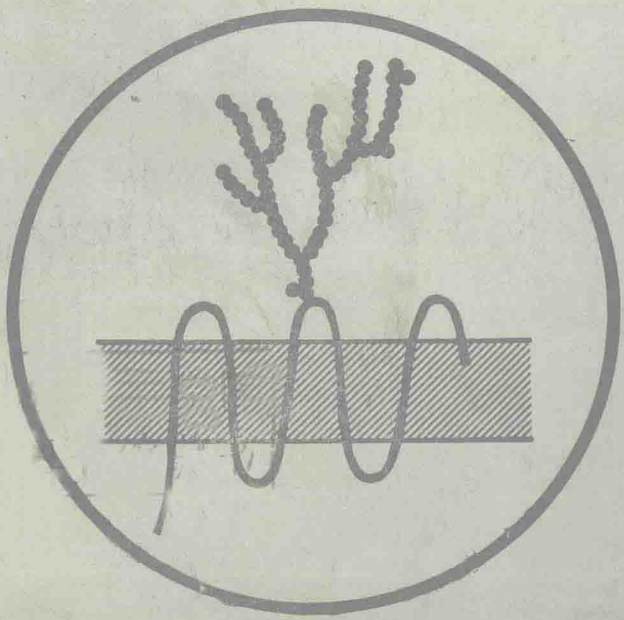
laboratory techniques

in biochemistry and molecular biology

general editors: R.H. BURDON and P.H. van KNIPPENBERG

glycoprotein and proteoglycan techniques

J.G. BEELEY



GLYCOPROTEIN AND PROTEOGLYCAN TECHNIQUES

J.G. Beeley

Department of Biochemistry,
University of Glasgow,
Glasgow G12 8QQ, U.K.

1985
ELSEVIER
AMSTERDAM · NEW YORK · OXFORD

ISBN 0-444-80652-2 (pocket edition)
ISBN 0-444-80652-0 (library edition)
ISBN 0-7204-4200-1 (series)

Published by
ELSEVIER SCIENCE PUBLISHERS B.V. (Biomedical Division)
PO BOX 211
1000 AE AMSTERDAM
THE NETHERLANDS

Sole distributors for the USA and Canada:
ELSEVIER SCIENCE PUBLISHING COMPANY, INC.
52 VANDERBILT AVENUE
NEW YORK, N.Y. 10017
USA

Library of Congress Cataloging in Publication Data

Beeley, J. G.
Glycoprotein and proteoglycan techniques.

(Laboratory techniques in biochemistry and molecular biology ; v. 16)
Bibliography: p.
Includes index.
1. Glycoproteins--Analysis. 2. Proteoglycans--Analysis. I. Title. II. Series.
QP519.L2 vol. 16 574.19'2'028 s [574.19'285] 85-10396
[QP552G59]
ISBN 0-444-80652-0
ISBN 0-444-80651-2 (pbk.)

Printed in Belgium

Glycoprotein and proteoglycan techniques

LABORATORY TECHNIQUES IN BIOCHEMISTRY AND MOLECULAR BIOLOGY

Volume 16

Edited by

R.H. BURDON - *Department of Bioscience and Biotechnology, University of Strathclyde, Glasgow*

P.H. van KNIPPENBERG - *Department of Biochemistry, University of Leiden, Leiden*

ELSEVIER
AMSTERDAM · NEW YORK · OXFORD

Contents

Introduction

There has been a remarkable expansion of interest in glycoproteins over the past two decades. From being a specialised area of structural research on the borderline between carbohydrate and protein chemistry the study of these molecules has become highly relevant to a wide range of biological phenomena. This change has come about because of developments in our knowledge of the distribution, biosynthesis, molecular organisation and functions of these molecules.

Glycoproteins (i.e. proteins containing covalently bound carbohydrate) are ubiquitous constituents of all living organisms with the possible exception of bacteria, in which they have to date only been unequivocally demonstrated in one genus (Sharon and Lis, 1972). Glycosylation is a very common modification of extracellular and integral membrane proteins of higher organisms. Covalently bound carbohydrate groups occur in glycoproteins which function as enzymes, antibodies, hormones, structural proteins, carrier proteins, mucins of epithelial secretions, membrane transport proteins and receptors.

