

Topics in Current Chemistry 367

Rita Gerardy-Schahn
Philippe Delannoy
Mark von Itzstein *Editors*

SialoGlyco Chemistry and Biology II

Tools and Techniques to Identify and
Capture Sialoglycans

 Springer

367

Topics in Current Chemistry

Editorial Board:

H. Bayley, Oxford, UK
K.N. Houk, Los Angeles, CA, USA
G. Hughes, CA, USA
C.A. Hunter, Sheffield, UK
K. Ishihara, Chikusa, Japan
M.J. Krische, Austin, TX, USA
J.-M. Lehn, Strasbourg Cedex, France
R. Luque, Córdoba, Spain
M. Olivucci, Siena, Italy
J.S. Siegel, Tianjin, China
J. Thiem, Hamburg, Germany
M. Venturi, Bologna, Italy
C.-H. Wong, Taipei, Taiwan
H.N.C. Wong, Shatin, Hong Kong
V.W.-W. Yam, Hong Kong, China
S.-L. You, Shanghai, China

Aims and Scope

The series *Topics in Current Chemistry* presents critical reviews of the present and future trends in modern chemical research. The scope of coverage includes all areas of chemical science including the interfaces with related disciplines such as biology, medicine and materials science.

The goal of each thematic volume is to give the non-specialist reader, whether at the university or in industry, a comprehensive overview of an area where new insights are emerging that are of interest to larger scientific audience.

Thus each review within the volume critically surveys one aspect of that topic and places it within the context of the volume as a whole. The most significant developments of the last 5 to 10 years should be presented. A description of the laboratory procedures involved is often useful to the reader. The coverage should not be exhaustive in data, but should rather be conceptual, concentrating on the methodological thinking that will allow the non-specialist reader to understand the information presented.

Discussion of possible future research directions in the area is welcome.

Review articles for the individual volumes are invited by the volume editors.

Readership: research chemists at universities or in industry, graduate students.

More information about this series at
<http://www.springer.com/series/128>

Rita Gerardy-Schahn • Philippe Delannoy •
Mark von Itzstein
Editors

SialoGlyco Chemistry and Biology II

Tools and Techniques to Identify
and Capture Sialoglycans

With contributions by

B. Ernst • J. Finne • K. Fukase • G. Herrler • J. Hirabayashi •
C.-H. Hsu • E. Jakobsson • A. Jokilammi • S. Kelm •
K. Kitajima • H.D. Klenk • A. Kuno • C.-H. Liang •
M. Matrosovich • S. Meinke • C. Sato • O. Schwardt •
D. Schwarzer • K. Tanaka • H. Tateno • J. Thiem •
N. Varki • C.-Y. Wu

 Springer

Editors

Rita Gerardy-Schahn
Institute for Cellular Chemistry
Hannover, Germany

Philippe Delannoy
Lille University of Science and Technology
Villeneuve d'Ascq Cedex, France

Mark von Itzstein
Institute for Glycomics
Griffith University
Southport, Queensland
Australia

ISSN 0340-1022

Topics in Current Chemistry

ISBN 978-3-319-21316-3

DOI 10.1007/978-3-319-21317-0

ISSN 1436-5049 (electronic)

ISBN 978-3-319-21317-0 (eBook)

Library of Congress Control Number: 2015944515

Springer Cham Heidelberg New York Dordrecht London

© Springer International Publishing Switzerland 2015

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, 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.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made.

Printed on acid-free paper

Springer International Publishing AG Switzerland is part of Springer Science+Business Media (www.springer.com)

Preface

Nature's enormous potential for the shaping of structures is made possible by the use of sugars.¹ These molecular building blocks are unique in providing permutation capacity. From a chemical point of view, sugars are polyhydroxy-aldehydes and -ketones which, under physiological conditions, form ring structures (hemiacetals). The relative position of hydroxyl groups (OH- groups) to the plane of the cyclic scaffold determines their chemical and biological properties. For instance, sugars with distinct biological functions such as glucose and galactose differ chemically in no more than the relative positioning of a single OH- group. Moreover, sugars are multivalent and, because of the anomeric freedom of the reducing end, can generate α - or β -linkages to any one of several positions on a second monosaccharide. Thus, the theoretical number of distinct trisaccharides that can be built by the combination of 3 monosaccharides can reach 27,648 whereas 3 different nucleotides or amino acids can form only 6 trimers (for review see [1]).

Sialic acids are acidic nine-carbon sugars that meet all the above discussed aspects and are special because the addition of sialic acid to glycoconjugates occurs exclusively at the non-reducing end. Because of this 'outstanding' position, sialoglycoconjugates form the 'communication front' of animal cells. Mandal and colleagues review this richness of the sialome with a focus on the O-acetylation of sialic acids. O-Acetylation represents a developmentally regulated modification and a marker of some cancer cells such as lymphoblasts in acute lymphoblastic leukemia.

In terms of evolution, sialic acids have an interesting history with an abundant occurrence in the deuterostome lineage and a scattered expression in bacterial pathogens. Importantly, viruses infecting these bacteria bear highly specific receptors recognizing the bacterial sialoglycans. In fact, the use of these unique viral tools in biochemistry laboratories has been of enormous value for the detection and characterization of sialoglycans, as is discussed by Jakobsson and colleagues

¹The term is synonymously used with the term carbohydrates.

(see the chapter titled “Endosialidases: Versatile Tools for the Study of Polysialic Acid”). Moreover, as reviewed by Matrosovich and colleagues, many viral pathogens recognize and intrude on their hosts by exploiting cellular sialoglycans (see the chapter titled “Sialic Acid Receptors of Viruses”).

An organ particularly rich in sialoglycans is the brain. As extensively reviewed by Hildebrandt and Dityatev, a large number of reports has demonstrated the essential nature of sialic acid for brain development and function. While the major form of sialic acid in humans is *N*-acetyl-neuraminic acid (Neu5Ac), most vertebrates, including the great apes, produce *N*-glycolyl-neuraminic acid (Neu5Gc) at similar or even higher concentrations. However, remarkably, in species where Neu5Gc represents the major sialic acid in peripheral tissues, Neu5Gc is rarely found in brain structures, thus raising the question as to whether Neu5Gc could be toxic in the brain (see the review by Davies and Varki). In the light of evolution, this tissue-selective expulsion of Neu5Gc is exciting because the gene encoding the hydroxylase needed for the conversion of CMP-Neu5Ac to CMP-Neu5Gc was lost before the emergence of the genus *Homo*. However, humans can acquire Neu5Gc from dietary sources and a number of studies have shown the presence of Neu5Gc in peripheral tissues. Whether the integration of Neu5Gc into human tissue causes major immune reactions is a matter of debate. Shilova et al. present a primary study in this volume, in which large cohorts of probands were screened for their patterns of natural antibodies. Surprisingly, only low concentrations of natural antibodies against sialoglycans were identified.

The cloning of the major components of the sialylation machinery in mammals generated new targets for the generation of knockout models which have been used to interrogate the role of sialic acids and sialoglycans in organ development and homeostasis. The lessons learned by the use of these mouse models to re-enact sialoglyco pathologies identified in humans are reviewed by Hinderlich and colleagues and Sellmeier and colleagues.

Control of sialoglycoconjugate expression in the mammalian system involves the activity of sialidases as well as anabolic pathways. The interest in this important group of enzymes was underestimated in the past but is currently escalating. This volume is directing major attention to these developments by providing a detailed review on structure-function and phylogenetic analyses prepared by Monti and Miyagi.

Considering that the outside of the animal cells is dominated by the presence of sialoglycans, it is not difficult to deduce that there must be numerous counter-receptors to decipher the information presented in the sialome. This fact requests both novel analytical techniques permitting the quantitative determination of individual glycotopes and techniques that allow a holistic monitoring of variations in the cellular sialome such as those occurring during cell differentiation or in cancerogenesis. This analytical area has generated a new research field which is reviewed in this volume by outstanding experts (Kitajima et al., see the chapter titled “Advanced Technologies in Sialic Acid and Sialoglycoconjugate Analysis” and Hirabayashi et al., see the chapter titled “Development and Applications of the Lectin Microarray”). The search for proteins (factors in general) that specifically

bind to sialoglycans is a success story that impressively demonstrates the added value of interdisciplinary collaborative research activities in the field of glycobiology. Innovative array technologies developed by synthetic organic chemists have paved the way for the search of binding molecules not only at the cellular level but also systematically. These topics are reviewed in three different chapters by Liang et al. (see the chapter titled “Sialoside Arrays: New Synthetic Strategies and Applications”), Tanaka and Fukase (see the chapter titled “Chemical Approach to a Whole Body Imaging of Sialo-N-Linked Glycans”), and Meinke and Thiem (see the chapter titled “Trypanosomal Trans-sialidases: Valuable Synthetic Tools and Targets for Medicinal Chemistry”). Last, but certainly not least, Schwadt et al. (see the chapter titled “SIGLEC-4 (MAG) Antagonists: From the Natural Carbohydrate Epitope to Glycomimetics”) focus on the specificity of sialylated carbohydrate structures and the development of sialic acid derivatives that mimic these structures. With myelin-associated glycoprotein (MAG) as an example, the development of mimetics that bind with low nanomolar affinity is described. The leads identified show great promise to be developed further to prevent the inhibitory activity of MAG on nerve regeneration.

This edition entitled *SialoGlyco Chemistry and Biology I and SialoGlyco Chemistry and Biology II* combines 15 chapters from distinguished authors that, together, form a unique reference book which should be of great interest to researchers and teachers.

Hannover, Germany
Villeneuve d’Ascq Cedex, France
Southport, Australia

Rita Gerardy-Schahn
Philippe Delannoy
Mark von Itzstein

Reference

1. Cohen M, Varki A (2010) *OMICS* 14(4):455–464. doi:10.1089/omi.2009.0148

Contents

Sialic Acid Receptors of Viruses	1
Mikhail Matrosovich, Georg Herrler, and Hans Dieter Klenk	
Endosialidases: Versatile Tools for the Study of Polysialic Acid	29
Elina Jakobsson, David Schwarzer, Anne Jokilammi, and Jukka Finne	
Advanced Technologies in Sialic Acid and Sialoglycoconjugate Analysis	75
Ken Kitajima, Nissi Varki, and Chihiro Sato	
Development and Applications of the Lectin Microarray	105
Jun Hirabayashi, Atsushi Kuno, and Hiroaki Tateno	
Sialoside Arrays: New Synthetic Strategies and Applications	125
Chi-Hui Liang, Che-Hsiung Hsu, and Chung-Yi Wu	
SIGLEC-4 (MAG) Antagonists: From the Natural Carbohydrate Epitope to Glycomimetics	151
Oliver Schwardt, Soerge Kelm, and Beat Ernst	
Chemical Approach to a Whole Body Imaging of Sialo-<i>N</i>-Linked Glycans	201
Katsunori Tanaka and Koichi Fukase	
Trypanosomal Trans-sialidases: Valuable Synthetic Tools and Targets for Medicinal Chemistry	231
Sebastian Meinke and Joachim Thiem	
Index	251

Sialic Acid Receptors of Viruses

Mikhail Matrosovich, Georg Herrler, and Hans Dieter Klenk

Abstract Sialic acid linked to glycoproteins and gangliosides is used by many viruses as a receptor for cell entry. These viruses include important human and animal pathogens, such as influenza, parainfluenza, mumps, corona, noro, rota, and DNA tumor viruses. Attachment to sialic acid is mediated by receptor binding proteins that are constituents of viral envelopes or exposed at the surface of non-enveloped viruses. Some of these viruses are also equipped with a neuraminidase or a sialyl-*O*-acetyl-esterase. These receptor-destroying enzymes promote virus release from infected cells and neutralize sialic acid-containing soluble proteins interfering with cell surface binding of the virus. Variations in the receptor specificity are important determinants for host range, tissue tropism, pathogenicity, and transmissibility of these viruses.

Keywords Ganglioside · Mucins · Neuraminidase · Receptor binding · Receptor-destroying enzyme · Sialate-*O*-acetyl-esterase · Virus

Contents

1	Introduction	2
2	Orthomyxoviridae	2
2.1	Influenza A and B Viruses	3
2.2	Influenza C virus	7
2.3	Isavirus	9
3	Coronaviridae	9
3.1	Betacoronaviruses	9

M. Matrosovich and H.D. Klenk (✉)
Institut für Virologie, Philipps-Universität, Hans-Meerwein-Str. 2, 35043 Marburg, Germany
e-mail: klenk@staff.uni-marburg.de

G. Herrler
Institut für Virologie, Stiftung Tierärztliche Hochschule Hannover, Bünteweg 17, 30559
Hannover, Germany

3.2	Alphacoronaviruses	11
3.3	Gammacoronaviruses	12
3.4	Torovirus	12
4	Paramyxoviridae	13
5	Caliciviridae	14
6	Picornaviridae	14
7	Reoviridae	15
8	Polyomaviridae	16
9	Adenoviridae	17
10	Parvoviridae	17
	References	18

1 Introduction

The initial step in the viral life cycle is the attachment of virus particles to the cell surface. Attachment is mediated by binding of the virus to a receptor. Sometimes co-receptors are also involved that might promote post-attachment events in the entry process. Receptor molecules are constituents of the cell membrane, and the receptor determinant, the structure to which the virus binds, may be either a protein epitope or the carbohydrate of a glycoprotein or a glycolipid. Soluble proteins present in body fluids and in mucus on respiratory and enteric epithelia may also contain such carbohydrates and therefore interfere with virus binding to the cell surface.

Sialic acid was the first virus receptor identified. Hirst and McClelland and Hare discovered that influenza virus is able to hemagglutinate and that adsorbed virus is eluted from erythrocytes on incubation at 37°C, indicating an enzymatic destruction of a receptor substance on the cells [1, 2]. When a similar enzymatic activity was subsequently detected in *Vibrio cholerae* cultures, the term “receptor-destroying enzyme” was introduced [3]. The substance released by the viral enzyme from soluble hemagglutination inhibitors was initially characterized as a carbohydrate of low molecular weight [4] and then identified in crystalline form as *N*-acetyl-D-neuraminic acid [5]. Thus, it was clear that the receptor determinant of influenza virus was sialic acid and that the viral enzyme was a neuraminidase. Furthermore, for the first time an important biological function of sialic acid had been identified.

