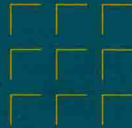


P. Crocker



Mammalian Carbohydrate Recognition Systems

Results
and Problems
in Cell
Differentiation



Springer

Paul R. Crocker (Ed.)

Mammalian Carbohydrate Recognition Systems

With 52 Figures and 10 Tables



Springer

July 2000

Dr. PAUL R. CROCKER
The Wellcome Trust Building
Dept. of Biochemistry
University of Dundee
Dundee DD1 4HN
Scotland

ISBN 3-540-67335-0 Springer-Verlag Berlin Heidelberg New York

Library of Congress Cataloging-in-Publication Data

Mammalian carbohydrate recognition systems / Paul R. Crocker (ed.).
p. cm. – (Results and problems in cell differentiation ; 33)
Includes bibliographical references (p.).
ISBN 3540673350

1. Lectins. I. Crocker, Paul R., 1959– II. Series.

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilm or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer-Verlag. Violations are liable for prosecution under the German Copyright Law.

Springer-Verlag Berlin Heidelberg New York
a member of Bertelsmann Springer Science+Business Media GmbH

© Springer-Verlag Berlin Heidelberg 2001
Printed in Germany

The use of general descriptive names, registered names, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

Cover design: Meta Design, Berlin
Typesetting: Best-set Typesetter Ltd., Hong Kong
Printed on acid-free paper SPIN 10718859 39/313Oas – 5 4 3 2 1 0

Results and Problems in Cell Differentiation

Series Editors:

W. Hennig, L. Nover, U. Scheer

29059890

33

Mammalian Carbohydrate Recognition Systems

With 52 Figures and 10 Tables

ger

Springer

Berlin

Heidelberg

New York

Barcelona

Hong Kong

London

Milan

Paris

Singapore

Tokyo



Springer

Results and Problems in Cell Differentiation

33

Series Editors:
W. Hennig, I. Nover, U. Scheer

Springer

Berlin

Heidelberg

New York

Barcelona

Hong Kong

London

Milan

Paris

Singapore

Tokyo

Preface

Over the last decade, the number of known mammalian carbohydrate binding proteins (CBPs) has increased considerably. With the recent completion of the human genome, this number is likely to increase further still. Fortunately, most of the known CBPs can be placed into a small number of distinct families based on the primary amino acid sequence. Currently, there are five recognised categories, namely C-type lectins, Galectins, I-type lectins, P-type lectins and legume-related lectins. In addition, there are 'stand-alone' CBPs such as the hyaluronic acid-binding protein, CD44, although here, too, recent evidence suggests that this CBP may be part of a larger gene family. Each class of CBP has a characteristic protein fold required for carbohydrate recognition and, in most cases, the three-dimensional structure of at least one member has been determined in the presence of a bound oligosaccharide ligand. Given the enormous complexity of carbohydrates themselves and the huge array of glycoproteins and glycolipids that present them, it is not surprising that CBPs are involved in remarkably diverse functions. This is exemplified by the Galectins which mediate specific functions in the nucleus, cytoplasm, cell surface and in the extracellular milieu. Other CBPs are found in discrete cellular and extracellular compartments in a way that is intrinsically linked to their known functions. The 12 chapters of this book are focused on structure-function relationships of these CBP families, including CD44, and are written by leading experts in the field. The order of chapters is based on the cellular localisation of CBPs, beginning with a review on the endoplasmic reticulum chaperones, calreticulin and calnexin, and ending with the Collectins, a group of remarkable molecules that function in host defence to pathogens. The intervening chapters cover Galectins and a range of membrane proteins (ERGIC-53, mannose-6-phosphate receptors, CD44, the mannose/GalNAc-SO₄ receptor, Siglecs and Selectins). Collectively, they discuss the molecular basis for carbohydrate recognition and how this triggers various biological responses, ranging from control of glycoprotein folding to initiation of leukocyte recruitment in inflammation. The new millennium offers exciting opportunities not only to identify and molecularly characterise new CBPs but also to deepen our understanding of how carbohydrate recognition by this diverse group of proteins leads to 'glycosignalling' and control of cellular behaviour. I would like to take this opportunity to thank all the authors for their excellent contributions.

