M., Johnston, S. H., Chen, J., Shao, L., Wislon, R., Wang, Y., Stanley, P., Irvine, K. D., Haltiwanger, R. S., Vogt, T. F. (2000) Fringe is a glycosyltransferase that modifies notch. *Nature*. 406, 369-375
17. Brückner, K., Perez, L., Clausen, H., Cohen, S. (2000) Glycosyltransferase activity of Fringe modulates Notch-Delta interactions. *Nature*. 406, 411-415
18. Hallgren, P., Lundblad, A., Svensson, S. (1975) A new type of carbohydrate-protein linkage in a glycopeptide from normal human urine. *J. Biol. Chem.* 250, 5312-5314

19. Macek, B., Hofsteenge, J., Peter-Katalinic, J. (2001) Direct determination of glycosylation sites in O-fucosylated glycopeptides using nano-electrospray quadrupole time-of-flight mass spectrometry. *Rapid Commun Mass Spectrom.* 15, 771-777
20. Hofsteenge, J., Huwiler, K. G., Macek, B., Hess, D., Lawler, J., Mosher, D. F., Peter-Katalinic, J. (2001) C-Mannosylation and O-Fucosylation of the Thrombospondin Type 1 Module. *J Biol Chem.* 276, 6485-6498
21. Klar, A., Baldassare, M., Jessell, T. M. (1992) F-sponding: A gene expressed at high levels in the floor plate encodes a secreted protein that promotes neural cell adhesion and neurite extension. *Cell.* 69, 95-110
22. Krieg, J., Gläsner, W., Vincentini, A., Doucey, M.-A., Löffler, A., Hess, D., Hofsteenge, J. (1997) Protein C-mannosylation is an intracellular process, performed by a variety of cultured cells. *J. Biol. Chem.* 272, 26687-26692
23. Hofsteenge, J., Blommers, M. M., Hess, D., Furmanek, A., Miroshnichenko, O. (1999) The four terminal components of the complement system are C-mannosylated on multiple tryptophan residues. *J. Biol. Chem.* 274, 32786-32794
24. Li, Q. M., van den Heuvel, H., Dillen, L., Claeys, M. (1992) Differentiation of 6-C and 8-C-glycosidic flavonoids by positive ion fast atom bombardment and tandem mass spectrometry. *Biological Mass Spectrometry.* 21, 213-221
25. Becchi, M., Fraisse, D. (1989) Fast atom bombardment and fast atom bombardment collision-activated dissociation/mass-analysed ion kinetic energy analysis of C-glycosidic flavonoids. *Biomedic.and Environmental Mass Spectrometry.* 18, 122-130
26. Löffler, A., Doucey, M.-A., Jansson, A. M., Müller, D. R., de Beer, T., Hess, D., Meldal, M., Richter, W. J., Vliegthart, J. F. G., Hofsteenge, J. (1996) Spectroscopic and protein chemical analyses demonstrate the presence of C-mannosylated tryptophan in intact human RNase 2 and its isoforms. *Biochemistry.* 35, 12005-12014

27. Doucey, M.-A., Hess, D., Blommers, M. J., Hofsteenge, J. (1999) Recombinant human interleukin-12 is the second example of a C-mannosylated protein. *Glycobiology*. 9, 435-441
28. Hofsteenge, J., Löffler, A., Müller, D. R., Richter, W. J., de Beer, T., Vliegthart, J. F. G. (1996) Protein C-glycosylation in *Techniques in protein chemistry VII* (Marshak, D. R., ed) pp. 163-171, Academic Press, New York
29. Gäde, G., Keller, R., Rinehart, K. L., Proefke, M. L. (1992) A tryptophan-substituted member of the AKH/RPCH family isolated from a stick insect corpus cardiacum. *Biochem.Biophys.Res.Comm.* 189, 1303-1309
30. Domon, B., Costello, C. E. (1988) A systematic nomenclature for carbohydrate fragmentation in FAB-MS/MS spectra of glycoconjugates. *Glycoconjugate J.* 5, 397-409
31. Bazan, J. F. (1990) Structural design and molecular evolution of a cytokine receptor superfamily. *Proc.Natl.Acad.Sci.* 87, 6934-6938
32. Hilton, D. J., Watowich, S. S., Katz, L., Lodish, H. F. (1996) Saturation mutagenesis of the WSXWS motif of the erythropoietin receptor. *Journal of Biological Chemistry.* 271, 4699-4708
33. Burstyn-Cohen, T., Tzarfaty, V., Frumkin, A., Feinstein, Y., Stoeckli, E., Klar, A. (1999) F-Spondin is required for accurate pathfinding of commissural axons at the floor plate. *Neuron.* 23, 233-246
34. Adams, J. C., Tucker, R. P. (2000) The thrombospondin type 1 repeat (TSR) superfamily: Diverse proteins with related roles in neuronal development. *Developm. Dynam.* 218, 280-299
35. Lawler, J. (2000) The functions of thrombospondin-1 and-2. *Curr Opin Cell Biol.* 12, 634-640