Sialic acid has later also been found to serve as receptor of a large spectrum of other viruses. Most of them will be addressed here, with emphasis, however, on influenza viruses. For additional information we refer to several excellent reviews that have been published in recent years on similar topics [6–10].

2 Orthomyxoviridae

The orthomyxoviruses are enveloped viruses with a single-stranded, segmented RNA genome of negative polarity [11, 12]. There are five genera in the family: *Influenza virus A*, *B*, and *C*, *Thogotovirus*, and *Isavirus*. Influenza A viruses are further divided into subtypes characterized by 16 different hemagglutinins

(H1–H16) and 9 different neuraminidases (N1–N9). Except for the *Thogotovirus* genus, all orthomyxoviruses bind to sialic acid receptors. The receptor of an influenza A virus of subtype H17N10 isolated recently from bats [13] is not known.

2.1 Influenza A and B Viruses

Influenza A viruses are important human and animal pathogens. Their primary natural hosts are aquatic birds from which they are occasionally transmitted to other species. In man they cause outbreaks of respiratory disease that occur as annual epidemics and less frequent pandemics. Influenza B viruses are also believed to be descendants of avian influenza A viruses, but are now largely restricted to humans where they cause respiratory infections as well. Influenza A and B viruses have two envelope glycoproteins, the hemagglutinin (HA) and the neuraminidase (NA), both of which interact with sialic acid.

2.1.1 Hemagglutinin

HA initiates infection by binding to cell surface receptors and by inducing fusion between viral and cellular membranes. HA is integrated in the virus envelope as a type I membrane protein. The ectodomain of HA represents 90% of the polypeptide chain. The residual 10% of the HA sequence accounts for the transmembrane domain and the cytosolic domain. HA is synthesized as a precursor molecule HA0 (75 kDa) which assembles to homotrimers. HA0 is *N*-glycosylated, palmitoylated, and proteolytically cleaved by host enzymes. The amino-terminal cleavage fragment HA1 (50 kDa) contains the receptor binding site and the carboxy-terminal fragment HA2 (25 kDa) is membrane anchored and responsible for fusion (reviewed in [14]).

The receptor determinant of influenza A and B viruses is sialic acid, mostly *N*-acetyl-neuraminic acid (Neu5Ac). The structures of complexes of HA of influenza A and B viruses with sialyloligosaccharides were determined by X-ray crystallography (reviewed in [15, 16]). The sialic acid-binding site is a shallow pocket located on the globular head of HA (Fig. 1). Virus binding depends not only on HA affinity for the terminal sialic acid residues, but also on the structure of the underlying oligosaccharide and protein or lipid moieties of the receptors, as well as on the abundance and accessibility of receptors on the cell surface. Because of this complex mode of binding, the receptor-binding properties of influenza viruses can be affected by amino acid substitutions inside the sialic acid-binding pocket, on the pocket rim, and by distant mutations resulting in altered glycosylation or altered electrostatic charge of the globular head of HA (reviewed in [17]). In natural glycoconjugates, sialic acids are α 2-3- or α 2-6-linked to Gal and GalNAc, α 2-6-linked to GlcNAc, or α 2-8-linked to the second Sia residue. Influenza viruses generally do not bind to α 2-8-linked

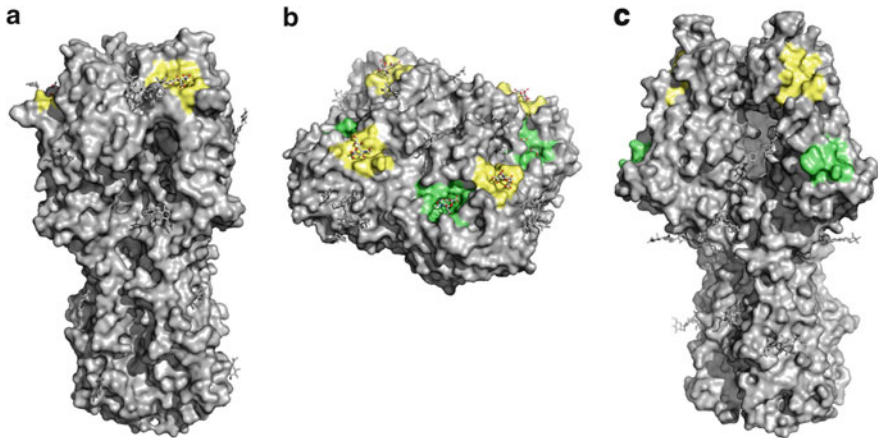


Fig. 1 Sialic acid binding sites of the hemagglutinin (a) and the neuraminidase (b) of influenza A virus and the hemagglutinin-esterase-fusion protein of influenza C virus (c). Molecular surfaces of HA and HEF trimers and the NA tetramer are shown. Receptor-binding sites of HA, HEF and the hemadsorption site of NA are colored *yellow*. The catalytic sites of NA and HEF are colored *green*. Sialic acid moieties in the binding sites of HA and NA are shown as stick models. The figure is based on crystal structures 1MQM, 1W20, and 1FLC from Protein Data Bank

Neu5Ac and can recognize only α 2-3- or α 2-6-linked sialic acid moieties such as Neu5Ac α 2-3/6Gal, Neu5Ac α 2-3/6GalNAc, and Neu5Ac α 2-6GlcNAc.

Differences in receptor-binding specificity of influenza viruses can contribute to viral host range restriction. Thus, human influenza viruses preferentially bind to α 2-6-linked sialic acids (Neu5Ac α 2-6Gal), whereas avian influenza viruses preferentially recognize Neu5Ac α 2-3Gal [18–20]. These preferences are matched by predominant expression of Neu5Ac α 2-6Gal on epithelial cells in the human airway epithelium and by abundance of Neu5Ac α 2-3Gal on epithelial cells in the intestinal and respiratory tract of birds [21–26]. The receptor-binding specificity of human and avian influenza viruses suggests that avian viruses need to acquire the ability to recognize human-type receptors to be able to replicate efficiently and transmit in humans. Indeed, the earliest isolates of the 1918, 1957, and 1968 pandemics possessed HA that, although of avian origin, recognized human-type receptors (reviewed in [27, 28]). In light of these findings, the infection of humans with highly pathogenic avian H5N1 viruses seemed to be surprising as H5N1 viruses isolated from infected individuals preferentially recognize Neu5Ac α 2-3Gal [29–31]. Studies on human and avian virus infection in differentiated cultures of human airway epithelial cells indicated, however, that some cells in the human airway epithelium express sufficient amounts of receptors to allow infection with avian viruses and that receptor specificity determines the viral cell tropism in the epithelium. Early in infection, human viruses preferentially infected non-ciliated cells, whereas avian viruses mainly infected ciliated cells [32]. Other groups studied expression of viral receptors in human biopsies and archival tissues using lectins *Sambucus nigra* agglutinin, *Maackia amurensis* agglutinins I and II, and human and avian influenza viruses as molecular probes

[26, 33–36]. The results obtained in these studies suggest that paucity of receptors for avian viruses in the upper respiratory tract in humans is one of the factors preventing efficient human-to-human transmission. This concept is supported by recent studies showing that H5N1 mutants binding to α 2-6 linked sialic acid are transmitted between ferrets through the air [37, 38].

Because pigs support replication of both avian and human viruses, they were considered to be a plausible intermediate host for the generation of human pandemic strains by gene reassortment (reviewed in [39]). This theory was further supported by the finding that both 3-linked and 6-linked sialic acid moieties were detected by staining on the histological sections of pig tracheal epithelium [23]. All early studies on swine influenza viruses were done using viruses that were grown in embryonated chicken eggs. However, similar to human influenza viruses, swine viruses appear to change their receptor specificity in eggs. Indeed, non-egg-adapted classical swine influenza viruses that were isolated and propagated solely in MDCK cells displayed a strict preference for 6-linked sialic acids and did not bind to 3-linked sialic acids [40]. This binding pattern is typical for non-egg-adapted human influenza viruses, and it is in discordance with the previously described ability of egg-adapted swine influenza viruses to recognize Neu5Ac α 2-3Gal [19, 23]. Thus, the receptor specificity of the pig viruses may be even closer to that of human viruses than originally thought. This notion agrees with recent data on a close similarity in the distribution of sialic acid receptors in the respiratory tract of pigs and humans [24, 26, 41].

The receptor specificity of the novel swine-origin H1N1/2009 pandemic influenza virus has been analyzed in studies employing carbohydrate microarrays. In some of these studies the virus was found to bind exclusively to α 2-6-linked sialyl sequences [42–45], whereas in another study using a different microarray some binding to probes containing α 2-3-linkages was also observed [46]. These studies showed also that the H1N1/2009 pandemic virus displayed the same binding profile as its putative swine precursors. The results indicate that no major change in receptor-binding specificity of HA was required for the emergent pandemic virus to acquire human-like characteristics and become established in the human population. Interestingly, mutations in the receptor-binding site of the HA of H1N1/2009 viruses have been detected sporadically, and the D222G substitution has been associated with severe or fatal disease [47, 48]. Compared to the parental virus, the D222G mutant virus displayed enhanced binding to α 2-3-linked sugars [45, 49], infected a higher proportion of ciliated cells in cultures of human airway epithelium [49], and showed an altered pattern of attachment to human respiratory tissues *in vitro*, in particular increased binding to macrophages and type II pneumocytes in the alveoli [50]. These results suggested that the association of the D222G mutation with severe disease in humans reflects receptor-mediated alteration of the cell tropism of the mutant in human respiratory epithelium with enhanced replication in the lower respiratory tract.

Based on early data [18, 20, 51], it was assumed that all avian influenza viruses have similar receptor-binding specificity. The first evidence against this theory was obtained in a study showing that H5N1 viruses isolated in Hong Kong in 1997 from

poultry and humans had a lower receptor binding affinity and a lower neuraminidase activity than closely related viruses of aquatic birds [30].

Subsequent detailed receptor-binding studies revealed that influenza viruses adapted to ducks, gulls, and land-based gallinaceous poultry differ in their ability to recognize the sub-terminal saccharides of Neu5Ac α 2-3Gal-terminated receptors (reviewed in [28, 52]). In particular, duck viruses preferentially bind to receptors with type 1 and type 3 oligosaccharide sequences, such as Neu5Ac α 2-3Gal β 1-3GlcNAc and Neu5Ac α 2-3Gal β 1-3GalNAc, and viruses isolated from gulls show high-avidity binding to fucosylated sialyloligosaccharides Neu5Ac α 2-3Gal β 1-4(Fuc α 1-3)GlcNAc and Neu5Ac α 2-3Gal β 1-3(Fuc α 1-4)GlcNAc. In contrast, poultry-adapted viruses preferentially bind to receptors with type 2 sequences, such as Neu5Ac α 2-3Gal β 1-4GlcNAc, with particularly strong binding to the corresponding sulfated analogues Neu5Ac α 2-3Gal β 1-4(6-O-HSO₃)GlcNAc and Neu5Ac α 2-3Gal β 1-4(Fuc α 1-3)(6-O-HSO₃)GlcNAc. Furthermore, some viruses of the Eurasian lineage of H9N2 poultry viruses bind to Neu5Ac α 2-6Gal terminated sialyloligosaccharides [28]. Thus it seems that influenza viruses circulating in different birds can have different receptor specificity owing to distinctions between the sialic acid receptors in these avian species.

2.1.2 Neuraminidase

NA is a type II membrane protein that is present in homotetrameric form in the viral envelope [53]. Each monomer consists of a cytoplasmic tail six amino acids in length, a stem region varying in length between 19 and 45 amino acids, and a carboxy-terminal globular head [53, 54]. The monomers are linked to dimers by disulfide bridges in the stalk region. The available evidence indicates that the neuraminidase has several functions in the life cycle of influenza virus. It was Burnet who proposed more than 60 years ago that the RDE allows the virus to penetrate the mucus layer coating the respiratory epithelium and thus to infect its target cells [3]. This concept has recently been shown to be correct when it was found that the neuraminidase inhibitor oseltamivir prevented initiation of infection of human tracheo-bronchial cell cultures [55]. The second function of the neuraminidase is at the end of the life cycle where it promotes virus release and prevents clumping of virions by removing receptors from the cell surface and viral glycoproteins, respectively [56].

Interspecies transmission of avian influenza viruses from aquatic birds to terrestrial poultry is often accompanied by a deletion in the stalk region of the NA and reduced catalytic activity [30, 57]. The observation that the reduced catalytic activity of NA is compensated by mutations in HA resulting in decreased receptor affinity led to the concept that optimal virus replication depends on a balance between receptor binding by HA and receptor destruction by NA [58–61].

The catalytic site of NA is located in the globular head region (Fig. 1). It is in the center of a propeller-like structure formed by four anti-parallel β -sheets [53]. *N*-Acetyl-neuraminic acid is bound by hydrogen bonds to amino acids

R118, D151, R152, R224, E276, R292, and R371 (N2 numbering). The acetamido group is linked by van der Waals forces to W178 and I222. The amino acids directly interacting with sialic acid are stabilized by contacts with amino acids E119, R156, S179, D/N198, N294, and E425. All of these amino acids are conserved among different NA subtypes.

NAs of avian influenza viruses have, in addition to the catalytic function, the capacity to agglutinate erythrocytes [62–64]. NAs of human viruses are unable to hemadsorb. The hemadsorption site is a shallow pocket located close by, but separately from the deep catalytic site (Fig. 1) [65]. It is formed by three amino acid loops, with residues S367, S370, and S372 in the first, N/I400 and W403 in the second, and E/K/Q/N432 in the third loop, directly interacting with the sialic acid moiety. Recently it could be shown that the hemadsorption function enhances the catalytic activity of NA. This study also revealed that the hemadsorption activity of the NAs of early human isolates of the pandemics of 1918 and 1957 was reduced or completely absent. Thus, it appears that loss of the hemadsorption site is the result of an adaptive mutation involved in interspecies transmission from bird to man and has therefore to be considered as a pandemic marker [66].

2.2 *Influenza C virus*

Influenza C viruses that cause mild respiratory infections in humans differ from other influenza viruses because (1) their preferred sialic acid is *N*-acetyl-9-*O*-acetylneuraminic acid (Neu5,9Ac₂), (2) their receptor-destroying enzyme is an acetyl esterase rather than a neuraminidase, and (3) three functions are combined in one surface glycoprotein, the hemagglutinin-esterase-fusion (HEF) protein: sialic acid binding, esterase and membrane fusion activity as compared to influenza A and B viruses where sialic acid binding and neuraminidase activity are distributed on two glycoproteins, the HA and NA proteins.

