Glycosylation Patterns of Proteins Studied by Liquid Chromatography-Mass Spectrometry and Bioinformatic Tools

Hansjörg Toll[1], Peter Berger[2*], Andreas Hofmann[3], Andreas Hildebrandt[3], Herbert Oberacher[1], Hans Peter Lenhof[3], and Christian G. Huber[1,*]

[1]Department of Chemistry, Instrumental Analysis and Bioanalysis, and [3]Center for Bioinformatics, Saarland University, 66123 Saarbrücken, Germany

[2]Institute of Biomedical Aging Research, Austrian Academy of Sciences, 6020 Innsbruck, Austria

Abstract. Due to their extensive structural heterogeneity, the elucidation of glycosylation patterns in glycoproteins such as the subunits of human chorionic gonadotropin (hCG), hCG-α and hCG-β remains one of the most challenging problems in the proteomic analysis of posttranslational modifications. In consequence, glycosylation is usually studied after decomposition of the intact proteins to the proteolytic peptide level. However, by this approach all information about the combination of the different glycopeptides in the intact protein is lost. In this study we have, therefore, attempted to combine the results of glycan identification after tryptic digestion with molecular mass measurements on the intact glycoproteins. Despite the extremely high number of possible combinations of the glycans identified in the tryptic peptides by high-performance liquid chromatography-mass spectrometry (> 1000 for hCG-α and > 10.000 for hCG-β), the mass spectra of intact hCG-α and CG-β revealed only a limited number of glycoforms present in hCG preparations from pools of pregnancy urines. Peak annotations for hCG-α were performed with the help of an algorithm that generates a database containing all possible modifications of the proteins (inclusive possible artificial modifications such as oxidation or truncation) and subsequent searches for combinations fitting the mass difference between the polypeptide backbone and the measured molecular masses. Fourteen different glycoforms of CG-α, including methionine-oxidized and N-terminally truncated forms, were readily identified. For hCG-β, however, the relatively high mass accuracy of ± 5 Da was still insufficient to unambiguously assign the possible combinations of posttranslational modifications. Finally, the mass spectrometric fingerprints of the intact molecules were shown to be very useful for the characterization of glycosylation patterns in different CG preparations.

Dagstuhl Seminar Proceedings 05471
Computational Proteomics
http://drops.dagstuhl.de/opus/volltexte/2006/543

1 Introduction

The detection and measurement of human chorionic gonadotropin (hCG) is the basis of several pregnancy testing and diagnostic procedures. Moreover, hCG has been recognized as a marker for trophoblastic and nontrophoblastic tumors. It consists of two non-covalently linked subunits, hCG-α and hCG-β whereby the α-subunit is common to all four human glycoprotein hormones (GPH) . Both subunits are heavily glycosylated: hCG-α contains two N-glycosylation sites on Asn 52 and Asn 78, hCG-β two N- glycosylation sites (Asn 13 and 30) on its core region and four O-glycosylation sites (Ser 121,127,132 and 138) (Mise & Bahl, 1980;Fujiki, Rathnam et al., 1980;Blithe & Iles, 1995;Kessler, Reddy et al., 1979;Kessler, Mise et al., 1979). Information about the glycosylation patterns is considered to be essential for the proposed use of particular forms of glycosylation as markers for cancer(Nemansky, Moy et al., 1998;Mizuochi, Nishimura et al., 1983). Our analytical strategy involved the identification of glycans present in gonadotropin first at the tryptic peptide level **(see Figure 1**) followed by characterization of the glycosylation pattern at the intact protein level **(see Figure 2**).

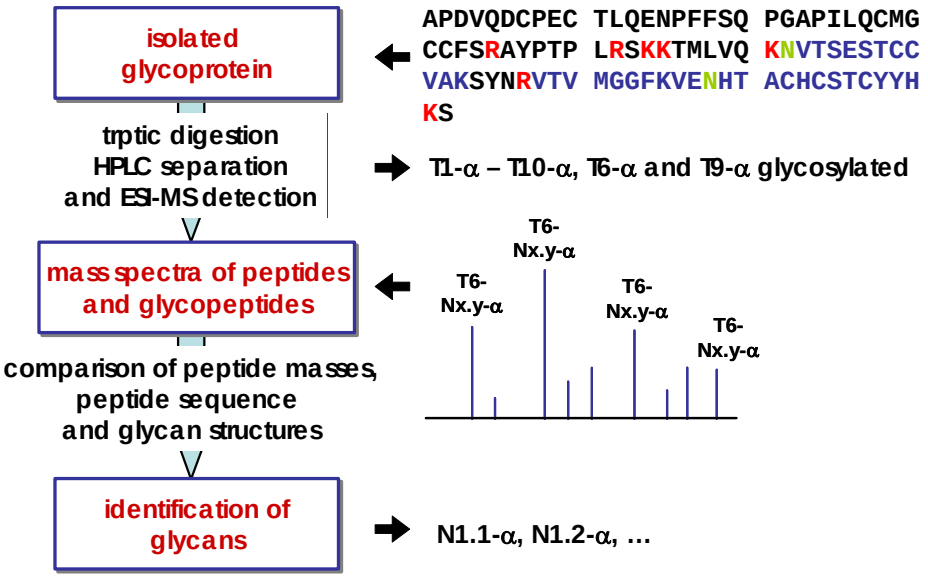

Figure 1. Analytical strategy for the identification of glycans by enzymatic digestion followed by high-performance liquid chromatography hyphenated to electrospray ionization mass spectrometry.