36. Nakakura, N., Hietter, H., Van Dorselaer, A., Luu, B. (1992) Isolation and structural determination of three peptides from the insect *Locusta migratoria*. Identification of a deoxyhexose-linked peptide. Eur. J. Biochem. 204, 147-153
37. Moloney, D. J., Shair, L. H., Lu, F. M., Xia, J., Locke, R., Matta, K., Haltiwanger, R. S. (2000) Mammalian Notch 1 is modified with two unusual forms of O-linked glycosylation found on epidermal growth factor-like modules. J. Biol. Chem. 275, 9604-9611
38. Rabbani, S. A., Mazar, A. P., Bernier, S. M., Haq, M., Bolivar, I., Henkin, J., Goltzman, D. (1992) Structural requirements for the growth factor activity of the amino-terminal domain of urokinase. J Biol Chem. 267, 14151-14156
39. Schnaper, H. W., Barnathan, E. S., Mazar, A., Maheshwari, S., Ellis, S., Cortez, S. L., Baricos, W. H., Kleinman, H. K. (1995) Plasminogen activators augment endothelial cell organization in vitro by two distinct pathways. J Cell Physiol. 165, 107-118
40. Koopman, J. L., Slomp, J., de Bart, A. C., Quax, P. H., Verheijen, J. H. (1998) Mitogenic effects of urokinase on melanoma cells are independent of high affinity binding to the urokinase receptor. J Biol Chem. 273, 33267-33272
41. Mer, G., Hietter, H., Lefevre, J.-F. (1996) Stabilization of proteins by glycosylation examined by NMR analysis of a fucosylated inhibitor. Nature Struct.Biol. 3, 45-52
42. Kao, Y. H., Lee, G. F., Wang, Y., Starovasnik, M. A., Kelley, R. F., Spellman, M. W., Lerner, L. (1999) The effect of O-fucosylation on the first EGF-like domain from human blood coagulation factor VII. Biochemistry. 38, 7097-110
43. Kilpeläinen, I., Kaksonen, M., Kinnunen, T., Avikainen, H., Fath, M., Linhardt, R., Raulo, E., Rauvala, H. (2000) Heparin-binding growth-associated molecule contains two heparin-binding β -sheet domains that are homologous to the thrombospondin type 1 repeat. J. Biol.Chem. 275, 13564-13570

44. Abeijon, C., Mandon, E. C., Hirschberg, C. B. (1997) Transporters of nucleotide sugars, nucleotide sulfate and ATP in the golgi apparatus. *TIBS*. 22, 203-207
45. Luhn, K., Wild, M. K., Eckhardt, M., Gerardy-Schahn, R., Vestweber, D. (2001) The gene defective in leukocyte adhesion deficiency II encodes a putative GDP-fucose transporter. *Nat Genet*. 28, 69-72
46. Lubke, T., Marquardt, T., Etzioni, A., Hartmann, E., von Figura, K., Korner, C. (2001) Complementation cloning identifies CDG-IIc, a new type of congenital disorders of glycosylation, as a GDP-fucose transporter deficiency. *Nat Genet*. 28, 73-76
47. Becker, D. J., Lowe, J. B. (1999) Leukocyte adhesion deficiency type II. *Biochim Biophys Acta*. 1455, 193-204
48. Hartmann, S., Hofsteenge, J. (2000) Properdin, the positive regulator of the complement, is highly C-mannosylated. *J. Biol. Chem*. 275, 28569-28574

Footnotes

¹ Gonzalez et al., manuscript in preparation

² Smilda et al. manuscript in preparation

³ the resolution was sufficient to distinguish the sugar oxonium ion from the y1 ion.

Figure legends

Figure 1

A. The structure of C2- α -mannopyranosyltryptophan (4,5)

B. Schematic structure of the domains of F-spondin. Thrombospondin type 1 repeats are indicated in black, and partial sequences relevant to the analyses have been shown. Potential C-mannosylation and O-glycosylation sites have been underlined. Note that the sixth TSR of F-spondin (21) is not present in the recombinant protein used in this study.