Paul R. Crocker, Dundee July 2000

Addresses of Senior Authors

Dr Klaus Ley
University of Virginia
Health Science Center
School of Medicine
Dept of Biomedical Engineering
BOX 377
Charlottesville, VA 22908
USA

Dr Hakon Leffler
Department of Medical Microbiology
Solvegatan 23
S 22362 Lund
Sweden

Dr Maureen Taylor
Glycobiology Institute
Biochemistry Department
University of Oxford
South Parks Road,
Oxford OX1 3QU
UK

Dr Nancy M. Dahms
Department of Biochemistry
Medical College of Wisconsin
Milwaukee
Wisconsin 53226
USA

Professor
Dr Ten Feizi
The Glycosciences Laboratory
Northwick Park Hospital, Watford Road, Harrow, Middx.
HA1 3UJ
UK

Professor
Dr Ken Reid
MRC Immunochemistry Unit
Department of Biochemistry
University of Oxford
South Parks Road
Oxford OX1 3QU
UK

Dr Soerge Kelm
NW2, University of Bremen
Postfach 33 04 40
D-28334 Bremen
Germany

Dr Jürgen Bajorath
Senior Director
Computer-Aided Drug Discovery
New Chemical Entities, Inc.
18804 North Creek Pkwy. South
Bothell, WA 98011
USA
Tel.: 001-425-424-7297
Fax: 001-425-424-7299
jbajorath@nce-mail.com

Professor
Dr Jacques U. Baenziger
Washington University Medical School
Department of Pathology
660 S. Euclid Ave.
St. Louis, MO 63110
USA

Dr John J.M. Bergeron
Department of Anatomy and Cell Biology
McGill University
Montreal, Quebec H3A 2B2
Canada

Dr Andy May
Department of Structural Biology
Fairchild Building, Room D-139
Stanford University School of Medicine
Stanford, CA 94305-5126
USA

Dr A-C. Roche
Glycobiologie
Centre de Biophysique Moleculaire
Bat. B, rue Charles Sadron
45071 Orleans, cedex 2
France

Contents

Lectins of the ER Quality Control Machinery

C.A. Jakob, E. Chevet, D.Y. Thomas, J.J.M. Bergeron

1 Introduction	1
2 Structural Aspects	1
3 Calnexin and Calreticulin in Quality Control of Glycoprotein Folding	6
4 Calnexin and Calreticulin in Glycoprotein Degradation	9
5 Regulation of Calnexin/Calreticulin-Substrate Interaction	12
5.1 Modifications of the Luminal Domain	12
5.2 Alterations of the Cytosolic Tail	13
6 Conclusions	13
References	14

MR60/ERGIC-53, a Mannose-Specific Shuttling Intracellular Membrane Lectin

Annie-Claude Roche, Michel Monsigny

1 Discovery	19
2 Structural Features	21
3 Intracellular Mannose-Specific Animal Lectins Are Homologous to Leguminous Plant Lectins	22
4 Oligomerization and Sugar Binding Activity	23
5 Cytological Features	26
6 Deciphering the Role of MR60/ERGIC-53: Looking for Highly Specific Oligosaccharide and Natural Glycoprotein Ligands	29
7 Concluding Remarks	34
References	35

The Cation-Dependent Mannose 6-Phosphate Receptor

Jung-Ja P. Kim, Nancy M. Dahms

1 Introduction	39
2 Intracellular Trafficking of the MPRs and Lysosomal Enzymes	40
2.1 Generation of the Mannose 6-Phosphate Recognition Marker	40

2.2 Subcellular Distribution of the MPRs	41
2.3 Targeting Signals in the Cytoplasmic Region of the MPRs	41
3 Primary Structure and Biosynthesis of the CD-MPR	42
3.1 Primary Structure	42
3.2 Genomic Structure	43
3.3 Oligomeric Structure	44
3.4 Co- and Post-Translational Modifications	44
3.4.1 Acylation	44
3.4.2 Phosphorylation	44
3.4.3 Glycosylation	45
4 Carbohydrate Recognition by the CD-MPR	45
4.1 Lysosomal Enzyme Recognition	45
4.2 Expression of Mutant Forms of the CD-MPR	47
4.3 Crystal Structure of the CD-MPR in the Presence of Bound Man-6-P	48
4.3.1 Polypeptide Fold	48
4.3.2 Structural Similarity to Biotin-Binding Proteins	49
4.3.3 Dimeric Structure	50
4.3.4 Carbohydrate Binding Pocket	50
4.3.5 Comparison of the CRDs of the CD-MPR and CI-MPR	53
5 Concluding Remarks	53
References	54