The HEF protein is a type I membrane protein of about 80 kDa [67, 68]. It is synthesized as a precursor (HEF0) that is post-translationally cleaved into the subunits HEF1 and HEF2. HEF1 comprises the sialic acid-binding and esterase activity and is connected via disulfide bonds to the membrane-bound HEF2 subunit. Despite little sequence similarity, HEF and HA show surprising structural similarity. The receptor domain of HEF is inserted into a surface loop of the esterase domain and the esterase domain is inserted into a surface loop of the stem which includes the hydrophobic peptide at the aminoterminal of HEF2 that is crucial for the fusion activity [69]. The sialic acid binding site is a cavity at the tip of each HEF1 subunit. The active site of the acetyl esterase is located at the base of the globular head region. In the viral spikes HEF is present in homotrimeric form [69] (Fig. 1).

2.2.1 The Esterase Activity of HEF

The receptor-destroying enzyme of influenza C viruses was identified as an acetyl esterase that releases the 9-*O*-acetyl residue from Neu5,9Ac₂ [70]. No or little activity was observed when the *O*-acetyl groups were linked to C-4 or C-7 of sialic acid. The enzyme belongs to the class of serine hydrolases with a catalytic triad formed by residues S57, D352, and H355 [69, 71–73]. The biological importance of the acetyl esterase activity of HEF is believed to be similar to that of the neuraminidase of influenza A virus, i.e., facilitating virus spread by inactivation of potential receptor determinants from the surface of the infected cells and from the viral surface. In the initial stage of the infection cycle, the receptor-destroying enzyme may facilitate virus entry, e.g., by enabling virus to penetrate the mucus layer covering the respiratory epithelium [74]. In the late stage of the growth cycle, inactivation of receptor determinants may promote release of viruses from the infected cell and may prevent the formation of virus aggregates [75]. Supporting evidence has been provided by studies involving enzyme inhibitors, sialic acid analogues, and de- and resialylation experiments [73, 76, 77].

2.2.2 Sialic Acid Binding Activity of HEF

The identification of the receptor-destroying enzyme of influenza C virus as a sialate 9-*O*-acetyl esterase indicated that Neu5,9Ac₂ is a receptor determinant for this virus [70]. Formal proof for the importance of 9-*O*-acetylated sialic acid was provided by desialylation and resialylation of cultured cells which abolished and regenerated agglutination of erythrocytes [78] as well as susceptibility of cultured cells to infection by influenza C virus [79]. The results demonstrated the role of Neu5,9Ac₂ for the cell tropism of the virus. As sialic acids are present on many cell surface glycoconjugates, attempts to identify a specific receptor for virus infection have failed so far for influenza A and B viruses. In the case of influenza C virus, overlay binding assays with immobilized membrane proteins indicated that the major interaction partner on the surface of the susceptible cell line MDCK I is gp40, a mucin-type glycoprotein with a high content of *O*-glycans [80, 81].

Crucial amino acids for substrate binding are residues Y127, T170, and G172 [69]. The specificity for the 9-*O*-acetyl group is determined by Y224 and R236 that interact with the carbonyl oxygen and by residues W225, W293, and P271 that form a pocket for the methyl group. Interestingly, influenza C virus can adapt to growth in cells with a low content of Neu5,9Ac₂. Passage in such cells or establishment of a persistent infection resulted in viruses with increased binding affinity to 9-*O*-acetylated sialic acids. These mutants or variant viruses had mutations at residues 269, 270, or 272, i.e., next to the above-mentioned P271 [82–85].

2.3 *Isavirus*

Infectious salmon anemia virus (ISAV) is an important pathogen in farmed Atlantic salmon. Similar to influenza viruses it has a hemagglutinating and a receptor-destroying activity. Unlike influenza A and B viruses, the RDE is not a neuraminidase but an acetylase [86]. The enzyme belongs to the class of serine hydrolases [86, 87]. Unlike the HEF protein of influenza C virus, the ISAV esterase releases the 4-*O*-acetyl group of 4-*O*-5-*N*-acetylneuraminic acid (Neu4,5Ac₂) [88]. This enzymatic activity corresponds to the preferred ligand of the ISAV hemagglutinin which is also Neu4,5Ac₂ [88]. Both the sialic acid binding activity and the acetylase activity are functions of the 38–43-kDa surface glycoprotein which has been designated HE protein [89–91].

3 Coronaviridae

Coronaviruses (order *Nidovirales*, family *Coronaviridae*) are a diverse group of viruses that cause enteric, respiratory, and neural infections in both mammalian and avian species. According to a current proposal to the International Committee of Taxonomy of Viruses, they are classified within the subfamily *Coronavirinae* which comprises four genera: *Alpha*-, *Beta*-, *Gamma*-, and *Deltacoronavirus*. The diversity of coronaviruses is also evident in the sialic acid binding activity. Some members of the *Betacoronavirus* genus, e.g., bovine coronavirus (BCoV), recognize *O*-acetylated sialic acids and contain an acetylase that functions as a receptor-destroying enzyme. On the other hand, some *alpha*- and *gamma*coronaviruses lack a comparable enzyme and have a preference for *N*-acetyl- or *N*-glycolylneuraminic acid, the best studied examples being transmissible gastroenteritis virus (TGEV) and infectious bronchitis virus (IBV). In addition to the above-mentioned viruses, both *alpha*- and *gamma*coronaviruses also include members that lack any sialic acid binding activity, e.g., SARS coronavirus and human coronavirus 229E. In the following, the sialic acid binding activities of BCoV, TGEV, and IBV will be described in more detail.

3.1 *Betacoronaviruses*

The presence of an acetylase in coronaviruses was first described by Vlasak and coworkers who showed that BCoV and HCoV-OC43 eluted from the erythrocytes during the course of a hemagglutination reaction, rendering the cells resistant to subsequent agglutination by either of the two coronaviruses or by influenza C virus. This finding demonstrated that BCoV and HCoV-OC43, similar to influenza C viruses, have a sialate 9-*O*-acetylase that functions as a receptor-destroying enzyme [92].

The acetylase activity was assigned to the HE surface glycoprotein of BCoV, hemagglutinating encephalomyelitis virus (HEF), and mouse-hepatitis virus [93–95]. The three-dimensional structure of the HE protein of BCoV has been determined showing an esterase site similar to that of the influenza C virus HEF protein [96]. By contrast, the sialic acid binding site of HE differs from that of the HEF protein with the ligand bound in the opposite orientation. An HE gene is present only in members of the *Betacoronavirus* genus. The different strains of murine coronaviruses contain an HE gene but differ widely in the amount of protein expressed. The acetylase of murine coronaviruses has been shown to have a different substrate specificity compared to that of BCoV, HEV, and HCoV-OC43, which release the *O*-acetyl residue from position C-9 of sialic acids. By contrast, murine coronaviruses – with the exception of the diarrhea virus of infant mice [97] – preferentially hydrolyze the ester linkage of 4-*O*-acetyl-*N*-acetylneuraminic acid [98–100].

The biological role of the acetylase of the betacoronaviruses is assumed to be similar to that of the receptor-destroying enzymes of influenza viruses, i.e., it may inactivate binding sites for the virus (1) on the cell surface and thus allow virus release from the infected cell, (2) on mucins covering the respiratory epithelial cells and thus facilitate the penetration of the mucus layer, and (3) on viral surface glycoproteins or glycolipids and thus prevent aggregate formation. Conflicting data have been reported concerning the role of the receptor-destroying enzyme in the initial stage of infection. Inhibition of the acetylase by diisopropyl fluorophosphate was shown to reduce the infectivity about a hundredfold in one report, and to have no effect in another report [94, 101].

Following the discovery of an acetylase in BCoV and HCoV-OC43 [92] it has been shown that 9-*O*-acetylated sialic acid serves as a receptor determinant not only for binding to erythrocytes but also for initiating infection of cultured cells [102]. When polarized epithelial cells such as MDCK I cells were analyzed for susceptibility to infection, BCoV was found to infect the cells via the apical but not via the basolateral side of the membrane [103, 104]. The inability of BCoV to infect MDCK I cells via the basolateral plasma membrane may reflect that the major glycoprotein recognized by BCoV, a mucin-like glycoprotein of 40 kDa, is predominantly present in the apical membrane domain [105]. An alternative explanation is that BCoV requires an additional receptor for initiation of infection, which is present only on the apical membrane. Such a secondary receptor has not yet been identified for BCoV.

The HE protein of BCoV has not only acetylase activity (see above); it can also function as a hemagglutinin [93, 106, 107]. However, BCoV agglutinates a wider spectrum of erythrocytes than does the isolated HE protein. HE only agglutinates cells that contain a high content of Neu5,9Ac₂ such as mouse and rat erythrocytes. Chicken erythrocytes are agglutinated by BCoV and HCoV-OC43, but not by the HE protein. The second surface glycoprotein of BCoV, the S protein, has an important function in virus entry by being involved in the attachment of virions to the cell surface and by mediating the subsequent fusion of the viral and the cell membrane. By contrast to the HE protein, isolated S protein is able to

agglutinate chicken erythrocytes [108]. Therefore, the S protein of these viruses is the major hemagglutinin and thus the major sialic acid binding protein. With murine coronaviruses, where the role of *O*-acetylated sialic acids as an essential receptor determinant has not been demonstrated [109], a sialic acid binding activity could be assigned only to the HE protein, not to the S protein [110].

3.2 *Alphacoronaviruses*

TGEV is an enteropathogenic virus which may affect pigs of all ages. Infections are especially severe in piglets up to two weeks of age which usually die unless they are protected by maternal antibodies. When Noda and co-workers [111, 112] first described the ability of TGEV to agglutinate erythrocytes, the virus appeared to contain a weak hemagglutinin. This is probably related to the absence of a receptor-destroying enzyme that may remove competitive inhibitors from the viral surface. In fact, when a virus or cells used for virus growth were pre-treated with neuraminidase, the resulting virions were able to agglutinate erythrocytes efficiently. In this way it was shown that the HA-activity of TGEV was due to a sialic acid-binding activity with a preference for α 2-3-linked *N*-glycolylneuraminic acid [104, 113].

The sialic acid binding activity of TGEV is located in the amino-terminal portion of the surface glycoprotein S between amino acids 20 and 244. Evidence is based on the hemagglutination-inhibiting effect of monoclonal antibodies and on the analysis of mutant proteins with one or more amino acid exchanges [104, 113]. Interestingly, all mutants that had lost hemagglutinating activity were strongly reduced in their enteropathogenic effect, indicating that the sialic acid binding activity is an important factor for the enteropathogenicity of TGEV [113–115].

In virus overlay binding assays with brush border membranes from suckling piglets, TGEV recognized a high molecular mass protein via its sialic acid binding activity [116]. This highly glycosylated protein was designated MGP (mucin-like glycoprotein) as it possesses typical characteristics of a mucin. In *in situ* binding assays with jejunal cryosections, TGEV bound in a sialic acid-dependent manner to a component that was mainly localized in the goblet cells which are known to synthesize and secrete mucins [116]. From these data it can be concluded that binding to the sialic acids of MGP helps the virus to penetrate the mucus layer and to proceed to the intestinal enterocytes for initiation of infection.

This explanation also applies to an interesting phenomenon related to TGEV. A respiratory variant of TGEV, the porcine respiratory coronavirus (PRCoV), was first isolated in Belgium [117] and found to be very similar to TGEV. The major difference was a deletion of 224 amino acids in the N-terminal half of the S protein. Both TGEV and PRCoV use pAPN (porcine aminopeptidase N) as a receptor to infect their host cells [118]. In contrast to TGEV, the S protein of PRCoV displays no hemagglutinating activity as the sialic acid binding site is located in the deleted region of the S protein [104]. PRCoV does not replicate efficiently in the gut [119]. As the S proteins of TGEV and PRCoV share the binding sites for

neutralizing antibodies, the spread of PRCoV in European pigs acted like the spread of a vaccine virus, resulting in drastic reduction of TGEV infection. Though PRCoV, similar to TGEV, uses pAPN as a cell surface receptor for entering host cells, PRCoV, unlike TGEV, is not an enteropathogenic virus. As in the case of the mutants mentioned above, the lack of sialic acid binding activity appears to be responsible for the lack of enteropathogenicity.

Though sialic acids are the receptor determinants for the HA activity of TGEV and are crucial for the enteropathogenicity of the virus, the sialic acid binding activity appears to be dispensable for growth of the virus in cell culture. TGEV mutants deficient in sialic acid binding activity grow well in cell culture using pAPN as receptor [113, 115]. However, in binding assays the amount of parental virus attached to sialic acids on the cell surface was increased sixfold compared to mutant virus that was only able to bind to pAPN [120]. Recent results demonstrated that binding to sialic acids is dispensable for infection of cultured cells, when a conventional adsorption time is applied, i.e., 60 min. However, when the adsorption time is reduced to 5 min, infection becomes sialic acid-dependent, as indicated by the effect of pretreatment of cells with neuraminidase, which resulted in a more than 80% reduction of infectivity. This result indicates that the sialic acid binding activity can facilitate infection under unfavorable conditions [121] and therefore may be necessary for infection of the intestine.

3.3 *Gammacoronaviruses*

Bingham and coworkers [122] reported that some IBV strains were able to agglutinate erythrocytes. Similar to TGEV, IBV requires pretreatment with neuraminidase for efficient hemagglutinating activity. Furthermore, it preferentially recognizes α -2-3-linked sialic acid [123]. Recently, it has been shown that sialic acid is also a crucial receptor-determinant for infection of cells [124]. Pretreatment with neuraminidase was found to result in decreased infectivity as indicated by a reduced number of infected cells and by lower titers of virus released into the supernatant. This finding was obtained with both a lab strain and strains circulating in poultry [124–126]. The sialic acid-dependence of the IBV infection was observed both with conventional cell cultures and differentiated airway epithelial cells from trachea and lung [126, 127].

