

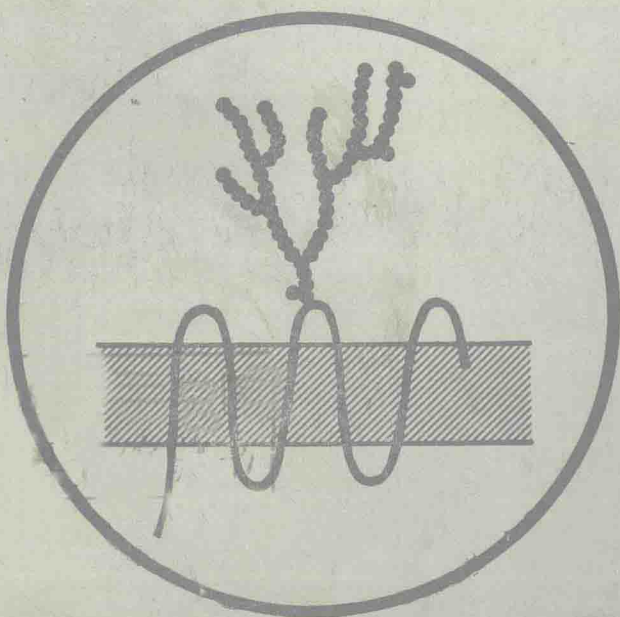
laboratory techniques

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

© 1985, Elsevier Science Publishers B.V. (Biomedical Division)

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior written permission of the publisher, Elsevier Science Publishers B.V. (Biomedical Division), P.O. Box 1527, 1000 BM Amsterdam, The Netherlands

Special regulations for readers in the USA:

This publication has been registered with the Copyright Clearance Center Inc. (CCC), Salem, Massachusetts

Information can be obtained from the CCC about conditions under which the photocopying of parts of this publication may be made in the USA. All other copyright questions, including photocopying outside the USA, should be referred to the publisher.

ISBN 0-444-8065 2 (pocket edition)

ISBN 0-444-80652-0 (library edition)

ISBN 0-720-4200-1 (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 ; no. 16)

Bibliography: p.

Includes in ex.

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*

Advisory board

P. BORST - *University of Amsterdam*

D.C. BURKE - *University of Warwick*

P.B. GARLAND - *University of Dundee*

M. KATES - *University of Ottawa*

W. SZYBALSKI - *University of Wisconsin*

H.G. WITTMAN - *Max-Planck Institut für Molekuläre Genetik, Berlin*



ELSEVIER
AMSTERDAM · NEW YORK · OXFORD

Contents

<i>Chapter 1. Introduction</i>	1
<i>Chapter 2. Glycoproteins and proteoglycans</i>	5
2.1. Introduction	5
2.2. Carbohydrate components	6
2.3. Protein-carbohydrate linkages	9
2.4. Structures of carbohydrate units	10
2.4.1. Periodic and aperiodic structures	10
2.4.2. Core structures	12
2.4.3. Peripheral residues	12
2.4.4. Simple, complex and mixed types of carbohydrate units	13
2.5. Peptide moieties	14
2.6. Secreted glycoproteins	15
2.6.1. Mucous glycoproteins	16
2.6.2. Serum glycoproteins	18
2.6.3. Structural glycoproteins	19
2.6.4. Membrane glycoproteins	20
2.7. Proteoglycans and glycosaminoglycans	21
2.8. Variability of carbohydrate structure	24
2.9. Nonenzymic glycosylation	26
2.10. Semi-synthetic glycoproteins (neoglycoproteins)	27
2.11. Relationships between glycoconjugates	28
<i>Chapter 3. Isolation and fractionation</i>	29
3.1. Introduction	29
3.2. Choice of starting material	31
3.3. Extraction and solubilisation	32