The amount of carbohydrate present in glycoproteins can vary from less than 1% to more than 85% of the dry weight of these molecules. Proteoglycans are a class of highly glycosylated glycoproteins which are important constituents of the extracellular matrix of animal connective tissues. Some proteoglycans are closely associated with the surfaces of animal cells. The presence of glycoproteins and proteoglycans in, or attached to, the surfaces of animal cells (Hughes, 1976) has stimulated many enquiries into the role of these molecules in cellular adhesiveness, differentiation, in the control of growth and

in disease processes such as neoplasia and the infection of tissues with bacteria and viruses.

The biosynthesis of glycoproteins occurs within the internal membrane systems of cells. Glycoprotein oligosaccharide units are assembled and modified as the molecules move through successive subcellular compartments *en route* for destinations outside the cell, as membrane components of the cell surface, or as components of the membranes or contents of cellular organelles. Studies of the biosynthesis of glycoproteins are therefore not only of inherent interest in showing the mechanisms by which these molecules are assembled but can provide insight into the biological problem of how macromolecules are directed to specific destinations. For example, the examination of glycoproteins synthesised by cells obtained from individuals with genetic disorders of proteoglycan catabolism ('mucopolysaccharidoses') has provided evidence that specific signals associated with the carbohydrate units of lysosomal enzymes are responsible for directing these glycoproteins to their destination as lysosomal contents (Neufeld and Ashwell, 1980). Studies of the biosynthesis and catabolism of proteoglycans and glycoproteins are also clearly of importance in understanding diseases of connective tissue, including the most widespread of all ailments, ageing.

The role of carbohydrate in the function of glycosylated proteins has been a continuing, and at times elusive, theme in research over many years. With such wide variation in both the protein components of these molecules and the carbohydrate units attached to them it would be surprising if several different functions had not evolved. In the proteoglycans and some glycoproteins (e.g. mucous glycoproteins) the physico-chemical properties associated with the carbohydrate units (such as visco-elastic behaviour, water retention and the exclusion of macromolecules from their solvent domain) are clearly of functional significance (Muir, 1983). The high negative charge on heparan sulphate present in the basement membrane of the renal glomerulus appears to have a physiological role in the retention of macromolecules in the bloodstream (Lemkin and Farquar, 1981). A role for oligosaccharide units of glycoproteins and proteoglycans in

the protection of the peptide chains of these molecules from proteolytic cleavage has also been established. The carbohydrate units of glycoproteins can also have a pronounced effect on the folding of the peptide chain to which they are attached (Rose et al., 1984). However, there is now also sound evidence that the cells of animal tissues contain several different types of specific receptors which can recognise and respond to oligosaccharide units of glycoproteins by enhanced pinocytosis (Ashwell and Harford, 1982). The occurrence of these well-characterised receptors gives credibility to some of the many suggestions which have been made regarding the potential role of specific recognition of glycoproteins in biological processes.

The aim of this book is to describe techniques which can be used to answer some of the basic questions about glycosylated proteins. Methods are discussed for isolation, characterisation, compositional analysis, for determination of the primary structure of carbohydrate units and the nature of protein-carbohydrate linkages of glycoproteins and proteoglycans. An attempt has been made to keep in mind the diverse nature of glycosylated proteins and the many different types of problem, alluded to in preceding paragraphs, which readers may have to tackle. For example, quite different approaches are required for the isolation of mucins and membrane glycoproteins. The amount of sample available is often a limitation in the analysis of membrane glycoproteins or the products of viruses or cultured cells, and some emphasis has therefore been placed on the quantities of sample required for particular procedures and a chapter has been devoted to radioactive labelling techniques. Lectins have become important tools for the investigation of glycoproteins and these reagents are considered in a separate chapter. Emphasis has been placed on describing techniques which can be applied in most laboratories without requiring highly sophisticated equipment. However, when the best approach to a problem would be to use major equipment (e.g. high-resolution NMR) this has been indicated in terms of the nature of the sample required and the type of information obtainable, but without detailed description of instrumentation or the theoretical background.

A feature of this book is that proteoglycans and other glycoproteins are considered together. For historical reasons they have usually, although not always (Spiro, 1973), been treated separately and, on the whole, research workers have tended to concentrate on one or the other of them. The artificiality of considering proteoglycans separately from other glycosylated proteins has become obvious since it has been shown that a single peptide chain can carry carbohydrate units of both the 'proteoglycan type' and units characteristic of other glycoproteins. It is hoped that the inclusion of proteoglycans and glycoproteins will help to counteract the dichotomy which has arisen in the study of these molecules. The reader should be aware that there are also marked similarities between glycoproteins and glycolipids (Chapter 2). Some of the methods described here could quite easily be adapted to glycolipids.