3.4 *Torovirus*

Toroviruses belong to the family *Coronaviridae* and are classified within the subfamily *Torovirinae* and the genus *Torovirus*. They cause mild infections of swine and cattle [6]. Toroviruses contain an HE protein that resembles the HE proteins of betacoronaviruses [128, 129]. Like the coronaviral counterparts,

torovirus HE proteins are acetylsterases. The enzyme of bovine torovirus releases the *O*-acetyl group from position C-9 of sialic acid and accepts as a substrate both Neu5,9Ac₂ and *N*-acetyl-7(8),9-*O*-acetylneuraminic acid; this specificity resembles those that have been reported for the HEF protein of influenza C virus and for the HE proteins of several coronaviruses [100]. By contrast, the HE protein of porcine torovirus has a narrower specificity, accepting Neu5,9Ac₂ but not Neu5,7(8),9Ac₃ as a substrate [100]. Analysis of the crystal structure revealed that the torovirus HE proteins have an esterase domain similar to those of the coronavirus HE and influenza C virus HEF proteins [130]; on the other hand, the sialic acid binding site is unique. The difference in substrate specificity is explained by a single amino acid, Thr73 in the porcine and Ser64 in the bovine HE protein [130].

4 Paramyxoviridae

The *Paramyxoviridae* family that is divided into two subfamilies and seven genera comprises a large group of enveloped viruses with non-segmented single-stranded RNA genomes of negative polarity. The members of the genera *Respirovirus*, *Rubulavirus*, and *Avulavirus* are viruses that share binding to sialic acid-containing cell receptors as a common feature. They include several major pathogens for man (human parainfluenza viruses (HPIV) 1-4, mumps virus) and animals (Newcastle disease virus (NDV)) as well as Sendai virus that became an important tool in genetic engineering because of its capacities as a membrane fusing agent and a gene vector.

Receptor interaction of these viruses is mediated by the hemagglutinin-neuraminidase (HN) glycoprotein, a type II integral membrane protein with an N-terminal cytoplasmic tail, a transmembrane domain, a membrane-proximal stalk domain, and a large C-terminal globular head domain that contains the sites responsible for hemagglutinating and neuraminidase activities. HN forms tetramers that are present as spikes on the surface of the virus particles (see [131]).

HN is believed not only to initiate infection by receptor binding but also to prevent aggregation and to promote release of mature virions by receptor removal. X-Ray crystallographic analysis of the HN glycoprotein of NDV [132], HPIV3 [133], and SV5 [134] has revealed a typical neuraminidase fold consisting of six antiparallel β strands organized as a super barrel with a centrally located active site located at the tip of the globular head domain. This exerts both the receptor binding and the catalytic function. A second sialic binding site has been observed on HN of NDV, the biological function of which, however, has not been clearly established yet [132]. HPIV1 HN also has a second binding site, but it is accessible only after removal of a nearby carbohydrate side chain [135].

The receptor specificity of Sendai virus was first analyzed in studies employing gangliosides [136, 137] and erythrocytes that contained defined sialyloligosaccharides after neuraminidase and subsequent sialyltransferase treatment [138]. These studies showed that Sendai virus has a preference for α 2-3-bound

N-acetylneuraminic acid. This receptor determinant appears to be present on both glycoproteins and gangliosides [139, 140]. HPIV1 also recognizes α 2-3 linkages, whereas HPIV3 has α 2-6 specificity [141].

5 Caliciviridae

Caliciviruses are small non-enveloped viruses that contain a single-stranded plus-sense RNA genome encapsidated by an icosahedral protein shell. The major capsid protein VP1 has a shell (S) domain and a protruding (P) domain [142, 143]. The P domain which forms arch-like structures on the virion surface is further subdivided into subdomains P1 and P2. P2 is the most variable region and contains carbohydrate binding motifs [144–148].

Caliciviruses which occur in a large variety of different hosts are subdivided into several genera, including the genus *Norovirus*. Human noroviruses are responsible for the majority of acute viral gastroenteritis. Although these infections are usually mild they can be a serious threat to the elderly and the immuno-compromised. Murine noroviruses share pathogenic properties with human noroviruses as they are enteric viruses that replicate in the intestine and are shed in feces [149].

Whereas most human noroviruses bind to non-charged histo-blood group antigens [150–152] or to heparan sulfate [153], some recognize sialyl-Lewis X neoglycoproteins. Binding to the sialyl-Lewis X group is strictly sialic acid-dependent, since a non-sialylated control glycan does not bind [154]. While the tropism of human norovirus remains unknown, murine noroviruses efficiently replicate in murine macrophages and dendritic cells [149]. Virus binding to the macrophage surface is partially neuraminidase-sensitive and ganglioside-dependent [155]. Murine macrophages express gangliosides GD1a and GM1, and murine norovirus binds to GD1a, but not to GM1, suggesting that the minimal binding epitope is the terminal sialic acid found in GD1a [148]. Only in a few other instances has sialic acid been identified as a calicivirus receptor. Thus, a feline calicivirus strain attaches to α 2-6-linked sialic acid on *N*-glycans [156].

6 Picornaviridae

Among the *Picornaviridae*, a large family of non-segmented positive-stranded RNA viruses comprising many animal and human pathogens, the use of sialic acid as a receptor component has been described for encephalomyocarditis virus [157], Theiler's murine encephalomyelitis virus (TMEV) [158], mengovirus [159], and bovine enterovirus 261 [160]. A human enterovirus also attaches to sialic acid, with a strong preference for *O*-linked glycans containing sialic acid α 2-3-linked to galactose [161]. Differences in receptor specificity appear to be virulence markers of TMEV. While strains with high neurovirulence bind to heparan sulfate, low

neurovirulence strains bind to α 2-3-linked sialic acid moieties on *N*-glycans [162]. Crystallographic studies revealed a positively charged area on the viral surface in contact with sialic acid through non-covalent hydrogen bonds to be important for the persistent infection of the non-neurovirulent strain [158].

7 Reoviridae

These viruses have a segmented double-stranded RNA genome that is encapsidated by one, two, or three protein layers. The icosahedral virions are not enveloped and have a diameter of about 80 nm. There are 12 genera in the virus family. Binding to sialic acid has been observed with many members of the *Rotavirus* genus (for references see below), and some viruses belonging to serotypes 1 and 3 of the *Orthoreovirus* genus [163] also recognize such receptors.

Orthoreoviruses occur with a variety of vertebrates. Infection in humans is generally benign, but may cause upper respiratory tract illness and possibly enteritis in infants and children. Infection is initiated by receptor binding of the sigma 1 protein located in the outer layer of the viral capsid. Sigma 1 forms trimers and is composed of a fibrous tail containing the sialic acid-binding site and a globular head domain that interacts with junctional adhesion molecule 1 (JAM-1) serving as a secondary receptor [164, 165]. The ability of the sigma 1 protein to bind to sialic acid depends on a point mutation (L204P) at the binding site that converts a sialic acid-negative into a sialic acid-positive binding phenotype [164]. Interaction with sialic acid appears to precede binding to JAM-1 and to be necessary for endocytosis of the virus [166].

Rotaviruses infect a wide range of avian and human species and they are the major cause of gastroenteritis in children. Virions possess an outer VP7 layer and large “spikes” or “turrets” at the 12 icosahedral vertices composed of VP4. Trypsin cleaves the C-terminal from the N-terminal domain of VP4, giving rise to VP5 and VP8, respectively, both of which remain associated with the virion. X-Ray crystallography and NMR spectroscopy of VP8 alone and complexed with 2-*O*-methyl- α -*D*-*N*-acetyl neuraminic acid revealed that the VP8 core is a globular domain of an 11-stranded anti-parallel β -sandwich with the sialic acid binding site located in an open-ended, shallow groove [167, 168]. The concept that rotaviruses attach to sialic acid is supported by the observation that binding of some strains to cells is abolished by neuraminidase treatment [169, 170]. In contrast, binding of many other strains is neuraminidase insensitive [171], but it is now clear that these viruses also use sialic acid, yet in a form resistant to neuraminidase treatment [172, 173]. Comparison of the crystal structures of VP8 of neuraminidase-sensitive and neuraminidase-insensitive strains revealed that they were very similar, differing only by the size of the sialic acid binding groove that was slightly wider with the neuraminidase-insensitive strain [174].

The following steps are believed to be involved in the cell entry of rotaviruses: The VP8 domain of VP4 binds first to sialic acid residues of gangliosides or

glycoproteins resulting in a conformational change of VP4 that exposes VP5. The VP5 domain then interacts with $\alpha 2\beta 1$ integrin. Finally, several additional interactions take place, involving VP5, VP7, integrins $\alpha v\beta 3$ and $\alpha x\beta 2$, and probably other cellular proteins [175]. Compatible with this concept is the observation that rotavirus binding to sialic acid is characterized by broad specificity and low affinity, suggesting that it mediates initial cell attachment prior to other interactions that determine host range and cell type specificity [176].

Different gangliosides have been found to be involved in rotavirus entry, and the results of these studies have recently been reviewed in detail [10]. Briefly, porcine rotaviruses have GM3 [177] and GD1a [173] receptors. Simian rotavirus 11 binds to GM3, GM2, and GD1a [178, 179]; GM3 containing both *N*-acetyl- and *N*-glycolyl-neuraminic acid may represent the receptors of bovine rotaviruses [178]. Human rotavirus bound to GM1 [173, 180].

8 Polyomaviridae

Polyomaviruses are DNA-tumor viruses. Most of them have oncogenic potential in rodents and non-human primates, and murine polyomavirus (MPyV) and simian polyomavirus 40 (SV40) have been widely used in experimental oncology. In immunocompromised patients, the human polyomaviruses JCPyV and BKPyV cause progressive multifocal leucoencephalopathy, a fatal demyelinating disease, and nephropathy, respectively. The recently discovered Merkel cell polyomavirus (MCPyV) is the causative agent of an aggressive form of human skin cancer. Polyomaviruses are small non-enveloped viruses containing a double-stranded DNA genome. VP1 is the major viral protein. It forms the outer capsid shell of the icosahedral virions and carries the receptor binding site [181–184].

Paulson and his group were the first to show that MPyV utilizes sialic acid as receptor. Employing reconstituted erythrocytes with defined sialic acid moieties they found that some strains specifically bound to $\alpha 2$ -3-linked sialic acid, whereas others also recognized branched $\alpha 2$ -6-linkages [185–187]. More recently, gangliosides GD1a and GT1b were identified as receptors in sucrose gradient floatation assays [188]. Crystallographic analysis has shown that a shallow groove composed of several loops of VP1 serves as the sialic acid binding site [184, 189]. The structural analysis also showed that the receptor pocket specifically accommodates a Neu5Ac $\alpha 2$ -3-Gal motif unbranched at the Gal position [183] which is compatible with the data obtained in the binding studies employing erythrocytes [186] and gangliosides [188]. MPyV also uses $\alpha 4\beta 1$ integrin as receptor [190] which appears to be mediated by an LDV integrin binding groove deep within VP1. This suggests that, after attachment to sialic acid, the virus has to undergo a conformational change that allows binding to integrin as a second step in the entry process [8]. Evidence has also been obtained that binding to gangliosides promotes virus entry via caveolin-mediated endocytosis [191, 192].

SV40 also binds to gangliosides, but it differs in its receptor specificity from MPyV by showing a specific requirement for GM1 [188]. Crystallographic analysis has revealed that both the Gal β 1-3GalNAc and Neu5Ac branches provide binding activity by directly contacting the protein [182]. Receptor binding of African green monkey lymphotropic papovavirus (LPV), another primate polyomavirus, has been shown to be neuraminidase sensitive, and it has been suggested that the sialic acid necessary for the receptor function is located on a mucin-type glycoprotein or on a ganglioside [193].

Knowledge on the receptors of the human polyomaviruses is less detailed. JCPyV binds to α 2-3- and α 2-6-linked sialic acid [194, 195], and there is some evidence that ganglioside GT1b is involved in the infection of human neuroblastoma cells [196]. Infection of glial cells depends on the serotonin receptor 5HT_{2a}, and this receptor function appears to be neuraminidase sensitive [195, 197]. BkPyV binds only to α 2-3-linked sialic acid [198], and floatation assays have shown that gangliosides GD1b and GT1b serve as receptors [199]. GT1b was also identified as a receptor of MCPyV, and the observation that GD1a and GD1b did not show this function suggests that both the α 2-3-linked and the α 2-8-linked sialic acid of GT1b are required [200].

9 Adenoviridae

This family contains non-enveloped DNA viruses that bind to their receptors via interactions with the distal knob of the penton fibers attached to the vertices of the icosahedral virions. Human adenoviruses mainly cause respiratory and gastrointestinal infections. Several adenoviruses also infect the eye where the most important disease is epidemic keratoconjunctivitis (EKC), caused primarily by Ad8, Ad19, and Ad37. Ad37 binds preferentially to α 2-3-linked sialic acid which is the most frequent type of sialic acid linkage in corneal and conjunctival cells [201]. The crystal structure of the Ad37 knob–sialic acid complex has been elucidated [202].

10 Parvoviridae

This family contains small icosahedral viruses with a single-stranded DNA genome that is encapsidated by a shell composed of two or three proteins. The *Parvoviridae* family is subdivided into two subfamilies (*Parvovirinae* and *Densovirinae*) comprising a total of nine genera, two of which contain viruses that recognize sialic acid receptors. These are the minute virus of mice in the *Parvovirus* genus and some adeno associated viruses (AAVs) in the *Densovirus* genus. AAVs are non-pathogenic agents that depend on adenoviruses for replication. Because of their inability to induce productive infection in the absence of a helper virus, AAVs are promising vectors in gene therapy.

Bovine AAV has been shown to depend on gangliosides for entry [203], and binding to α 2-3-linked sialic acid has been reported for AAV type 5, whereas AAV4 appears to bind to α 2-6-linked sialic acid [204]. It has also been suggested that sialic acid serves not just as an attachment factor but is also required for virus internalization [205]. On the whole, however, the role of sialic acid in the AAV infection process is still poorly understood.

Acknowledgements Our own recent studies were supported by the Deutsche Forschungsgemeinschaft (SFB 587, SFB 593, SFB 621 and SFB 1021), the Bundesministerium fuer Bildung und Forschung (BMBF, FluResearchNet), the Von Behring-Roentgen-Stiftung, the LOEWE Program of the State of Hessen (Universities of Giessen and Marburg Lung Center), the Wellcome Trust grant WT085572MF, and the European Commission FP7 projects FLUPIG and PREDEMICS.