3.3.1. Mucous glycoproteins	32
3.3.1.1. Centrifugation and homogenisation	33
3.3.1.2. Sonication	33
3.3.1.3. Extraction with salts and urea	34
3.3.1.4. Phenol extraction	34
3.3.1.5. Reducing agents	35
3.3.1.6. Proteolysis	36
3.3.2. Membrane glycoproteins	36
3.3.2.1. Detergents	38
3.3.2.2. Factors affecting detergent extraction of membrane glycoproteins	39
3.3.2.3. Chaotropes and protein denaturants	42
3.3.2.4. Organic solvents	42
3.3.2.5. Proteolysis	43
3.3.3. Proteoglycans	44
3.3.3.1. Extraction conditions	44
3.3.3.2. Effects of shearing forces and sonication	44
3.3.3.3. Extractants	45
3.3.3.4. Reducing agents	48
3.3.3.5. Proteases	48
3.4. Isolation and fractionation	48
3.4.1. Assays for glycoproteins and proteoglycans	49
3.4.2. Fractionation based on solubility	49
3.4.3. Fractionation based on size and shape	51
3.4.4. Fractionation based on charge	52
3.4.5. Fractionation based on density differences	53
3.4.6. Affinity chromatography	54
3.4.7. Examples of glycoprotein isolation	55
<i>Membrane glycoproteins</i>	
3.4.7.1. Glycophorin — chloroform-methanol extraction	55
3.4.7.2. Glycophorin — LIS-phenol extraction	56
3.4.7.3. Band 3 extraction with Triton X-100	56
3.4.7.4. Purification of Thy-1 antigen by affinity chromatography	57
<i>Mucins</i>	
3.4.7.5. Ovine submaxillary mucin	58
3.4.7.6. Glycoprotein from gastric mucus	59
<i>Proteoglycans</i>	
3.4.7.7. Proteoglycan 'subunit' from bovine nasal cartilage	60
3.4.7.8. Isolation of low buoyant density dermatan sulphate proteoglycan synthesised by cultured cells	60
3.4.7.9. Purification of keratan sulphate proteoglycan from monkey cornea	61
 <i>Chapter 4. Physico-chemical characterisation</i>	 63
4.1. Introduction	63
4.2. Problems in characterising glycoproteins and proteoglycans	65

4.3. Molecular weight determination	67
4.3.1. Gel filtration	67
4.3.2. Electrophoretic methods	73
4.3.3. Sedimentation equilibrium ultracentrifugation	79
4.3.4. Other methods	87
4.4. Characterisation of charge	88
4.5. Shape, interactions and flexibility	94

Chapter 5. Analysis of constituents 100

5.1. Introduction	100
5.2. Analytical methods	101
5.3. Hydrolysis and methanolysis	103
5.4. Amino acid composition	105
5.5. Hexosamines and hexosaminotols	107
5.6. Uronic acids	113
5.7. Sialic acids	117
5.7.1. Identification of sialic acids	119
5.7.2. Purification of sialic acid	120
5.7.3. Choice of sialic acid assay	121
5.7.4. Colourimetric assays for sialic acid	122
5.7.5. Enzymic assays for sialic acid	125
5.7.6. Analysis of sialic acid by GLC	125
5.8. Neutral sugars	127
5.8.1. Colourimetric assays for neutral sugars	129
5.8.2. Hydrolysis procedure for neutral sugars	133
5.8.3. Ion-exchange chromatography of borate complexes	134
5.8.4. Gas-liquid chromatography	137
5.8.4.1. Estimation of neutral sugars as their alditol acetates	138
5.8.4.2. Estimation of methyl glycosides	140
5.9. Microdetermination of ^3H -labelled sugar alcohols	144
5.10. Isotope dilution methods	147
5.11. Acetyl and sulphate	147
5.12. Reducing sugars	151

Chapter 6. Structural analysis 153

6.1. Introduction	153
6.2. Strategy for structural analysis	153
6.3. Is carbohydrate covalently linked to protein?	157
6.4. Release of carbohydrate units	159
6.4.1. The isolation of glycopeptides	159