Most research workers concentrate their efforts in a limited field of glycoconjugate research. The author's own interests, and limitations, may well be apparent in the selection of methods for this book, though an attempt has been made to cover a wide variety of different types of molecule.

Preparation of this book has been greatly aided and encouraged by a number of people. These include the secretarial staff of the Biochemistry Department at the University of Glasgow, who have typed the manuscript with great skill and perseverance. The illustrations have been prepared by Mr. Ian Ramsden and his staff in the Medical Illustration Unit. Dr. R. Eason kindly agreed to read and comment on Chapter 4 of the manuscript. Professor A. Kobata, Professor J. Montreuil and Professor N. Sharon generously provided copies of manuscripts prior to publication. In addition, a number of authors and publishers have given permission for the reproduction of original material subject to copyright which occurs, sometimes in slightly modified form, in several of the figures and tables in this book. Authorship of this material is indicated by the references given in the legends to figures and tables, and full details of the publications can be found in the references at the end of the book. Finally, I would like to thank the editors and publishers for their patience and encouragement.

Glycoproteins and proteoglycans

2.1. *Introduction*

This chapter is intended to provide a brief guide to the molecular structures of glycosylated proteins. More extensive discussion of the structures of these molecules can be found in reviews of proteoglycans (Kennedy, 1979; Rodén, 1980) and other glycoproteins (Marshall, 1972; Spiro, 1973; Kornfeld and Kornfeld, 1976; Montreuil, 1980; Sharon and Lis, 1982).

Glycoproteins are proteins to which carbohydrates are covalently linked through glycosidic bonds (Spiro, 1973). Proteoglycans are a subclass of glycoproteins with distinctive features of carbohydrate structure (Spiro, 1973; Sharon and Lis, 1982). In describing methodology it is, however, convenient to differentiate between proteoglycans and other glycoproteins. For this reason the term glycoprotein will, subsequently in this book, be used to describe enzymically glycosylated proteins excluding proteoglycans.

Glycoproteins and proteoglycans have carbohydrate units which vary in size from monosaccharides to polysaccharides and there may be from one to some hundreds of carbohydrate units attached to a single polypeptide chain. Subsequent sections of this chapter contain descriptions of the components from which the carbohydrate units are built up (Section 2.2), types of protein-carbohydrate linkage (Section 2.3), the structural organisation of carbohydrate units (Section 2.4), polypeptide components (Section 2.5), molecular organisation of the different types of glycosylated protein (Sections 2.6 and 2.7), structural microheterogeneity (Section 2.8) and the relation-

ship between glycoproteins, proteoglycans and other glycoconjugates (Section 2.11). Non-enzymatic glycosylation of proteins (Section 2.9) and chemically synthesised glycoproteins (Section 2.10) are also discussed.

2.2. *Carbohydrate components*

The structures of the monosaccharides which have been isolated from glycoproteins and proteoglycans are shown in Fig. 2.1 When the sugar residues are glycosidically linked they occur as six membered pyranoside rings, with the exception of L-arabinofuranoside, which has been found in plant glycoproteins. The stereoisomers of monosaccharide given in Fig. 2.1 are those which have been identified in glycoproteins and proteoglycans.

The monosaccharide residues of glycoproteins often carry substituent groups Hexosamines in glycoproteins and proteoglycans other than heparin and heparan sulphate are *N*-acetylated. Both heparin and heparan sulphate contain *N*-sulphated as well as *N* acetylated hexosamine. Most proteoglycans and several glycoproteins have *O*-sulphate substituents. A few glycoproteins (lysosomal enzymes) have been found to have mannose residues esterified with *O* phosphate (one β-*N* acetylglucosaminyl-*O*-phosphate). Sialic acids can carry a wide range of substituents. Either the *N*-acetyl or *N*-glycolyl ($-COCH_2OH$) derivative of neuraminic acid can be present and there may also be a variety of *O*-acetyl and/or *O*-glycolyl groups

The Haworth perspective formulae given in Fig. 2.1 show substituents and stereochemistry clearly but the conformations of sugar rings are not represented accurately. For consideration of the molecular shape and reactivity of the carbohydrate units of glycosylated proteins conformational formulae are preferable. The Haworth perspective formula and the conformational formula for the most stable $^1C_4$ structure in aqueous solution of β-*N*-acetyl-D-glucosamine-pyranose are illustrated in Fig 2.2. Both Haworth and conformational representations of structures will be employed elsewhere in this book