References

1. Hirst GK (1941) The agglutination of red cells by allantoic fluid of chick embryos infected with influenza virus. *Science* 94:22–23
2. McClelland L, Hare R (1941) The adsorption of influenza virus by red cells and a new in vitro method of measuring antibodies for influenza virus. *Can J Public Health* 32:530–538
3. Burnet FM, Stone JD (1947) The receptor-destroying enzyme of *V. cholerae*. *Aust J Exp Biol Med Sci* 25:227–233
4. Gottschalk A, Lind PE (1949) Product of interaction between influenza virus enzyme and ovomucin. *Nature* 164:232
5. Klenk E, Faillard H, Lempfrid H (1955) Über die enzymatische Wirkung von Influenza Virus. *Z physiol Chem* 301:235–246
6. de Groot RJ (2006) Structure, function and evolution of the hemagglutinin-esterase proteins of corona- and toroviruses. *Glycoconj J* 23:59–72
7. Lehmann F, Tiralongo E, Tiralongo J (2006) Sialic acid-specific lectins: occurrence, specificity and function. *Cell Mol Life Sci* 63:1331–1354
8. Neu U, Stehle T, Atwood WJ (2009) The Polyomaviridae: contributions of virus structure to our understanding of virus receptors and infectious entry. *Virology* 384:389–399
9. Olofsson S, Bergstrom T (2005) Glycoconjugate glycans as viral receptors. *Ann Med* 37:154–172
10. Taube S, Jiang M, Wobus CE (2010) Glycosphingolipids as receptors for non-enveloped viruses. *Viruses* 2:1011–1049
11. McCauley JW, Hongo S, Kaverin NV, Kochs G, Lamb RA, Matrosovich MN, Perez DR, Palese P, Presti RM, Rimstadt E et al. (2011) Orthomyxoviridae. In: King AMQ, Adams MJ, Carstens EB, Lefkowitz EJ (eds) *Virus taxonomy*. Elsevier, Oxford, pp 749–762
12. Palese P, Shaw ML (2007) Orthomyxoviridae: the viruses and their replication. In: Knipe DM, Howley PM (eds) *Fields virology*. Lippincott, Williams and Wilkins, Philadelphia, pp 1647–1689
13. Tong S, Li Y, Rivailler P, Conrardy C, Castillo DA, Chen LM, Recuenco S, Ellison JA, Davis CT, York IA et al. (2012) A distinct lineage of influenza A virus from bats. *Proc Natl Acad Sci U S A* 109:4269–4274
14. Klenk HD (2011) Influenza virology. In: von Itzstein M (ed) *Influenza virus sialidase – a drug discovery target*. Springer, Basel, pp 1–29
15. Gamblin SJ, Skehel JJ (2010) Influenza hemagglutinin and neuraminidase membrane glycoproteins. *J Biol Chem* 285:28403–28409

16. Skehel JJ, Wiley DC (2000) Receptor binding and membrane fusion in virus entry: the influenza hemagglutinin. *Annu Rev Biochem* 69:531–569
17. Matrosovich MN, Klenk H-D, Kawaoka Y (2006) Receptor specificity, host range and pathogenicity of influenza viruses. In: Kawaoka Y (ed) *Influenza virology: current topics*. Caister Academic, Wymondham, pp 95–137
18. Connor RJ, Kawaoka Y, Webster RG, Paulson JC (1994) Receptor specificity in human, avian, and equine H2 and H3 influenza virus isolates. *Virology* 205:17–23
19. Matrosovich M, Tuzikov A, Bovin N, Gambarian A, Klimov A, Cox N, Castrucci M, Donatelli I, Kawaoka Y (2000) Alterations of receptor-binding properties of H1, H2 and H3 avian influenza virus hemagglutinins upon introduction into mammals. *J Virol* 74:8502–8512
20. Rogers GN, Paulson JC (1983) Receptor determinants of human and animal influenza virus isolates: differences in receptor specificity of the H3 hemagglutinin based on species of origin. *Virology* 127:361–373
21. Baum LG, Paulson JC (1990) Sialyloligosaccharides of the respiratory epithelium in the selection of human influenza virus receptor specificity. *Acta Histochem Suppl* 40:35–38
22. Gambaryan A, Webster R, Matrosovich M (2002) Differences between influenza virus receptors on target cells of duck and chicken. *Arch Virol* 147:1197–1208
23. Ito T, Couceiro JN, Kelm S, Baum LG, Krauss S, Castrucci MR, Donatelli I, Kida H, Paulson JC, Webster RG et al. (1998) Molecular basis for the generation in pigs of influenza A viruses with pandemic potential. *J Virol* 72:7367–7373
24. Kuchipudi SV, Nelli R, White GA, Bain M, Chang KC, Dunham S (2009) Differences in influenza virus receptors in chickens and ducks: implications for interspecies transmission. *J Mol Genet Med* 3:143–151
25. Pillai SP, Lee CW (2010) Species and age related differences in the type and distribution of influenza virus receptors in different tissues of chickens, ducks and turkeys. *Virol J* 7:5
26. Shinya K, Ebina M, Yamada S, Ono M, Kasai N, Kawaoka Y (2006) Avian flu: influenza virus receptors in the human airway. *Nature* 440:435–436
27. Matrosovich M, Stech J, Klenk HD (2009) Influenza receptors, polymerase and host range. *Rev Sci Tech* 28:203–217
28. Matrosovich MN, Gambarian AS, Klenk HD (2008) Receptor specificity of influenza viruses and its alteration during interspecies transmission. In: Klenk HD, Matrosovich MN, Stech J (eds) *Avian Influenza*. Karger, Basel, pp 134–155
29. Gambaryan A, Tuzikov A, Pazynina G, Bovin N, Balish A, Klimov A (2006) Evolution of the receptor binding phenotype of influenza A (H5) viruses. *Virology* 344:432–438
30. Matrosovich M, Zhou N, Kawaoka Y, Webster R (1999) The surface glycoproteins of H5 influenza viruses isolated from humans, chickens, and wild aquatic birds have distinguishable properties. *J Virol* 73:1146–1155
31. Stevens J, Blixt O, Tumpey TM, Taubenberger JK, Paulson JC, Wilson IA (2006) Structure and receptor specificity of the hemagglutinin from an H5N1 influenza virus. *Science* 312:404–410
32. Matrosovich MN, Matrosovich TY, Gray T, Roberts NA, Klenk HD (2004) Human and avian influenza viruses target different cell types in cultures of human airway epithelium. *Proc Natl Acad Sci U S A* 101:4620–4624
33. Nicholls JM, Chan MC, Chan WY, Wong HK, Cheung CY, Kwong DL, Wong MP, Chui WH, Poon LL, Tsao SW et al. (2007) Tropism of avian influenza A (H5N1) in the upper and lower respiratory tract. *Nat Med* 13:147–149
34. van Riel D, Munster VJ, de Wit E, Rimmelzwaan GF, Fouchier RA, Osterhaus AD, Kuiken T (2006) H5N1 virus attachment to lower respiratory tract. *Science* 312:399
35. van Riel D, Munster VJ, de Wit E, Rimmelzwaan GF, Fouchier RA, Osterhaus AD, Kuiken T (2007) Human and avian influenza viruses target different cells in the lower respiratory tract of humans and other mammals. *Am J Pathol* 171:1215–1223

36. Yao L, Korteweg C, Hsueh W, Gu J (2007) Avian influenza receptor expression in H5N1-infected and noninfected human tissues. *FASEB J*. doi:[10.1096/fj.1006-7880com](https://doi.org/10.1096/fj.1006-7880com)
37. Herfst S, Schrauwen EJ, Linster M, Chutinimitkul S, de Wit E, Munster VJ, Sorrell EM, Bestebroer TM, Burke DF, Smith DJ et al. (2012) Airborne transmission of influenza A/H5N1 virus between ferrets. *Science* 336:1534–1541
38. Imai M, Watanabe T, Hatta M, Das SC, Ozawa M, Shinya K, Zhong G, Hanson A, Katsura H, Watanabe S et al. (2012) Experimental adaptation of an influenza H5 HA confers respiratory droplet transmission to a reassortant H5 HA/H1N1 virus in ferrets. *Nature* 486:420–428
39. Scholtissek C, Hinshaw VS, Olsen CW (1998) Influenza in pigs and their role as the intermediate host. In: Nicholson KG, Webster RG, Hay A (eds) *Textbook of influenza*. Blackwell Science, London, pp 137–145
40. Gambaryan AS, Karasin AI, Tuzikov AB, Chinarev AA, Pazynina GV, Bovin NV, Matrosovich MN, Olsen CW, Klimov AI (2005) Receptor-binding properties of swine influenza viruses isolated and propagated in MDCK cells. *Virus Res* 114:15–22
41. Van Poucke SG, Nicholls JM, Nauwynck HJ, Van Reeth K (2010) Replication of avian, human and swine influenza viruses in porcine respiratory explants and association with sialic acid distribution. *Virology* 413:169–182
42. Bradley KC, Jones CA, Tompkins SM, Tripp RA, Russell RJ, Gramer MR, Heimburg-Molinaro J, Smith DF, Cummings RD, Steinhauer DA (2011) Comparison of the receptor binding properties of contemporary swine isolates and early human pandemic H1N1 isolates (Novel 2009 H1N1). *Virology* 413:169–182
43. Chen LM, Rivaille P, Hossain J, Carney P, Balish A, Perry I, Davis CT, Garten R, Shu B, Xu X et al. (2011) Receptor specificity of subtype H1 influenza A viruses isolated from swine and humans in the United States. *Virology* 412:401–410
44. Maines TR, Jayaraman A, Belser JA, Wadford DA, Pappas C, Zeng H, Gustin KM, Pearce MB, Viswanathan K, Shriver ZH et al. (2009) Transmission and pathogenesis of swine-origin 2009 A(H1N1) influenza viruses in ferrets and mice. *Science* 325:484–487
45. Yang H, Carney P, Stevens J (2010) Structure and receptor binding properties of a pandemic H1N1 virus hemagglutinin. *PLoS Curr* 2. doi:[10.1371/currents.RRN1152](https://doi.org/10.1371/currents.RRN1152)
46. Childs RA, Palma AS, Wharton S, Matrosovich T, Liu Y, Chai W, Campanero-Rhodes MA, Zhang Y, Eickmann M, Kiso M et al. (2009) Receptor-binding specificity of pandemic influenza A (H1N1) 2009 virus determined by carbohydrate microarray. *Nat Biotechnol* 27:797–799
47. Kilander A, Rykkvin R, Dudman SG, Hungnes O (2010) Observed association between the HA1 mutation D222G in the 2009 pandemic influenza A(H1N1) virus and severe clinical outcome, Norway 2009–2010. *Euro Surveill* 15(9):pii=19498
48. Puzelli S, Facchini M, De Marco MA, Palmieri A, Spagnolo D, Boros S, Corcioli F, Trotta D, Bagnarelli P, Azzi A, Cassone A, Rezza G, Pompa MG, Oleari F, Donatelli I, the Influenza Surveillance Group for Pandemic A(H1N1) 2009 Influenza Virus in Italy (2010) Molecular surveillance of pandemic influenza A(H1N1) viruses circulating in Italy from May 2009 to February 2010: association between haemagglutinin mutations and clinical outcome. *Euro Surveill* 15(43):pii=19696
49. Liu Y, Childs RA, Matrosovich T, Wharton S, Palma AS, Chai W, Daniels R, Gregory V, Uhlenendorff J, Kiso M et al. (2010) Altered receptor specificity and cell tropism of D222G hemagglutinin mutants isolated from fatal cases of pandemic A(H1N1) 2009 influenza virus. *J Virol* 84:12069–12074
50. Chutinimitkul S, Herfst S, Steel J, Lowen AC, Ye J, van Riel D, Schrauwen EJ, Bestebroer TM, Koel B, Burke DF et al. (2010) Virulence-associated substitution D222G in the hemagglutinin of 2009 pandemic influenza A(H1N1) virus affects receptor binding. *J Virol* 84:11802–11813
51. Matrosovich MN, Gambaryan AS, Teneberg S, Piskarev VE, Yamnikova SS, Lvov DK, Robertson JS, Karlsson KA (1997) Avian influenza A viruses differ from human viruses by