Figure 2

Peptide mapping of rat rF-spondin. A. Peptides from an endoproteinase Lys-C/trypsin digest of reduced and aminoethylated rF-spondin were fractionated by LC-ESIMS, using a C₁₈ column. The experiment was performed on an API 300 triple quadrupole mass spectrometer. The broken line shows the result of the extraction of the MS data for m/z 779, corresponding to the doubly charged ion of peptide KT61 including one hexosyl residue. B. Portion of the LC-MS spectrum of peptide KT61, showing the loss of 120 Da from the doubly charged ion (m/z = 60), characteristic for aromatic C-glycosides.

Figure 3

Comparison of the behaviour of C- and O-linked mannose in LC-ESIMS experiments. The experiment was performed on an API 300 triple quadrupole mass spectrometer. A. Spectrum of the peptide FT(C2-Man-)WA isolated from human RNase 2. B. Spectrum of the synthetic peptide F(O-Man-)TWA.

Figure 4

Low energy CID tandem-ESIMS of the C-mannosylated peptide KT61 from rF-spondin. The experiment was performed on an API 300 triple quadrupole mass spectrometer. The 120-Da loss, typical for aromatic C-glycosides has been indicated with *120* and *60* for singly and doubly charged ions, respectively. Losses of H₂O have been indicated with asterisks. W#: C-mannosylated tryptophan. Cysteine residues have been aminoethylated.

Figure 5

Low energy CID nano-ESI Q-TOF tandem MS sequencing of the glycopeptide T81D from rF-spondin. For details of the method see 'Experimental Procedures'. The cysteines residues have been carbamidomethylated.

Table 1 Summary of C-mannosylated sequences.

Protein	Function	Number of (C2-Man-)Trp	Sequence ^a	Reference
RNase 2	Host defence (?)	1	TWAQWF	(4)
Interleukin 12 β	Cytokine receptor homologue	1	SWSEWA	(27)
Hypertrehalosemic hormone	Neuropeptide	1 (?) ^b	F'FPNWGTG	(29)
<u>Proteins with TSRs:</u>				
C6	Complement; forms together with C5, C7, C8 α , C8 β and C9 the membrane attack complex	6	DHY AWTQWT LLGDFGP WS Q WGCWSSWS	(23)
C7	Complement	4	QWDFY APWS G WSCWSSWS	(23)
C8 α	Complement	4	QLSN WSEWT S WSCWSSWS	(23)
C8 β	Complement	4	ELSS WSSWT KWNC WSNWS	(23)
C9	Complement	2	RMSP WSEWS	(23)
Properdin	Positive regulator of complement	15	RW SLWSTWA G WSGWGPWE A WATWGPWT G WGPWGPVS E WDSWGEWS S WSEWSTWG	(48)
Thrombospondin-1	TGF- β activation Angiogenesis inhibitor	4	G WSPWSEWT G WSHWSPWS G WGPWSPWD	(20)
F-spondin	Axonal guidance	8	I YSNWSPWS T MSEWITWS L VTEWGEWD L LLSPWSEWS E LSEWSQWS	Footnote 2

Footnotes

^a C-mannosylated Trp residues have been indicated in bold (fully modified) or bold, italic (partially modified). For estimates of the degree of modification, see the original references.

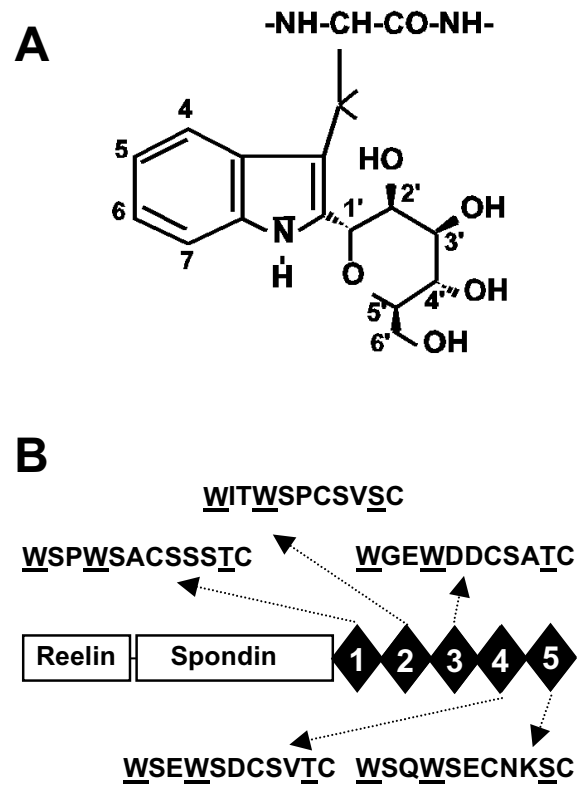
^b Gäde et al hypothesize that the unknown hexose is attached to N1 of the indole.