Galectins Structure and Function – A Synopsis

Hakon Leffler

1 Introduction	57
1.1 Discovery of Galectins – Past and Present	59
2 Structure, Specificity, and Endogenous Ligands	59
2.1 The Carbohydrate Recognition Domain (CRD) and Carbohydrate Binding Site	59
2.2 Domain Organization, Oligomerization, and Valency	61
2.3 Endogenous Galectin Glycoconjugate Ligands	62
3 Genes, Expression, and Targeting	63
3.1 Galectin Genes	63
3.2 Galectin Distribution in Cells and Tissues	64
3.3 Synthesis, Intracellular Targeting and Secretion	64
4 Functional Effects	65
4.1 Cell Adhesion	66
4.2 Galectin Induced Signaling	66
4.3 Galectins in Apoptosis	68
4.4 Galectins and Galectin Inhibitors In Vivo	68
4.5 Nuclear Functions	69
4.6 Other Galectin Effects	69
4.7 Galectin Null-Mutant Mice	70

5	Biological Roles and Biomedical Use	70
5.1	Immunity and Inflammation	70
5.2	Host-Pathogen Interaction	71
5.3	Cancer	71
5.4	Galectin Serology	72
5.5	Tissue Organization and Repair	73
5.6	Galectins in the Nervous System	73
6	Summary and Conclusions	73
	References	74

Structure and Function of CD44: Characteristic Molecular Features and Analysis of the Hyaluronan Binding Site

Jürgen Bajorath

1	Synopsis	85
2	Cell Adhesion Proteins	85
2.1	Representative Families and Characteristic Features	85
3	Molecular Structure of CD44	86
3.1	Cloning of CD44 and Domain Organization	86
3.2	Genomic Structure and Isoforms	88
3.3	Glycosylation	88
4	Biological Functions and Ligands of CD44	88
4.1	Functional Diversity	88
4.2	Hyaluronan and Other Ligands	89
4.3	Regulation of Hyaluronan Binding	90
4.4	Hyaluronan Binding Domain	90
5	The Link Protein Module	91
5.1	Three-Dimensional Structure of TSG-6	91
5.2	Molecular Model of the Link Module of CD44	91
6	Analysis of the Hyaluronan Binding Site in CD44	93
6.1	Mutagenesis Strategy and Experimental Approach	93
6.2	Classification of Targeted Residues	95
6.3	Mapping and Characterization of the Binding Site	96
6.4	Opportunities and Limitations	97
7	Comparison of Carbohydrate Binding Sites	99
7.1	Link Proteins and C-Type Lectins	99
8	Conclusions	100
	References	101

Structure and Function of the Macrophage Mannose Receptor

Maureen E. Taylor

1	Functions and Biological Ligands of the Mannose Receptor	105
1.1	Identification and Localization of the Mannose Receptor	105
1.2	Roles of the Mannose Receptor in the Immune Response	105
1.3	Clearance of Soluble Endogenous Ligands by the Mannose Receptor	106

2	Structure of the Mannose Receptor	107
2.1	Primary Structure	107
2.2	Features of Individual Domains	107
2.2.1	The Cytoplasmic Tail	107
2.2.2	The N-Terminal Cysteine-Rich Domain and the Fibronectin Type II Repeat	108
2.2.3	The C-Type Carbohydrate Recognition Domains	109
3	Mechanisms of Carbohydrate Binding by the Mannose Receptor	110
3.1	Roles of Individual Domains	110
3.2	Molecular Mechanism of Monosaccharide Binding to the Fourth Carbohydrate-Recognition Domain	111
3.2.1	Interaction of Ca ²⁺ with CRD-4	111
3.2.2	Involvement of a Stacking Interaction in Sugar Binding to CRD-4	112
3.2.3	Determinants of Specificity and Orientation of Monosaccharides Bound to CRD-4	114
3.3	Spatial Arrangement of Domains	114
4	The Mannose Receptor Family	116
4.1	Members of the Family	116
4.2	Ligand Binding by Members of the Mannose Receptor Family	117
4.3	Evolution of the Mannose Receptor Family	117
5	Conclusions	118
	References	119