- recognition of sialyloligosaccharides and gangliosides and by a higher conservation of the HA receptor-binding site. *Virology* 233:224–234
52. Nicholls JM, Chan RW, Russell RJ, Air GM, Peiris JS (2008) Evolving complexities of influenza virus and its receptors. *Trends Microbiol* 16:149–157
 53. Varghese JN, Laver WG, Colman PM (1983) Structure of the influenza virus glycoprotein antigen neuraminidase at 2.9 Å resolution. *Nature* 303:35–40
 54. Blok J, Air GM (1982) Block deletions in the neuraminidase genes from some influenza A viruses of the N1 subtype. *Virology* 118:229–234
 55. Matrosovich MN, Matrosovich TY, Gray T, Roberts NA, Klenk HD (2004) Neuraminidase is important for the initiation of influenza virus infection in human airway epithelium. *J Virol* 78:12665–12667
 56. Palese P, Tobita K, Ueda M, Compans RW (1974) Characterization of temperature sensitive influenza virus mutants defective in neuraminidase. *Virology* 61:397–410
 57. Banks J, Speidel ES, Moore E, Plowright L, Piccirillo A, Capua I, Cordioli P, Fioretti A, Alexander DJ (2001) Changes in the haemagglutinin and the neuraminidase genes prior to the emergence of highly pathogenic H7N1 avian influenza viruses in Italy. *Arch Virol* 146:963–973
 58. Kaverin NV, Gambaryan AS, Bovin NV, Rudneva IA, Shilov AA, Khodova OM, Varich NL, Sinitsin BV, Makarova NV, Kropotkina EA (1998) Postreassortment changes in influenza A virus hemagglutinin restoring HA-NA functional match. *Virology* 244:315–321
 59. Mitnaul LJ, Matrosovich MN, Castrucci MR, Tuzikov AB, Bovin NV, Kobasa D, Kawaoka Y (2000) Balanced hemagglutinin and neuraminidase activities are critical for efficient replication of influenza A virus. *J Virol* 74:6015–6020
 60. Wagner R, Matrosovich M, Klenk HD (2002) Functional balance between haemagglutinin and neuraminidase in influenza virus infections. *Rev Med Virol* 12:159–166
 61. Wagner R, Wolff T, Herwig A, Pleschka S, Klenk HD (2000) Interdependence of hemagglutinin glycosylation and neuraminidase as regulators of influenza virus growth: a study by reverse genetics. *J Virol* 74:6316–6323
 62. Hausmann J, Kretzschmar E, Garten W, Klenk HD (1995) N1 neuraminidase of influenza virus A/FPV/Rostock/34 has haemadsorbing activity. *J Gen Virol* 76(Pt 7):1719–1728
 63. Kobasa D, Rodgers ME, Wells K, Kawaoka Y (1997) Neuraminidase hemadsorption activity, conserved in avian influenza A viruses, does not influence viral replication in ducks. *J Virol* 71:6706–6713
 64. Laver WG, Colman PM, Webster RG, Hinshaw VS, Air GM (1984) Influenza virus neuraminidase with hemagglutinin activity. *Virology* 137:314–323
 65. Varghese JN, Colman PM, van Donkelaar A, Blick TJ, Sahasrabudhe A, McKimm-Breschkin JL (1997) Structural evidence for a second sialic acid binding site in avian influenza virus neuraminidases. *Proc Natl Acad Sci U S A* 94:11808–11812
 66. Uhlenndorff J, Matrosovich T, Klenk HD, Matrosovich M (2009) Functional significance of the hemadsorption activity of influenza virus neuraminidase and its alteration in pandemic viruses. *Arch Virol* 154:945–957
 67. Herrler G, Hausmann J, Klenk HD (1995) Sialic acid as receptor determinant of ortho- and paramyxoviruses. In: Rosenberg A (ed) *Biology of the Sialic acids*. Plenum, New York, pp 315–336
 68. Herrler G, Klenk HD (1991) Structure and function of the HEF glycoprotein of influenza C virus. *Adv Virus Res* 40:213–234
 69. Rosenthal PB, Zhang X, Formanowski F, Fitz W, Wong CH, Meier-Ewert H, Skehel JJ, Wiley DC (1998) Structure of the haemagglutinin-esterase-fusion glycoprotein of influenza C virus. *Nature* 396:92–96
 70. Herrler G, Rott R, Klenk HD, Muller HP, Shukla AK, Schauer R (1985) The receptor-destroying enzyme of influenza C virus is neuraminidase-O-acetyltransferase. *EMBO J* 4:1503–1506

71. Herrler G, Multhaup G, Beyreuther K, Klenk HD (1988) Serine 71 of the glycoprotein HEF is located at the active site of the acetyltransferase of influenza C virus. *Arch Virol* 102:269–274
72. Pleschka S, Klenk HD, Herrler G (1995) The catalytic triad of the influenza C virus glycoprotein HEF esterase: characterization by site-directed mutagenesis and functional analysis. *J Gen Virol* 76:2529–2537
73. Vlasak R, Muster T, Lauro AM, Powers JC, Palese P (1989) Influenza C virus esterase: analysis of catalytic site, inhibition, and possible function. *J Virol* 63:2056–2062
74. Hofling K, Klenk HD, Herrler G (1997) Inactivation of inhibitors by the receptor-destroying enzyme of influenza C virus. *J Gen Virol* 78:567–570
75. Hofling K, Brossmer R, Klenk H, Herrler G (1996) Transfer of an esterase-resistant receptor analog to the surface of influenza C virions results in reduced infectivity due to aggregate formation. *Virology* 218:127–133
76. Herrler G, Gross HJ, Brossmer R (1995) A synthetic sialic acid analog that is resistant to the receptor-destroying enzyme can be used by influenza C virus as a receptor determinant for infection of cells. *Biochem Biophys Res Commun* 216:821–827
77. Herrler G, Gross HJ, Imhof A, Brossmer R, Milks G, Paulson JC (1992) A synthetic sialic acid analogue is recognized by influenza C virus as a receptor determinant but is resistant to the receptor-destroying enzyme. *J Biol Chem* 267:12501–12505
78. Rogers GN, Herrler G, Paulson JC, Klenk HD (1986) Influenza C virus uses 9-*O*-acetyl-*N*-acetylneuraminic acid as a high affinity receptor determinant for attachment to cells. *J Biol Chem* 261:5947–5951
79. Herrler G, Klenk HD (1987) The surface receptor is a major determinant of the cell tropism of influenza C virus. *Virology* 159:102–108
80. Zimmer G, Klenk HD, Herrler G (1995) Identification of a 40-kDa cell surface sialoglycoprotein with the characteristics of a major influenza C virus receptor in a Madin–Darby canine kidney cell line. *J Biol Chem* 270:17815–17822
81. Zimmer G, Lottspeich F, Maisner A, Klenk HD, Herrler G (1997) Molecular characterization of gp40, a mucin-type glycoprotein from the apical plasma membrane of Madin–Darby canine kidney cells (type I). *Biochem J* 326:99–108
82. Marschall M, Herrler G, Boswald C, Foerst G, Meier-Ewert H (1994) Persistent influenza C virus possesses distinct functional properties due to a modified HEF glycoprotein. *J Gen Virol* 75:2189–2196
83. Matsuzaki M, Sugawara K, Adachi K, Hongo S, Nishimura H, Kitame F, Nakamura K (1992) Location of neutralizing epitopes on the hemagglutinin-esterase protein of influenza C virus. *Virology* 189:79–87
84. Szepanski S, Gross HJ, Brossmer R, Klenk HD, Herrler G (1992) A single point mutation of the influenza C virus glycoprotein (HEF) changes the viral receptor-binding activity. *Virology* 188:85–92
85. Umetsu Y, Sugawara K, Nishimura H, Hongo S, Matsuzaki M, Kitame F, Nakamura K (1992) Selection of antigenically distinct variants of influenza C viruses by the host cell. *Virology* 189:740–744
86. Falk K, Namork E, Rimstad E, Mjaaland S, Dannevig BH (1997) Characterization of an infectious salmon anemia virus, an orthomyxo-like virus isolated from Atlantic salmon (*Salmo salar* L.). *J Virol* 71:9016–9023
87. Kristiansen M, Froystad MK, Rishovd AL, Gjoen T (2002) Characterization of the receptor-destroying enzyme activity from infectious salmon anaemia virus. *J Gen Virol* 83:2693–2697
88. Hellebo A, Vilas U, Falk K, Vlasak R (2004) Infectious salmon anemia virus specifically binds to and hydrolyzes 4-*O*-acetylated sialic acids. *J Virol* 78:3055–3062
89. Falk K, Aspehaug V, Vlasak R, Endresen C (2004) Identification and characterization of viral structural proteins of infectious salmon anemia virus. *J Virol* 78:3063–3071
90. Krossoy B, Devold M, Sanders L, Knappskog PM, Aspehaug V, Falk K, Nylund A, Koumans S, Endresen C, Biering E (2001) Cloning and identification of the infectious salmon anaemia virus haemagglutinin. *J Gen Virol* 82:1757–1765

91. Rimstad E, Mjaaland S, Snow M, Mikalsen AB, Cunningham CO (2001) Characterization of the infectious salmon anemia virus genomic segment that encodes the putative hemagglutinin. *J Virol* 75:5352–5356
92. Vlasak R, Luytjes W, Spaan W, Palese P (1988) Human and bovine coronaviruses recognize sialic acid-containing receptors similar to those of influenza C viruses. *Proc Natl Acad Sci U S A* 85:4526–4529
93. Schultze B, Wahn K, Klenk HD, Herrler G (1991) Isolated HE-protein from hemagglutinating encephalomyelitis virus and bovine coronavirus has receptor-destroying and receptor-binding activity. *Virology* 180:221–228
94. Vlasak R, Luytjes W, Leider J, Spaan W, Palese P (1988) The E3 protein of bovine coronavirus is a receptor-destroying enzyme with acetylase activity. *J Virol* 62:4686–4690
95. Yokomori K, La Monica N, Makino S, Shieh CK, Lai MM (1989) Biosynthesis, structure, and biological activities of envelope protein gp65 of murine coronavirus. *Virology* 173:683–691
96. Zeng Q, Langereis MA, van Vliet AL, Huizinga EG, de Groot RJ (2008) Structure of coronavirus hemagglutinin-esterase offers insight into corona and influenza virus evolution. *Proc Natl Acad Sci U S A* 105:9065–9069
97. Sugiyama K, Kasai M, Kato S, Kasai H, Hatakeyama K (1998) Haemagglutinin-esterase protein (HE) of murine corona virus: DVIM (diarrhea virus of infant mice). *Arch Virol* 143:1523–1534
98. Klausegger A, Strobl B, Regl G, Kaser A, Luytjes W, Vlasak R (1999) Identification of a coronavirus hemagglutinin-esterase with a substrate specificity different from those of influenza C virus and bovine coronavirus. *J Virol* 73:3737–3743
99. Regl G, Kaser A, Iwersen M, Schmid H, Kohla G, Strobl B, Vilas U, Schauer R, Vlasak R (1999) The hemagglutinin-esterase of mouse hepatitis virus strain S is a sialate-4-*O*-acetylase. *J Virol* 73:4721–4727
100. Smits SL, Gerwig GJ, van Vliet AL, Lissenberg A, Briza P, Kamerling JP, Vlasak R, de Groot RJ (2005) Nidovirus sialate-*O*-acetylases: evolution and substrate specificity of coronaviral and toroviral receptor-destroying enzymes. *J Biol Chem* 280:6933–6941
101. Storz J, Zhang XM, Rott R (1992) Comparison of hemagglutinating, receptor-destroying, and acetylase activities of avirulent and virulent bovine coronavirus strains. *Arch Virol* 125:193–204
102. Schultze B, Herrler G (1992) Bovine coronavirus uses *N*-acetyl-9-*O*-acetylneuraminic acid as a receptor determinant to initiate the infection of cultured cells. *J Gen Virol* 73:901–906
103. Lin X, O'Reilly KL, Storz J (1997) Infection of polarized epithelial cells with enteric and respiratory tract bovine coronaviruses and release of virus progeny. *Am J Vet Res* 58:1120–1124
104. Schultze B, Krempel C, Ballesteros ML, Shaw L, Schauer R, Enjuanes L, Herrler G (1996) Transmissible gastroenteritis coronavirus, but not the related porcine respiratory coronavirus, has a sialic acid (*N*-glycolylneuraminic acid) binding activity. *J Virol* 70:5634–5637
105. Schultze B, Zimmer G, Herrler G (1996) Virus entry into a polarized epithelial cell line (MDCK): similarities and dissimilarities between influenza C virus and bovine coronavirus. *J Gen Virol* 77(Pt 10):2507–2514
106. King B, Potts BJ, Brian DA (1985) Bovine coronavirus hemagglutinin protein. *Virus Res* 2:53–59
107. Yoo D, Graham FL, Prevec L, Parker MD, Benko M, Zamb T, Babiuk LA (1992) Synthesis and processing of the haemagglutinin-esterase glycoprotein of bovine coronavirus encoded in the E3 region of adenovirus. *J Gen Virol* 73:2591–2600
108. Schultze B, Gross HJ, Brossmer R, Herrler G (1991) The S protein of bovine coronavirus is a hemagglutinin recognizing 9-*O*-acetylated sialic acid as a receptor determinant. *J Virol* 65:6232–6237

109. Gagneten S, Gout O, Dubois-Dalcq M, Rottier P, Rossen J, Holmes KV (1995) Interaction of mouse hepatitis virus (MHV) spike glycoprotein with receptor glycoprotein MHVR is required for infection with an MHV strain that expresses the hemagglutinin-esterase glycoprotein. *J Virol* 69:889–895
110. Langereis MA, van Vliet AL, Boot W, de Groot RJ (2010) Attachment of mouse hepatitis virus to *O*-acetylated sialic acid is mediated by hemagglutinin-esterase and not by the spike protein. *J Virol* 84:8970–8974
111. Noda M, Koide F, Asagi M, Inaba Y (1988) Physicochemical properties of transmissible gastroenteritis virus hemagglutinin. *Arch Virol* 99:163–172
112. Noda M, Yamashita H, Koide F, Kadoi K, Omori T, Asagi M, Inaba Y (1987) Hemagglutination with transmissible gastroenteritis virus. *Arch Virol* 96:109–115
113. Kreml C, Schultze B, Laude H, Herrler G (1997) Point mutations in the S protein connect the sialic acid binding activity with the enteropathogenicity of transmissible gastroenteritis coronavirus. *J Virol* 71:3285–3287
114. Bernard S, Laude H (1995) Site-specific alteration of transmissible gastroenteritis virus spike protein results in markedly reduced pathogenicity. *J Gen Virol* 76:2235–2241
115. Kreml C, Ballesteros ML, Zimmer G, Enjuanes L, Klenk HD, Herrler G (2000) Characterization of the sialic acid binding activity of transmissible gastroenteritis coronavirus by analysis of haemagglutination-deficient mutants. *J Gen Virol* 81:489–496
116. Schwegmann-Wessels C, Zimmer G, Schroder B, Breves G, Herrler G (2003) Binding of transmissible gastroenteritis coronavirus to brush border membrane sialoglycoproteins. *J Virol* 77:11846–11848
117. Pensaert M, Callebaut P, Vergote J (1986) Isolation of a porcine respiratory, non-enteric coronavirus related to transmissible gastroenteritis. *Vet Q* 8:257–261
118. Delmas B, Gelfi J, L'Haridon R, Vogel LK, Sjostrom H, Noren O, Laude H (1992) Amino-peptidase N is a major receptor for the entero-pathogenic coronavirus TGEV. *Nature* 357:417–420
119. Cox E, Pensaert MB, Callebaut P, van Deun K (1990) Intestinal replication of a porcine respiratory coronavirus closely related antigenically to the enteric transmissible gastroenteritis virus. *Vet Microbiol* 23:237–243
120. Schwegmann-Wessels C, Zimmer G, Laude H, Enjuanes L, Herrler G (2002) Binding of transmissible gastroenteritis coronavirus to cell surface sialoglycoproteins. *J Virol* 76:6037–6043
121. Schwegmann-Wessels C, Bauer S, Winter C, Enjuanes L, Laude H, Herrler G (2011) The sialic acid binding activity of the S protein facilitates infection by porcine transmissible gastroenteritis coronavirus. *Virology* 418:435
122. Bingham RW, Madge MH, Tyrrell DA (1975) Haemagglutination by avian infectious bronchitis virus – a coronavirus. *J Gen Virol* 28:381–390
123. Schultze B, Cavanagh D, Herrler G (1992) Neuraminidase treatment of avian infectious bronchitis coronavirus reveals a hemagglutinating activity that is dependent on sialic acid-containing receptors on erythrocytes. *Virology* 189:792–794
124. Winter C, Schwegmann-Wessels C, Cavanagh D, Neumann U, Herrler G (2006) Sialic acid is a receptor determinant for infection of cells by avian infectious bronchitis virus. *J Gen Virol* 87:1209–1216
125. Abd El Rahman S, El-Kenawy AA, Neumann U, Herrler G, Winter C (2009) Comparative analysis of the sialic acid binding activity and the tropism for the respiratory epithelium of four different strains of avian infectious bronchitis virus. *Avian Pathol* 38:41–45
126. Abd El Rahman S, Winter C, El-Kenawy A, Neumann U, Herrler G (2010) Differential sensitivity of well-differentiated avian respiratory epithelial cells to infection by different strains of infectious bronchitis virus. *J Virol* 84:8949–8952
127. Winter C, Herrler G, Neumann U (2008) Infection of the tracheal epithelium by infectious bronchitis virus is sialic acid dependent. *Microbes Infect* 10:367–373

