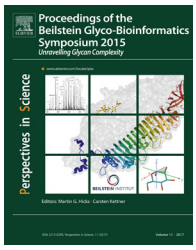
Perspectives in Science (2017) 11, 3—10

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/pisc



REVIEW

The impact of defining glycan structures☆

Johannes F.G. Vliegenthart

Bijvoet Center for Biomolecular Research, Utrecht University, Padualaan 8, 8435CH Utrecht, The Netherlands

Received 15 December 2015; accepted 5 February 2016
Available online 24 November 2016

KEYWORDS
NMR spectroscopy;
Structural reporter groups;
N- and O-type glycoproteins;
Gastrointestinal mucins;
Tamm—Horsfall glycoprotein;
Biologics;
Glycoconjugate vaccines

Summary For gaining insight in the mode of action at the molecular level of glycans in biological systems precise knowledge of the structure of the glycans is indispensable. To obtain this fundamental information well-defined starting material, optimal fractionation methods and adequate identification techniques are essential. In this review, the emphasis is on the application of high resolution ^{1}H NMR spectroscopy to the structure determination of glycans of different origin. The power of ^{1}H NMR spectroscopy is the possibility to determine in a non-destructive way all structural parameters of glycans. This is illustrated for glycans that differ in structural complexity.

Contents

E-mail address: j.f.g.vliegenthart@uu.nl

http://dx.doi.org/10.1016/j.pisc.2016.02.003

Introduction

Carbohydrates are ubiquitously found in nature as mono-, oligo- and polysaccharides or as parts of glycoconjugates (Montreuil et al., 1995, 1997; Varki et al., 2009). For gaining insight at the molecular level in the functioning and behaving of these compounds in biological systems as well as in the influence of glycans on the physicochemical properties or in pharmaceutical applications, it is essential to have reliable and precise information on the structure. Obviously, the different types of problems determine to which degree of detail the information is needed. This review is restricted to glycans in biological context. For these compounds the primary and 3-D structures are relevant. Structure analysis requires pure, preferably homogeneous, compounds as starting material. The natural heterogeneity of most glycans is a result of the biosynthetic pathways, which lead to heterogeneity in site occupancy as well as in the structure of individual glycan chains. Therefore, obtaining single molecular entities is almost impossible, despite of the formidable development in fractionation techniques. Impressive is the progress in the last decade in physical identification techniques like NMR spectroscopy and mass spectrometry, allowing the unambiguous determination of carbohydrate structures. ^{1}H and ^{13}C NMR-spectroscopy provide more detailed structural information than mass spectrometry, but require more material. The primary structure of glycans is relevant for characterising the function of free and covalently bound glycans, for the understanding of metabolic pathways in health and disease, for identification of epitopes for recognition by the immune system and more generally for the study of interaction with complementary compounds. Also for defining optimal building blocks as starting material for the synthesis of carbohydrate based vaccines it is a prerequisite to know the structure of the immune-epitope exactly. Determination of the primary structure requires the identification of the constituting monosaccharides including ring size and absolute configuration and the determination of glycosidic linkages in terms of configuration and position.

Here, the focus is mainly on results obtained by application of high resolution ^{1}H NMR spectroscopy as a tool for structural identification. For this purpose, we developed the structural-reporter-group concept comprising resonances that can be used to derive the structures (Vliegenthart et al., 1980, 1981, 1983a). Obviously, NMR can be applied to the analysis of the three-dimensional structure in solution and for the study of recognition and interaction with complementary molecules, as well.

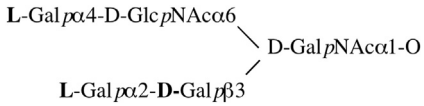


Figure 1 Structure of the *O*-glycan of conopeptide CeTx from *Conus comsors* containing **D**- and **L**-Gal residues.

Structure analysis

Absolute configuration

The determination of the absolute configuration of constituting monosaccharides is an often neglected aspect. For the determination of **L** or **D** configuration we developed a GLC procedure based on the separation of diastereomers obtained as (-)butan-2-ol derivatives of the monosaccharides (Gerwig et al., 1978, 1979). For example, the *O*-glycosylated conopeptide CeTx from *Conus comsors* has a structure containing **L**- and **D**-Gal (Hocking et al., 2013), see Fig. 1.

In particular, for glycoconjugates from lower organisms the determination of the absolute configuration of the monosaccharide components is compulsory. For mammalian systems usually the most common absolute configurations are assumed.

Glycoproteins

The starting material for structural studies of glycoproteins should be as pure as possible. This requires first that the material to be studied stems from a well-defined source in terms of organ, tissue or cell. Second, having a mixture of compounds fractionation to individual compounds should be realised as far as feasible. Third, contaminations with salts and metal ions that can disturb the recording of NMR spectra have to be removed (Leeflang and Vliegenthart, 2012). Apart from the natural (micro) heterogeneity, the main source of glycan heterogeneity is formed by the contamination with small amounts of (related) glycoproteins and have not been removed in the fractionation procedure. These contaminating compounds can be a source of wrong conclusions as to origin of the glycan structures that are present.

The glycans of glycoproteins are distinguished by the type of linkage to the protein namely N type, linked via the amide side chain of asparagine, O-type, linked via the hydroxyl group of hydroxyl amino acids, or C-type, linked to C-2 of the indole nucleus of tryptophan.

Mucin-type glycoproteins

The relevance of well-defined starting material can be illustrated for gastro-intestinal mucin. The analysis is performed