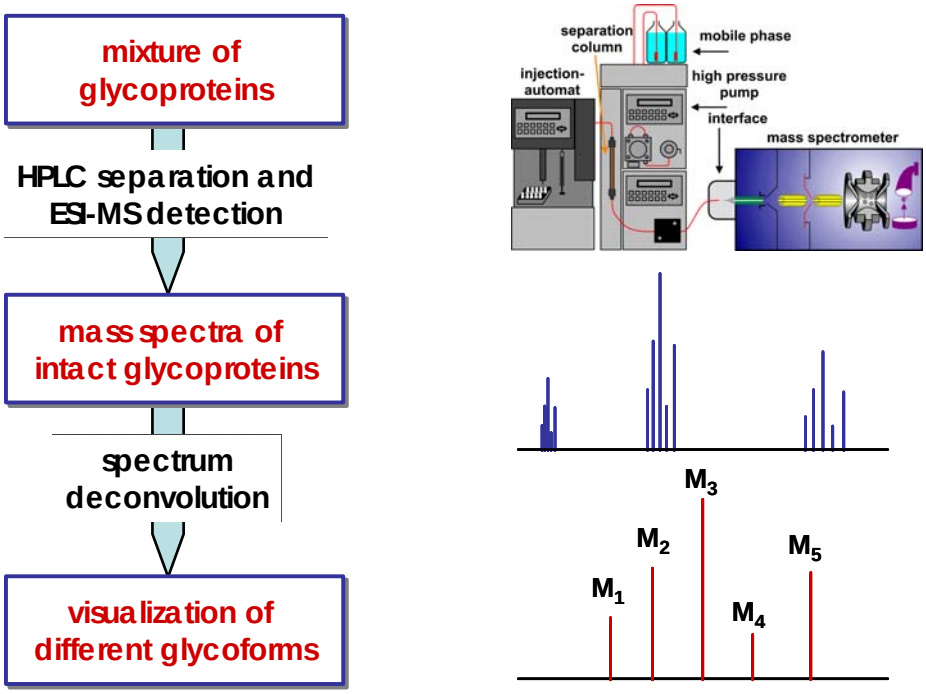

Figure 2. Analytical strategy for the characterization of glycosylation patterns by high-performance liquid chromatography hyphenated to electrospray ionization mass spectrometry.

2 Identification of Glycans on the Basis of Tryptic Glycopeptides

We followed the approach of Liu et al. (Liu & Bowers, 1997) to detect the glycopeptides of hCG-α and hCG-β in order to obtain information about the glycan compositions present in our hCG preparation. Ion-pair reversed-phase HPLC-ESI-MS analysis of the reduced and pyridylethylated tryptic peptides of purified hCG-α using 200 µm i.d. monolithic separation columns revealed the tryptic peptides of hCG-α and β. Sequence coverages of 95.7% and 96.6%, respectively were obtained for hCG-α and hCG-β. The glycan compositions associated with the peptides were assigned upon comparison of the measured masses with the theoretical masses of the peptides modified with the glycan compositions. However, investigation at the tryptic peptide level yielded no information about the assembly of different peptides and glycopeptides in the intact protein molecule. Therefore, the next step of the investigation aimed at the characterization of hCG glycosylation at the intact molecule level.

3 Characterization of hCG Glycosylation Patterns by Intact Molecular Mass Analysis

Upon multi-step purification and tuning the mass spectrometer with the 7+ charge state of lysozyme at m/z 2044, abundant signals were detected in the m/z range of 1500-3000 for hCG-α and hCG β. Deconvolution of the raw mass spectra yielded intact molecular masses in a range of 13000-14110 Da and 22000-24000 Da, which clearly identified hCG-α and hCG β. As far as we know, this is the first mass spectrometric investigation that was able to detect and resolve the heterogeneous glycoforms both of native hCG-α and hCG-β (Liu & Bowers, 1997). Considering the 31 different glycans known for hCG-α, two possible sites of N-glycosylation, oxidation, as well as carboxy-terminal truncation (see below) as additional modifications of the polypeptide backbone, combinatoric analysis tells us that there are

$$\binom{31+2-1}{2}\binom{2}{1}\binom{2}{1} = 2112$$

possible combinations to obtain the intact molecular mass of hCG-α.