128. Cornelissen LA, Wierda CM, van der Meer FJ, Herrewegh AA, Horzinek MC, Egberink HF, de Groot RJ (1997) Hemagglutinin-esterase, a novel structural protein of torovirus. *J Virol* 71:5277–5286
129. Duckmanton L, Tellier R, Richardson C, Petric M (1999) The novel hemagglutinin-esterase genes of human torovirus and Breda virus. *Virus Res* 64:137–149
130. Langereis MA, Zeng Q, Gerwig GJ, Frey B, von Itzstein M, Kamerling JP, de Groot RJ, Huizinga EG (2009) Structural basis for ligand and substrate recognition by torovirus hemagglutinin esterases. *Proc Natl Acad Sci U S A* 106:15897–15902
131. Lamb RA, Parks GD (2007) Paramyxoviridae: the viruses and their replication. In: Knipe DM, Howley PM (eds) *Fields Virology*. Lippincott, Williams and Wilkins, Philadelphia, pp 1449–1496
132. Zaitsev V, von Itzstein M, Groves D, Kiefel M, Takimoto T, Portner A, Taylor G (2004) Second sialic acid binding site in Newcastle disease virus hemagglutinin-neuraminidase: implications for fusion. *J Virol* 78:3733–3741
133. Lawrence MC, Borg NA, Streltsov VA, Pilling PA, Epa VC, Varghese JN, McKimm-Breschkin JL, Colman PM (2004) Structure of the haemagglutinin-neuraminidase from human parainfluenza virus type III. *J Mol Biol* 335:1343–1357
134. Yuan P, Thompson TB, Wurzburg BA, Paterson RG, Lamb RA, Jardetzky TS (2005) Structural studies of the parainfluenza virus 5 hemagglutinin-neuraminidase tetramer in complex with its receptor, sialyllactose. *Structure* 13:803–815
135. Alymova IV, Taylor G, Mishin VP, Watanabe M, Murti KG, Boyd K, Chand P, Babu YS, Portner A (2008) Loss of the N-linked glycan at residue 173 of human parainfluenza virus type 1 hemagglutinin-neuraminidase exposes a second receptor-binding site. *J Virol* 82:8400–8410
136. Holmgren J, Svennerholm L, Elwing H, Fredman P, Strannegard O (1980) Sendai virus receptor: proposed recognition structure based on binding to plastic-adsorbed gangliosides. *Proc Natl Acad Sci U S A* 77:1947–1950
137. Markwell MA, Svennerholm L, Paulson JC (1981) Specific gangliosides function as host cell receptors for Sendai virus. *Proc Natl Acad Sci U S A* 78:5406–5410
138. Markwell MA, Paulson JC (1980) Sendai virus utilizes specific sialyloligosaccharides as host cell receptor determinants. *Proc Natl Acad Sci U S A* 77:5693–5697
139. Suzuki Y, Suzuki T, Matsumoto M (1983) Isolation and characterization of receptor sialoglycoprotein for hemagglutinating virus of Japan (Sendai virus) from bovine erythrocyte membrane. *J Biochem* 93:1621–1633
140. Suzuki Y, Suzuki T, Matsunaga M, Matsumoto M (1985) Gangliosides as paramyxovirus receptor. Structural requirement of sialo-oligosaccharides in receptors for hemagglutinating virus of Japan (Sendai virus) and Newcastle disease virus. *J Biochem* 97:1189–1199
141. Suzuki T, Portner A, Scroggs RA, Uchikawa M, Koyama N, Matsuo K, Suzuki Y, Takimoto T (2001) Receptor specificities of human respiroviruses. *J Virol* 75:4604–4613
142. Prasad BV, Hardy ME, Dokland T, Bella J, Rossmann MG, Estes MK (1999) X-Ray crystallographic structure of the Norwalk virus capsid. *Science* 286:287–290
143. Prasad BV, Rothnagel R, Jiang X, Estes MK (1994) Three-dimensional structure of baculovirus-expressed Norwalk virus capsids. *J Virol* 68:5117–5125
144. Bu W, Mamedova A, Tan M, Xia M, Jiang X, Hegde RS (2008) Structural basis for the receptor binding specificity of Norwalk virus. *J Virol* 82:5340–5347
145. Cao S, Lou Z, Tan M, Chen Y, Liu Y, Zhang Z, Zhang XC, Jiang X, Li X, Rao Z (2007) Structural basis for the recognition of blood group trisaccharides by norovirus. *J Virol* 81:5949–5957
146. Choi JM, Hutson AM, Estes MK, Prasad BV (2008) Atomic resolution structural characterization of histo-blood group antigens by Norwalk virus. *Proc Natl Acad Sci U S A* 105:9175–9180
147. Katpally U, Voss NR, Cavazza T, Taube S, Rubin JR, Young VL, Stuckey J, Ward VK, Virgin HWT, Wobus CE et al. (2010) High-resolution cryo-electron microscopy structures of

- murine norovirus 1 and rabbit hemorrhagic disease virus reveal marked flexibility in the receptor binding domains. *J Virol* 84:5836–5841
148. Taube S, Rubin JR, Katpally U, Smith TJ, Kendall A, Stuckey JA, Wobus CE (2010) High-resolution X-ray structure and functional analysis of the murine norovirus 1 capsid protein protruding domain. *J Virol* 84:5695–5705
 149. Wobus CE, Thackray LB, Virgin HWT (2006) Murine norovirus: a model system to study norovirus biology and pathogenesis. *J Virol* 80:5104–5112
 150. Estes MK, Prasad BV, Atmar RL (2006) Noroviruses everywhere: has something changed? *Curr Opin Infect Dis* 19:467–474
 151. Le Pendu J, Ruvoen-Clouet N, Kindberg E, Svensson L (2006) Mendelian resistance to human norovirus infections. *Semin Immunol* 18:375–386
 152. Tan M, Jiang X (2007) Norovirus-host interaction: implications for disease control and prevention. *Expert Rev Mol Med* 9:1–22
 153. Tamura M, Natori K, Kobayashi M, Miyamura T, Takeda N (2004) Genogroup II noroviruses efficiently bind to heparan sulfate proteoglycan associated with the cellular membrane. *J Virol* 78:3817–3826
 154. Rydell GE, Nilsson J, Rodriguez-Diaz J, Ruvoen-Clouet N, Svensson L, Le Pendu J, Larson G (2009) Human noroviruses recognize sialyl Lewis X neoglycoprotein. *Glycobiology* 19:309–320
 155. Taube S, Perry JW, Yetming K, Patel SP, Auble H, Shu L, Nawar HF, Lee CH, Connell TD, Shayman JA et al. (2009) Ganglioside-linked terminal sialic acid moieties on murine macrophages function as attachment receptors for murine noroviruses. *J Virol* 83:4092–4101
 156. Stuart AD, Brown TD (2007) Alpha2,6-linked sialic acid acts as a receptor for Feline calicivirus. *J Gen Virol* 88:177–186
 157. Tavakkol A, Burness AT (1990) Evidence for a direct role for sialic acid in the attachment of encephalomyocarditis virus to human erythrocytes. *Biochemistry* 29:10684–10690
 158. Zhou L, Luo Y, Wu Y, Tsao J, Luo M (2000) Sialylation of the host receptor may modulate entry of demyelinating persistent Theiler's virus. *J Virol* 74:1477–1485
 159. Anderson K, Bond CW (1987) Biological properties of mengovirus: characterization of avirulent, hemagglutination-defective mutants. *Arch Virol* 93:31–49
 160. Stoner GD, Williams B, Kniazeff A, Shimkin MB (1973) Effect of neuraminidase pretreatment on the susceptibility of normal and transformed mammalian cells to bovine enterovirus 261. *Nature* 245:319–320
 161. Nokhbeh MR, Hazra S, Alexander DA, Khan A, McAllister M, Suuronen EJ, Griffith M, Dimock K (2005) Enterovirus 70 binds to different glycoconjugates containing alpha2,3-linked sialic acid on different cell lines. *J Virol* 79:7087–7094
 162. Lipton HL, Kumar AS, Trottier M (2005) Theiler's virus persistence in the central nervous system of mice is associated with continuous viral replication and a difference in outcome of infection of infiltrating macrophages versus oligodendrocytes. *Virus Res* 111:214–223
 163. Helander A, Silvey KJ, Mantis NJ, Hutchings AB, Chandran K, Lucas WT, Nibert ML, Neutra MR (2003) The viral sigma1 protein and glycoconjugates containing alpha2-3-linked sialic acid are involved in type 1 reovirus adherence to M cell apical surfaces. *J Virol* 77:7964–7977
 164. Barton ES, Connolly JL, Forrest JC, Chappell JD, Dermody TS (2001) Utilization of sialic acid as a coreceptor enhances reovirus attachment by multistep adhesion strengthening. *J Biol Chem* 276:2200–2211
 165. Prota AE, Campbell JA, Schelling P, Forrest JC, Watson MJ, Peters TR, Aurrand-Lions M, Imhof BA, Dermody TS, Stehle T (2003) Crystal structure of human junctional adhesion molecule 1: implications for reovirus binding. *Proc Natl Acad Sci U S A* 100:5366–5371
 166. Connolly JL, Barton ES, Dermody TS (2001) Reovirus binding to cell surface sialic acid potentiates virus-induced apoptosis. *J Virol* 75:4029–4039

167. Dormitzer PR, Sun ZY, Wagner G, Harrison SC (2002) The rhesus rotavirus VP4 sialic acid binding domain has a galectin fold with a novel carbohydrate binding site. *EMBO J* 21:885–897
168. Kraschnefski MJ, Bugarcic A, Fleming FE, Yu X, von Itzstein M, Coulson BS, Blanchard H (2009) Effects on sialic acid recognition of amino acid mutations in the carbohydrate-binding cleft of the rotavirus spike protein. *Glycobiology* 19:194–200
169. Bastardo JW, Holmes IH (1980) Attachment of SA-11 rotavirus to erythrocyte receptors. *Infect Immun* 29:1134–1140
170. Spence L, Fauvel M, Petro R, Bloch S (1976) Haemagglutinin from rotavirus. *Lancet* 2:1023
171. Ciarlet M, Estes MK (1999) Human and most animal rotavirus strains do not require the presence of sialic acid on the cell surface for efficient infectivity. *J Gen Virol* 80:943–948
172. Banda K, Kang G, Varki A (2009) ‘Sialidase sensitivity’ of rotaviruses revisited. *Nat Chem Biol* 5:71–72
173. Haselhorst T, Fleming FE, Dyason JC, Hartnell RD, Yu X, Holloway G, Santegoets K, Kiefel MJ, Blanchard H, Coulson BS et al. (2009) Sialic acid dependence in rotavirus host cell invasion. *Nat Chem Biol* 5:91–93
174. Monnier N, Higo-Moriguchi K, Sun ZY, Prasad BV, Taniguchi K, Dormitzer PR (2006) High-resolution molecular and antigen structure of the VP8* core of a sialic acid-independent human rotavirus strain. *J Virol* 80:1513–1523
175. Lopez S, Arias CF (2006) Early steps in rotavirus cell entry. *Curr Top Microbiol Immunol* 309:39–66
176. Dormitzer PR, Sun ZY, Blixt O, Paulson JC, Wagner G, Harrison SC (2002) Specificity and affinity of sialic acid binding by the rhesus rotavirus VP8* core. *J Virol* 76:10512–10517
177. Kuhlenschmidt MS, Rolsma MD, Kuhlenschmidt TB, Gelberg HB (1997) Characterization of a porcine enterocyte receptor for group A rotavirus. *Adv Exp Med Biol* 412:135–143
178. Delorme C, Brussow H, Sidoti J, Roche N, Karlsson KA, Neeser JR, Teneberg S (2001) Glycosphingolipid binding specificities of rotavirus: identification of a sialic acid-binding epitope. *J Virol* 75:2276–2287
179. Superti F, Donelli G (1991) Gangliosides as binding sites in SA-11 rotavirus infection of LLC-MK2 cells. *J Gen Virol* 72:2467–2474
180. Guo CT, Nakagomi O, Mochizuki M, Ishida H, Kiso M, Ohta Y, Suzuki T, Miyamoto D, Hidari KI, Suzuki Y (1999) Ganglioside GM(1a) on the cell surface is involved in the infection by human rotavirus KUN and MO strains. *J Biochem* 126:683–688
181. Liddington RC, Yan Y, Moulai J, Sahli R, Benjamin TL, Harrison SC (1991) Structure of simian virus 40 at 3.8-Å resolution. *Nature* 354:278–284
182. Neu U, Woellner K, Gauglitz G, Stehle T (2008) Structural basis of GM1 ganglioside recognition by simian virus 40. *Proc Natl Acad Sci U S A* 105:5219–5224
183. Stehle T, Harrison SC (1997) High-resolution structure of a polyomavirus VP1-oligosaccharide complex: implications for assembly and receptor binding. *EMBO J* 16:5139–5148
184. Stehle T, Yan Y, Benjamin TL, Harrison SC (1994) Structure of murine polyomavirus complexed with an oligosaccharide receptor fragment. *Nature* 369:160–163
185. Cahan LD, Paulson JC (1980) Polyoma virus adsorbs to specific sialyloligosaccharide receptors on erythrocytes. *Virology* 103:505–509
186. Cahan LD, Singh R, Paulson JC (1983) Sialyloligosaccharide receptors of binding variants of polyoma virus. *Virology* 130:281–289
187. Fried H, Cahan LD, Paulson JC (1981) Polyoma virus recognizes specific sialyloligosaccharide receptors on host cells. *Virology* 109:188–192
188. Tsai B, Gilbert JM, Stehle T, Lencer W, Benjamin TL, Rapoport TA (2003) Gangliosides are receptors for murine polyoma virus and SV40. *EMBO J* 22:4346–4355
189. Stehle T, Harrison SC (1996) Crystal structures of murine polyomavirus in complex with straight-chain and branched-chain sialyloligosaccharide receptor fragments. *Structure* 4:183–194