The Man/GalNAc-4-SO₄-Receptor has Multiple Specificities and Functions

Alison Woodworth, Jacques U. Baenziger

1	Introduction	123
2	Oligosaccharides Terminating with the Sequence SO ₄ -4GalNAcβ1,4GlcNAcβ1,2Manα (S4GGnM) Are Found on the Pituitary Glycoprotein Hormones of All Vertebrates	124
3	Terminal β1,4-Linked GalNAc-4-SO ₄ Determines the Circulatory Half-life of LH	125
4	The GalNAc-4-SO ₄ -Receptor: A Receptor with Multiple Carbohydrate Specificities and Functions	126
5	Functional Significance of the Man/GalNAc-4-SO ₄ -Receptor	131
6	Future Directions	134
	References	135

Sialoadhesin Structure

Andrew P. May, E. Yvonne Jones

1	Introduction	139
---	--------------------	-----

2 Sialoadhesin Carbohydrate-Binding Domain Structure and Function	141
3 Comparison with Other Sialic Acid-Binding Proteins	146
4 Sialic Acid Mediated Cell Adhesion	147
5 Conclusion	148
References	149

Ligands for Siglecs

Soerge Kelm

1 Introduction	153
1.1 Sialic Acids in Cell Recognition	153
1.2 Selectins	154
1.3 Siglecs	155
2 Structures of Siglecs	155
2.1 Sialic Acid-Binding Domain	156
2.1 Glycosylation	157
3 Carbohydrate Recognition	157
3.1 Methodology	157
3.1.1 Cell Surface Resialylation and Neoglycoconjugates	160
3.1.2 Cell Binding Assays	160
3.1.3 Fc-Chimeras	161
3.2 Sialic Acid Interactions with Siglecs	161
3.2.1 Functional Groups of Sialic Acids and Amino Acids	162
3.3 Glycan Interactions with Siglecs	165
3.3.1 Linkage Specificity	165
3.4 Ligands	169
3.4.1 Cell Surface Glycoproteins	169
3.4.2 Extracellular Glycoproteins	169
3.4.3 Glycolipids	169
3.5 Potential Regulation by <i>cis</i> -Interactions or Soluble Competitors	170
4 Perspectives	172
References	173

Functions of Selectins

Klaus Ley

1 Selectin Structure and Expression	177
1.1 Selectin Structure	177
1.2 Expression of E-selectin	177
1.3 Expression of P-selectin	178
1.4 Expression of L-selectin	178
2 Functions Common to All Selectins	178
2.1 Adhesion Under Flow	178
2.2 Leukocyte Adhesion Cascade	179

2.3	Selectin Bond Mechanics	181
3	Differential Functions of Selectins	181
3.1	P-selectin	182
3.2	E-selectin	182
3.3	L-selectin	183
3.4	Homing of Bone Marrow Stem Cells	183
4	Signaling through Selectins	184
5	Selectin Ligands	184
5.1	PSGL-1	185
5.2	Other Selectin Ligands	185
6	Phenotype of Selectin-Deficient Mice	186
6.1	E- and P-Selectin Double Deficient Mice	186
6.2	Triple Selectin Deficient Mice	186
6.3	L-Selectin-Deficient Mice	187
6.4	P-Selectin Deficient Mice	187
6.5	E-Selectin-Deficient Mice	187
6.6	L- and E- and L- and P-Selectin Double Deficient Mice	188
7	Selectin Ligand Deficiencies	188
7.1	Selectin Ligand Deficiencies in Mice	188
7.2	Selectin Ligand Deficiency in Humans	188
8	Selectins and Disease	189
8.1	Ischemia and Reperfusion	189
8.2	Cancer Metastasis	189
8.3	Autoimmune Diseases	190
8.4	Atherosclerosis	190
9	Summary	191
	References	191

Carbohydrate Ligands for the Leukocyte-Endothelium Adhesion Molecules, Selectins

Ten Feizi

1	Introduction	201
2	Carbohydrate Ligands for E-Selectin	203
2.1	The Initial Evaluations of the Le ^x and Le ^a Systems as Ligands	203
2.2	The Search for Epithelial and Myeloid Cell Ligands	207
2.3	Conclusions and Questions Regarding the Carbohydrate Sequences Recognized by E-Selectin	209
3	Carbohydrate Ligands for L-Selectin	210
3.1	Initial Explorations of Carbohydrate Sequences Recognized by L-Selectin	210
3.2	Novel Sulfated Sequences Detected on the Endothelial Glycoprotein GlyCAM-1	210
3.3	Chemically Synthesized Sulfated Forms of the Le ^x Pentasaccharide, and Their Interactions with L-Selectin	211