Table 2 O-fucosylation sites in the TSRs of TSP-1, properdin and rF-spondin^a

tsp1-tsr1	DGWSPWSEWTSC <u>ST</u> SCGNG
tsp1-tsr2	GGWSHWSPWSSC <u>SV</u> TCGDG
tsp1-tsr3	GGWGPWSPWDIC <u>SV</u> TCGGG
properdin-tsr1	PRWSLWSTWAPC <u>SV</u> TCSEG
properdin-tsr2	GGWSGWGPWEPC <u>SV</u> TCSKG
properdin-tsr3	GAWATWGPWTPC <u>SAS</u> CHGG
properdin-tsr4	GGWGPWGPVSPC <u>PV</u> TCGLG
properdin-tsr5	GEWDSWGEWSPCIRRNMKSISCQ
properdin-tsr6	GSWSEWSTWGLC MPPC . . .
f-spondin-tsr1	CIYSNWSPWSAC <u>SS</u> <u>t</u> CEK G ^b
f-spondin-tsr2	CTMSEWITWSPC <u>SV</u> SCGMG
f-spondin-tsr3	CLVTEWGEWDDC <u>SAT</u> CGMG
f-spondin-tsr4	CLLSPWSEWSDC <u>SV</u> TCGKG
f-spondin-tsr4	CELSEWSQWSEC NKSCGKG
consensus	W X ₅ C X _{2/3} <u>S/TC</u> X ₂ G

^a O-glycosylated Ser or Thr residues have been underlined, and C-mannosylation motifs are shown in italics. The consensus sequence shows the completely conserved residues.

^b TSR1 of rF-spondin is O-glycosylated, but the exact site of attachment could not yet be determined. This has been indicated in lower case.



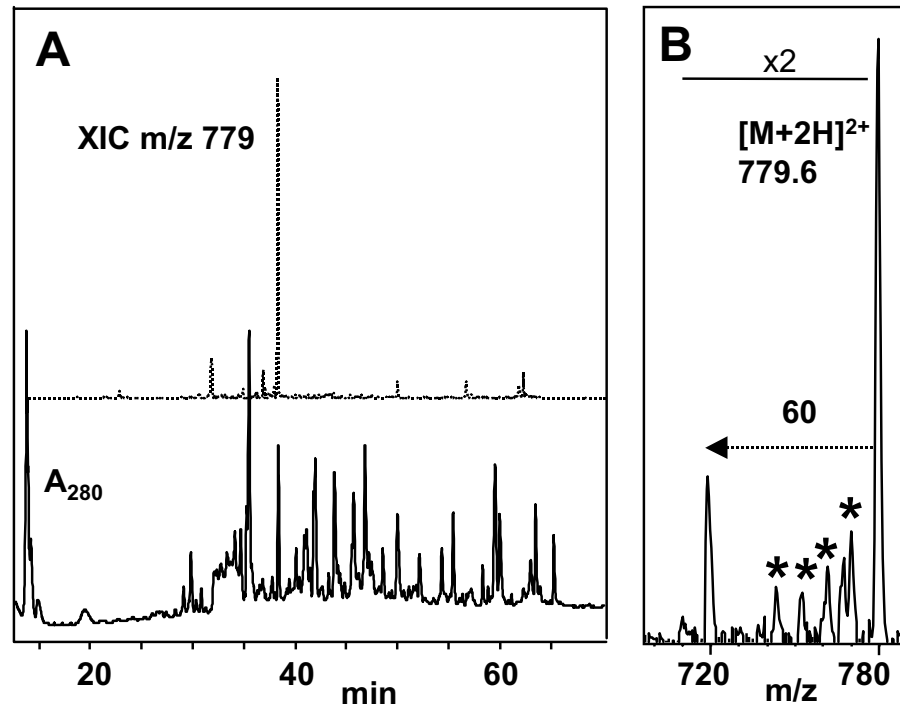


Fig.2 Hofsteenge et al.