6.4.2. Non-specific cleavage	160
6.4.3. Specific cleavage	162
6.5. Release of carbohydrate units by cleavage of protein-carbohydrate linkages	165
6.5.1. Alkaline cleavage of protein-carbohydrate linkages	167
6.5.1.1. Alkali-sensitive <i>O</i> -linked carbohydrate units	167
6.5.1.2. <i>N</i> -linked carbohydrate units	170
6.5.2. Cleavage of protein-carbohydrate linkages by hydrazinolysis	171
6.5.3. Trifluoroacetolysis	173
6.5.4. Enzymic cleavage of protein-carbohydrate linkages	175
6.6. Release of carbohydrate units with endoglycosidases	177
6.6.1. Endo- β - <i>N</i> -acetylglucosaminidases — specificity	179
6.6.2. Endo- β - <i>N</i> -acetylhexosaminidases — assays	181
6.6.3. Endo- β - <i>N</i> -acetylhexosaminidases — applications	182
6.6.4. Endo- α - <i>N</i> -acetylgalactosaminidase	184
6.6.5. Endo- β -galactosidases	184
6.6.6. Problems encountered in the use of endoglycosidases	186
6.7. Fractionation of glycopeptides and oligosaccharides	187
6.7.1. Glycoproteins obtained by non-selective cleavage of glycoproteins and proteoglycans	187
6.7.2. Glycopeptides obtained by selective cleavage of glycoproteins	191
6.7.3. Glycopeptides and glycosaminoglycans from cells or tissues	193
6.7.4. Fractionation of oligosaccharides obtained by chemical or enzymic cleavage	198
6.7.4.1. Oligosaccharide separation — paper electrophoresis and ion-exchange chromatography	200
6.7.4.2. Gel filtration (or gel permeation)	202
6.7.4.3. Paper and thin-layer chromatography	203
6.7.4.4. Adsorption chromatography on charcoal	205
6.7.4.5. High-performance liquid chromatography	206
6.7.4.6. Gas-liquid chromatography	209
6.7.4.7. Lectin affinity chromatography	209
6.8. Determination of the nature of protein-carbohydrate linkages in glycoproteins and proteoglycans	210
6.8.1. Glycosylamine linkage to asparagine	211
6.8.2. <i>O</i> -Glycosidic linkages to serine or threonine	214
6.8.2.1. <i>N</i> -Acetylgalactosaminyl linkage to serine or threonine	215
6.8.2.2. Galactosyl-serine linkage	219
6.8.2.3. Mannosyl-threonine or -serine linkages	220
6.8.2.4. Xylosyl-serine or -threonine linkages	221
6.8.3. Galactosyl linkages to hydroxylysine	222
6.8.4. Arabinosyl and galactosyl linkages to hydroxyproline	223
6.9. Strategies for the structural analysis of carbohydrate moieties	225
6.9.1. Methylation	228
6.9.2. Mass spectrometry	239
6.9.3. Nuclear magnetic resonance	245
6.9.4. Methods for partial degradation	253

6.9.4.1. Enzymic digestion	253
6.9.4.2. Partial acid hydrolysis	272
6.9.4.3. Acetolysis	277
6.9.4.4. Periodate oxidation and Smith degradation	279
6.9.4.5. Deamination with nitrous acid	288
6.9.4.6. Oxidation with CrO_3	292
6.10. Location of substituents	293
6.11. Deglycosylation of glycoproteins and proteoglycans	296

Chapter 7. Lectin techniques 301

7.1. Introduction	301
7.1.1. Isolation of lectins	304
7.1.2. Molecular properties	305
7.2. Lectin binding	314
7.2.1. Introduction	314
7.2.2. Radioactive labelling of lectins	318
7.2.3. Measurement of lectin binding	321
7.3. Agglutination methods	327
7.3.1. Introduction	327
7.3.2. Quantitation of agglutination	330
7.4. Lectin affinity chromatography	333
7.4.1. Setting up an affinity chromatography system	334
7.4.2. Purification of soluble glycoproteins	341
7.4.3. Fractionation of glycoprotein species with different carbohydrate groups	342
7.4.4. Affinity chromatography of membrane glycoproteins	344
7.4.5. Affinity chromatography of glycopeptides	345
7.5. Lectin staining methods	350
7.6. Lectin immunoprecipitation	357
7.7. Lectin precipitation methods	361
7.7.1. Introduction	361
7.7.2. Precipitation of glycoproteins or polysaccharides from solution	361
7.7.3. Lectin precipitation in gels (lectin immunodiffusion)	362
7.7.4. Electrophoresis of glycoproteins into lectin-containing gels	362
7.7.5. Electrophoresis through gels containing insolubilised lectin	363
7.8. Covalent linking of lectins to receptors	364

Chapter 8. Radioactive labelling techniques 365

8.1. Introduction	365
8.2. Biosynthetic labelling	367

8.2.1. Labelling intact cells	371
8.2.2. Determination of the extent of interconversion of label	377
8.2.3. Labelling for structural characterisation	378
8.2.4. Labelling to examine biosynthetic pathways	378
8.2.5. Measuring rates of synthesis	391
8.2.6. In vitro (cell-free) labelling of glycoproteins	395
8.2.7. In vitro (cell-free) labelling of proteoglycans	406
8.2.8. Inhibitors of glycoprotein and proteoglycan biosynthesis	410
8.3. Labelling terminal sialic acid, galactose and <i>N</i> -acetylgalactosamine	418
8.3.1. Sialic acid labelling with NaIO_4 and NaB^3H_4	418
8.3.2. Galactose and <i>N</i> -acetylgalactosamine labelling	422
 <i>Appendix</i>	 426
 <i>References</i>	 433
 <i>Subject index</i>	 456

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 substituents. 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 (e.g. β -*N*-acetylglucosaminyl-*O*-phosphate). Sialic acids can carry a wide range of substituents. Either the *N*-acetyl or *N*-glycolyl ($-\text{COCH}_2\text{OH}$) 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 $^1\text{C}_4$ 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.