3.4 Clues to the Existence of Novel Biosynthetic Pathways for Selectin Ligands	213
3.5 The Second Class of Sulfated L-Selectin Ligand	214
3.6 Possible Co-Operativity Between the Long and Short Ligands for L-Selectin	214
4 Oligosaccharide Ligands for P-Selectin	215
4.1 P-selectin Interactions with Defined Saccharide Sequences	215
4.2 PSGL-1, the Counter-Receptor Displaying Two Classes of Ligand for P-Selectin	216
5 Perspectives	217
References	218

Structures and Functions of Mammalian Collectins

Uday Kishore, Kenneth B.M. Reid

1 Introduction	225
2 Mannose-Binding Lectin (MBL)	226
2.1 Molecular Structure and Assembly of MBL	227
2.2 Biological Functions of MBL	227
2.3 Interaction of MBL with Micro-organisms	228
2.4 Gene Organisation and Genetics of MBL	229
2.5 Crystal Structure of Trimeric CRDs of MBL	229
3 Surfactant Protein A (SP-A)	232
3.1 SP-A Suprastructure and Assembly	232
3.2 SP-A Gene and Genomic Organisation	233
3.3 SP-A-Carbohydrate Interaction	234
3.4 SP-A-Phospholipid Interactions	234
3.5 SP-A-Type II Cell Interaction	235
3.6 Interaction of SP-A with Phagocytes	235
3.7 Interaction of SP-A with Pathogens and Allergens	236
4 Surfactant Protein D (SP-D)	238
4.1 Molecular Structure and Assembly of SP-D	238
4.2 Interaction of SP-D with Carbohydrate and Lipid Ligands	239
4.3 Interaction of SP-D with Pathogens and Allergens	239
4.4 SP-D Gene Organisation and Genetics	240
4.5 SP-D Crystal Structure	240
5 Cell Surface Receptors for Collectins	241
6 SP-A and SP-D Gene Knock-out Mice	242
7 SP-A and SP-D in Human Diseases	243
8 Bovine Collectins: Conglutinin (BC) and Collectin-43 (CL-43)	243
References	244
Index	249

Lectins of the ER Quality Control Machinery

C. A. Jakob, E. Chevet^{1,2}, D. Y. Thomas, J. J. M. Bergeron¹

Abbreviations. CNX: calnexin CRT: calreticulin CPY: carboxypeptidase Y CK2: casein kinase 2 ER: endoplasmic reticulum UGGT:UDP-glucose glycoprotein glucosyltransferase

1 Introduction

Calnexin was first identified as a type I transmembrane phosphoprotein located in the endoplasmic reticulum (ER; Wada et al. 1991). In subsequent work, this protein has been shown to belong to a new family of chaperones or proteins with lectin-like properties with orthologs present in all eukaryotes so far studied (Wada et al. 1991; Hebert et al. 1995). Calnexin and calreticulin, an ER luminal homologue, share extensive sequence similarity and both bind calcium.

Calreticulin was originally isolated as a calcium storage protein from the sarcoplasmic reticulum and the ER. Only later was it discovered that calreticulin also binds oligosaccharides and hence belongs to the calnexin-like lectins. Two calnexin homologues, calmegin (Watanabe et al. 1994) and/or calnexin-t (Ohsako et al. 1994), have also been described in testis and display the same 'lectin' properties as calnexin in their luminal domains and also cytoplasmic C-terminal regions. Regions of these proteins share amino acid sequence identity ranging from 42–78% (Wada et al. 1991).

2 Structural Aspects

Common features in the luminal domain of all calnexin and calreticulin-like proteins are the high affinity Ca^{2+} -binding site and the proline-rich repeat motifs 1 (I-DPD/EA-KPEDWDD/E) and 2 (G-W-P-I-NP-Y; Fig. 1A). The proline-rich motif is also known as the P-domain in calreticulin (Trombetta and Helenius 1998). Calnexin, calreticulin and calmegin all contain multiple

¹ Department of Anatomy and Cell Biology, McGill University, Montreal, QC, H3A 2B2, Canada

² Genetics Group, Biotechnology Research Institute, National Research Council of Canada, 6100 Royalmount Avenue, Montreal, QC, H4P 2R2, Canada