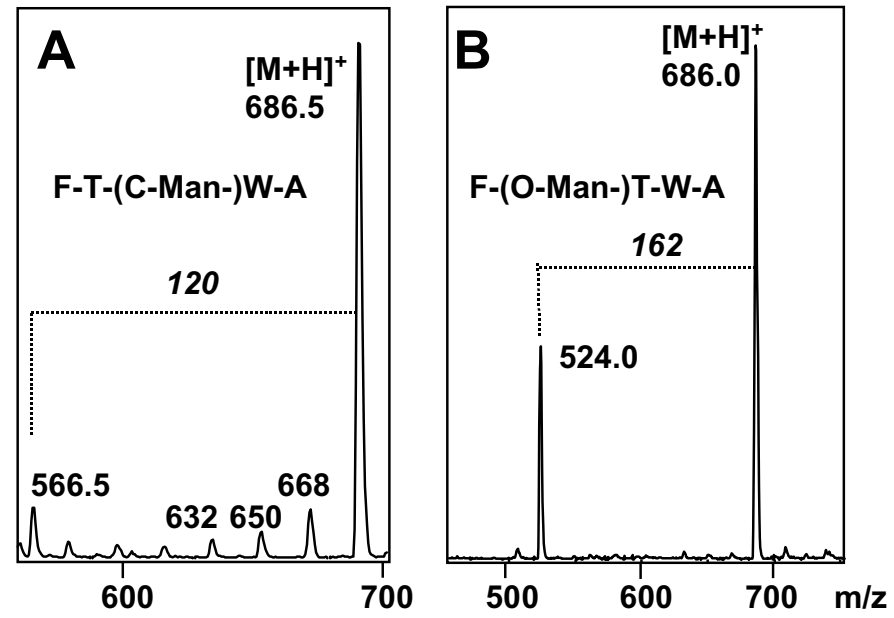


Fig.3 Hofsteenge et al.

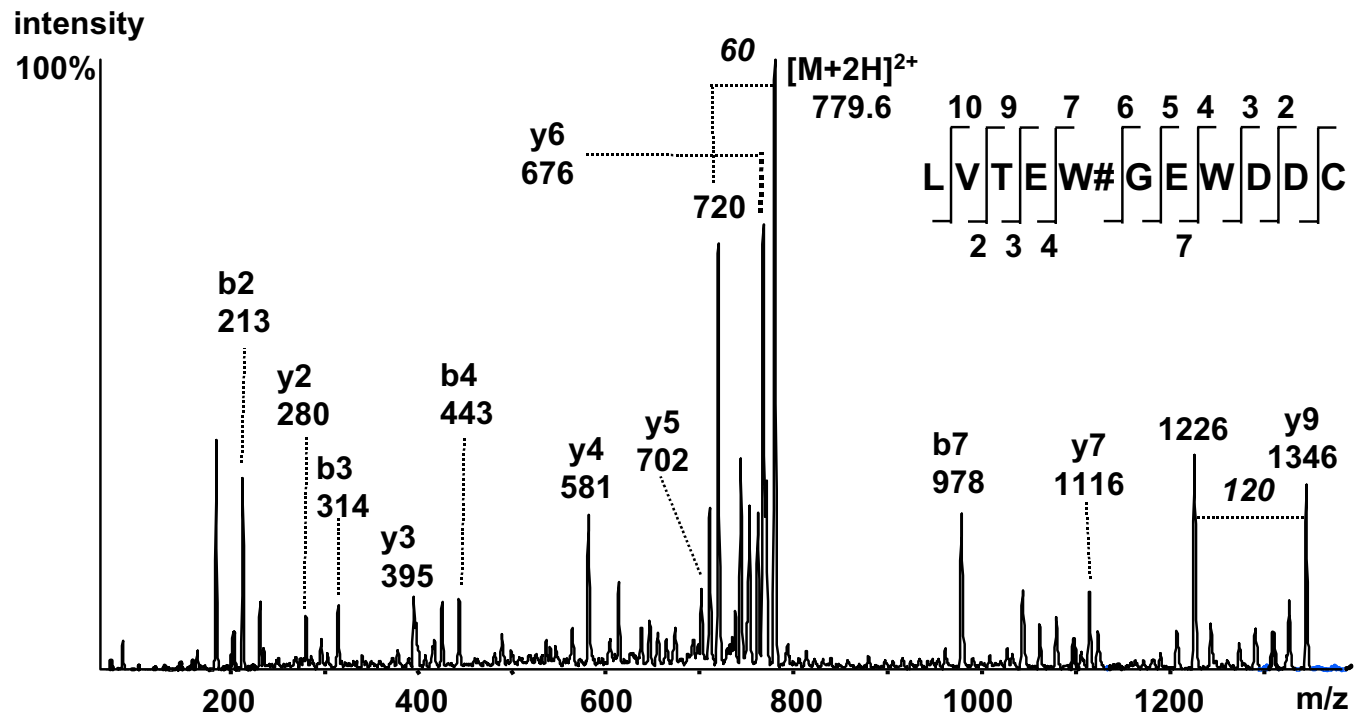


Fig.4 Hofsteenge et al.