A bioinformatic algorithm that generated a database containing all possible modifications of the proteins, including modifications possibly introduced during sample preparation such as oxidation or truncation, was utilized to find peak annotations for hCG-α. Combinations were searched that fitted the mass difference between the polypeptide backbone and the measured molecular masses (**Figure 3**).

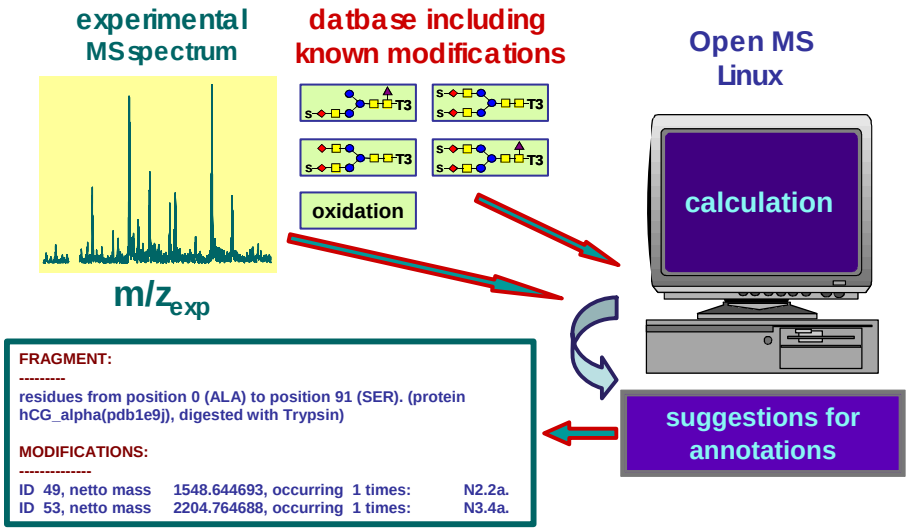

Figure 3. General scheme for the assignment of glycosylation patterns to intact glycoproteins characterized by electrospray ionization mass spectrometry.

In total, fourteen different glycoforms of hCG-α, containing biantennary, partly sialylized hybrid-type glycans, including methionine-oxidized and N-terminally truncated forms, were identified using the intact molecular mass measurements and the algorithm for peak annotation. Although mass spectra of high quality were also obtained for hCG-β, a database search mass accuracy of ± 5 Da was found insufficient to unambiguously assign the possible combinations of posttranslational modifications.

2 Conclusions

Our investigations proofed that mass spectrometry at the intact molecule level in combination with bioinformatic interpretation of databases containing possible posttranslational modifications is a useful tool for the study of glycosylated proteins. The analyses not only directly revealed the different glycoforms present in hCG preparations in pools of pregnancy urines, but also indicated the introduction of additional modifications in the course of sample workup and purification, including truncation and oxidation. To solve the problem of ambiguous peak annotation for hCG-β, the molecular masses need to be determined with greater accuracy.

3. References

Part of this work has been accepted for publication in Electrophoresis, 2006.

Blithe, D. L.Iles, R. K.," The role of glycosylation in regulating the glycoprotein hormone free alpha-subunit and free beta-subunit combination in the extraembryonic coelomic fluid of early pregnancy", Endocrinology, 136, 903-10 (1995).

Fujiki, Y., Rathnam, P.,Saxena, B. B.," Studies on the disulfide bonds in human pituitary follicle-stimulating hormone", Biochim. Biophys. Acta, 624, 428-35 (1980).

Kessler, M. J., Mise, T., Ghai, R. D.,Bahl, O. P.," Structure and location of the O-glycosidic carbohydrate units of human chorionic gonadotropin", J. Biol. Chem., 254, 7909-14 (1979).

Kessler, M. J., Reddy, M. S., Shah, R. H.,Bahl, O. P.," Structures of N-glycosidic carbohydrate units of human chorionic gonadotropin", J. Biol. Chem., 254, 7901-8 (1979).

Liu, C.Bowers, L. D.," Mass spectrometric characterization of the beta-subunit of human chorionic gonadotropin", J. Mass. Spectrom., 32, 33-42 (1997).

Mise, T.Bahl, O. P.," Assignment of disulfide bonds in the alpha subunit of human chorionic gonadotropin", J. Biol. Chem., 255, 8516-22 (1980).

Mizuochi, T., Nishimura, R., Derappe, C., Taniguchi, T., Hamamoto, T., Mochizuki, M.,Kobata, A.," Structures of the asparagine-linked sugar chains of human chorionic gonadotropin produced in choriocarcinoma. Appearance of triantennary sugar chains and unique biantennary sugar chains", J. Biol. Chem., 258, 14126-9 (1983).

Nemansky, M., Moy, E., Lyons, C. D., Yu, I.,Blithe, D. L.," Human endometrial stromal cells generate uncombined alpha-subunit from human chorionic gonadotropin, which can synergize with progesterone to induce decidualization", J. Clin. Endocrinol. Metab., 83, 575-81 (1998).