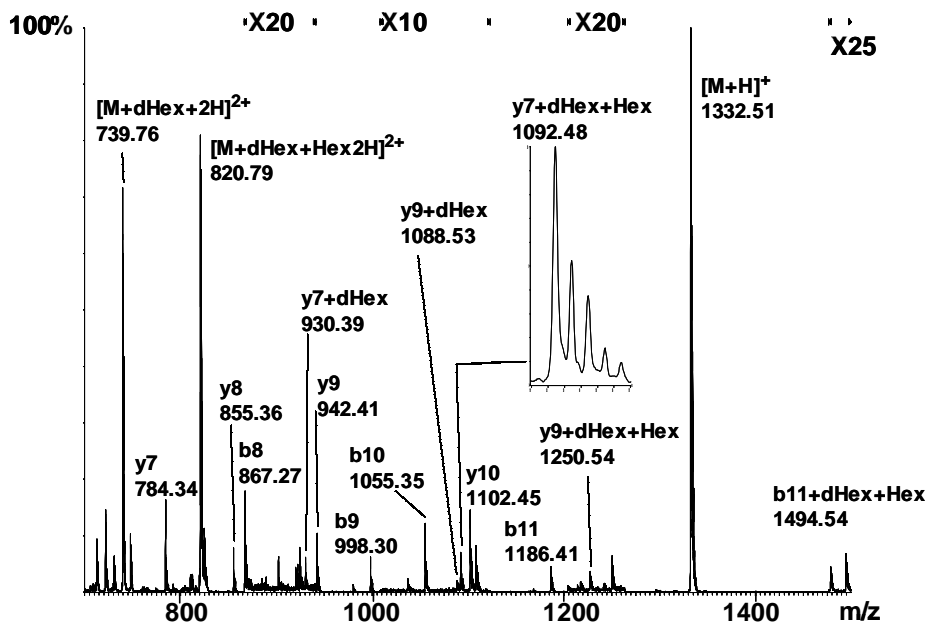
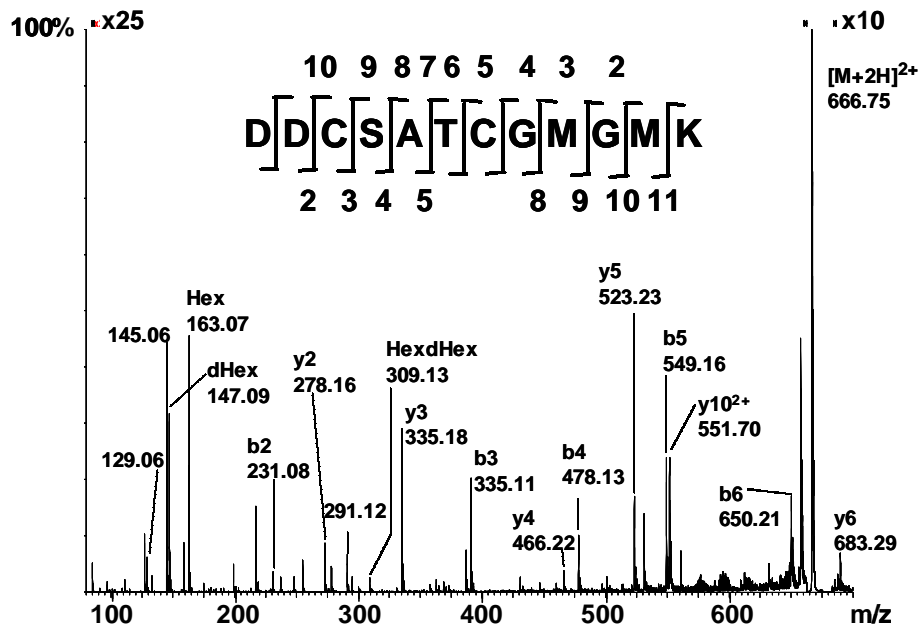


Fig.5 Hofsteenge et al.