190. Caruso M, Belloni L, Sthandier O, Amati P, Garcia MI (2003) Alpha4beta1 integrin acts as a cell receptor for murine polyomavirus at the postattachment level. *J Virol* 77:3913–3921
191. Ewers H, Romer W, Smith AE, Bacía K, Dmitrieff S, Chai W, Mancini R, Kartenbeck J, Chambon V, Berland L et al. (2010). GM1 structure determines SV40-induced membrane invagination and infection. *Nat Cell Biol* 12,11–18; sup pp 11–12.
192. Qian M, Cai D, Verhey KJ, Tsai B (2009) A lipid receptor sorts polyomavirus from the endolysosome to the endoplasmic reticulum to cause infection. *PLoS Pathog* 5:e1000465
193. Keppler OT, Herrmann M, Oppenlander M, Meschede W, Pawlita M (1994) Regulation of susceptibility and cell surface receptor for the B-lymphotropic papovavirus by N glycosylation. *J Virol* 68:6933–6939
194. Dugan AS, Gasparovic ML, Atwood WJ (2008) Direct correlation between sialic acid binding and infection of cells by two human polyomaviruses (JC virus and BK virus). *J Virol* 82:2560–2564
195. Liu CK, Wei G, Atwood WJ (1998) Infection of glial cells by the human polyomavirus JC is mediated by an N-linked glycoprotein containing terminal alpha(2-6)-linked sialic acids. *J Virol* 72:4643–4649
196. Komagome R, Sawa H, Suzuki T, Suzuki Y, Tanaka S, Atwood WJ, Nagashima K (2002) Oligosaccharides as receptors for JC virus. *J Virol* 76:12992–13000
197. Elphick GF, Querbes W, Jordan JA, Gee GV, Eash S, Manley K, Dugan A, Stanifer M, Bhatnagar A, Kroeze WK et al. (2004) The human polyomavirus, JCV, uses serotonin receptors to infect cells. *Science* 306:1380–1383
198. Dugan AS, Eash S, Atwood WJ (2005) An N-linked glycoprotein with alpha(2,3)-linked sialic acid is a receptor for BK virus. *J Virol* 79:14442–14445
199. Low JA, Magnuson B, Tsai B, Imperiale MJ (2006) Identification of gangliosides GD1b and GT1b as receptors for BK virus. *J Virol* 80:1361–1366
200. Erickson KD, Garcea RL, Tsai B (2009) Ganglioside GT1b is a putative host cell receptor for the Merkel cell polyomavirus. *J Virol* 83:10275–10279
201. Amberg N, Kidd AH, Edlund K, Nilsson J, Pring-Akerblom P, Wadell G (2002) Adenovirus type 37 binds to cell surface sialic acid through a charge-dependent interaction. *Virology* 302:33–43
202. Burmeister WP, Guilligay D, Cusack S, Wadell G, Amberg N (2004) Crystal structure of species D adenovirus fiber knobs and their sialic acid binding sites. *J Virol* 78:7727–7736
203. Schmidt M, Chiorini JA (2006) Gangliosides are essential for bovine adeno-associated virus entry. *J Virol* 80:5516–5522
204. Kaludov N, Brown KE, Walters RW, Zabner J, Chiorini JA (2001) Adeno-associated virus serotype 4 (AAV4) and AAV5 both require sialic acid binding for hemagglutination and efficient transduction but differ in sialic acid linkage specificity. *J Virol* 75:6884–6893
205. Dickey DD, Excoffon KJ, Koerber JT, Bergen J, Steines B, Klesney-Tait J, Schaffer DV, Zabner J (2011) Enhanced sialic acid-dependent endocytosis explains the increased efficiency of infection of airway epithelia by a novel adeno-associated virus. *J Virol* 85:9023–9030

Endosialidases: Versatile Tools for the Study of Polysialic Acid

Elina Jakobsson, David Schwarzer, Anne Jokilammi, and Jukka Finne

Abstract Polysialic acid is an α 2,8-linked *N*-acetylneuraminic acid polymer found on the surface of both bacterial and eukaryotic cells. Endosialidases are bacteriophage-borne glycosyl hydrolases that specifically cleave polysialic acid. The crystal structure of an endosialidase reveals a trimeric mushroom-shaped molecule which, in addition to the active site, harbors two additional polysialic acid binding sites. Folding of the protein crucially depends on an intramolecular C-terminal chaperone domain that is proteolytically released in an intramolecular reaction. Based on structural data and previous considerations, an updated catalytic mechanism is discussed. Endosialidases degrade polysialic acid in a processive mode of action, and a model for its mechanism is suggested. The review summarizes the structural and biochemical elucidations of the last decade and the importance of endosialidases in biochemical and medical applications. Active endosialidases are important tools in studies on the biological roles of polysialic acid, such as the pathogenesis of septicemia and meningitis by polysialic acid-encapsulated bacteria, or its role as a modulator of the adhesion and interactions of neural and other cells. Endosialidase mutants that have lost their polysialic acid cleaving activity while

This work has been supported by grants from the Academy of Finland and the Deutsche Forschungsgemeinschaft. Elina Jakobsson and David Schwarzer have contributed equally to this review.

E. Jakobsson and A. Jokilammi

Department of Medical Biochemistry and Genetics, University of Turku, Kiinamyllynkatu 10, 20520 Turku, Finland

D. Schwarzer

Institute for Cellular Chemistry (OE 4330), Hannover Medical School, Carl-Neuberg-Street 1, 30625 Hannover, Germany

J. Finne (✉)

Department of Biosciences, Division of Biochemistry and Biotechnology, University of Helsinki, P. O. Box 56, 00014 Helsinki, Finland
e-mail: jukka.finne@helsinki.fi

retaining their polysialic acid binding capability have been fused to green fluorescent protein to provide an efficient tool for the specific detection of polysialic acid.

Keywords Bacteria · Bacteriophages · Cell adhesion · Endosialidase · Polysialic acid

Contents

1	Introduction	30
2	Bacteriophages	31
3	<i>Escherichia coli</i> K1 and Polysialic Acid	32
4	K1-Specific Bacteriophages and Endosialidases	33
	4.1 Nomenclature of Exo- and Endosialidases	34
	4.2 Specificity of Endosialidase	35
5	The Endosialidase Gene and Its Modular Architecture	36
6	Crystal Structure of Endosialidase F	39
7	The Active Site of Endosialidase	42
	7.1 Catalytic Mechanism of Endosialidase	45
8	Oligosialic Acid Binding Sites	48
9	Intramolecular Chaperone of Endosialidase	49
10	Processivity of Endosialidase	52
11	Applications of Active Endosialidase	55
	11.1 Preparation and Properties of Catalytically Active Endosialidase	55
	11.2 Applications of Active Endosialidase for the Study of NCAM	56
	11.3 Other Applications of Active Endosialidase	58
12	Catalytically Inactive Endosialidase	60
	12.1 Generation and Properties of Inactive Endosialidase	60
	12.2 Use of Inactive Endosialidase for the Detection of Polysialic Acid	61
	References	62

Abbreviations

CTD	C-terminal chaperone domain
endoN	Endosialidase (endo- <i>N</i> -acetylneuraminidase)
NCAM	Neural cell adhesion molecule
Neu5Ac	<i>N</i> -Acetylneuraminic acid
Neu5Gc	<i>N</i> -Glycolylneuraminic acid
Sia	Sialic acid

1 Introduction

Polysialic acid is a linear carbohydrate homopolymer consisting of *N*-acetylneuraminic acid units and is found on the surface of both bacterial and eukaryotic cells [1–3]. In eukaryotes, the neural cell adhesion molecule (NCAM) is a major carrier of polysialic acid chains [4]. Polysialic acid has been ascribed important roles in neural development, cell differentiation, and plasticity [5]. It is also expressed in malignancies and has been associated with tumor growth and

metastasis [6–8]. In prokaryotes polysialic acid is found as a capsular polysaccharide and virulence factor of neurotrophic bacterial pathogens such as group B meningococci and *Escherichia coli* K1 [9]. In contrast to many other capsular polysaccharides, polysialic acid is a poor immunogen [10].

Bacteriophages are the most abundant organism in nature and are isolated from a wide variety of sources such as coastal waters and sewage [11]. *E. coli* K1-specific bacteriophages contain the enzyme endosialidase as their tailspike protein [12, 13] which is required for the recognition of the bacteria, binding to and degrading their polysialic acid capsule [14]. Phage mutants have been isolated that have lost their polysialic acid cleaving activity while retaining their binding activity for polysialic acid [15].

The active endosialidase with its ability to degrade specifically polysialic acid represents a powerful tool for studying the chemistry, metabolism, and biological roles of polysialic acid. The inactive endosialidase, on the other hand, specifically able to recognize and remain bound to polysialic acid, represents a sensitive and convenient tool for the specific detection of polysialic acid in different applications.

2 Bacteriophages

Viruses have escorted life since its very first moment 3.8 billion years ago. They are found in all domains of life – plants, animals, fungi, protists, prokaryotes, as well as archaea. The viruses of prokaryotes are named bacteriophages, or in short “phages,” and were discovered in the early twentieth century independently by Twort and d’Herelle [16, 17]. With an estimated population size of more than 10^{31} phage particles (virions) in the oceans, bacteriophages are likely to be numerically the most abundant organisms in the biosphere, outnumbering their bacterial hosts by up to two orders of magnitude [18–21]. Phages have followed prokaryotes into all known habitats including soil, marine and lake waters, animals, and plants. They have been isolated from a wide variety of sources – amongst others, coastal waters and sewage [11]. About 96% of known phages are ascribed to the most well characterized group of bacteriophages, the order of *Caudovirales*, also referred to as tailed phages [21, 22].

Caudovirales are composed of a head and a tail. They exclusively contain proteins and linear, double-stranded DNA. The outer shell of the head – the capsid – basically exhibits an icosahedral symmetry and encloses the DNA. The tail, essentially a hollow tube with sixfold symmetry, can contain additional structures, namely a baseplate with six copies of tailspike, tail fibers, or collars. According to microscopically observable features of the tail morphology, the *Caudovirales* are subdivided into three different families: the *Myoviridae*, the *Siphoviridae*, and the *Podoviridae*, reproducing Bradley groups A, B, and C, respectively [23, 24]. The tail of myoviruses (*Myoviridae*) comprises a long, rigid tail tube that is surrounded by a contractile sheath. Mostly they contain a sixfold symmetric baseplate composed of more than 16 different gene products. In total, myoviruses are the most

complex phages [25–27]. Siphoviruses contain a long, flexible tail structure that is not contractile and a threefold symmetric baseplate [28]. In contrast to *Myo-* or *Siphoviridae*, the podoviruses have a short non-contractile tail that forms a rigid sixfold symmetric structure [13, 29].

All tailed phages have evolved tailspike and fiber proteins for efficient virus-host-interactions. These specialized adhesions mediate the recognition and attachment to the bacterial surface and constitute the key determinants for host specificity. Interestingly, many spikes and fibers are composed of homotrimeric complexes which remain stable even in the presence of sodium dodecyl sulfate (SDS) [12, 14, 30–34]. Several phages have developed tailspike proteins with an enzymatic activity in order to penetrate the thick layer of lipopolysaccharides or capsular polysaccharides of many pathogenic bacteria. These capsule-specific depolymerases (hydrolases or lyases) are required to gain access to and to fix the phage at the bacterial outer membrane [13, 14, 35–38].

3 *Escherichia coli* K1 and Polysialic Acid

Capsular polysaccharides are important virulence factors of many pathogenic bacteria causing severe invasive infections including meningitis, pneumonia, septicemia, osteomyelitis, septic arthritis, and pyelonephritis [39–41]. With a high water-binding potential, capsular polysaccharides provide a thick layer (400 nm or more) protecting the bacterium from desiccation and other harsh environmental conditions. The polysaccharides mask underlying cellular surface molecules and thereby confer resistance against the host immune defense like complement-mediated lysis, phagocytosis, and opsonization (see [42] for review). It has recently been suggested that capsular polysaccharides may be essential in forming intracellular bacterial communities during urinary tract infections [43].

One prominent example of an encapsulated bacterium is *E. coli* K1, a Gram-negative neuroinvasive bacterium that causes meningitis, septicemia as well as urinary tract infections [42, 44, 45]. The pathogen is surrounded by a thick capsule of polysialic acid, a linear carbohydrate homopolymer consisting of up to 200 α 2,8-linked sialic acid units per chain. Sialic acids are derivatives of neuraminic acid (5-amino-3,5-dideoxy-D-glycero-D-galacto-non-2-ulosonic acid), which is a nine-carbon monosaccharide, a ketononose. Various substitutions like acylation, sulfation, and methylation and their combinations result in more than 50 known structural variants in nature [46]. The most common member of the sialic acids group is *N*-acetylneuraminic acid.

Polysialic acid can be composed of 5-*N*-acetylneuraminic acid (Neu5Ac), 5-*N*-glycolylneuraminic acid (Neu5Gc), or deaminated neuraminic acid (KDN) residues, which are usually joined by α 2,8, α 2,9, or alternating α 2,8/ α 2,9 ketosidic linkages [47–53].

The *E. coli* K1 polysaccharide is composed of α 2,8-linked Neu5Ac. Identical capsules have also been found in *Neisseria meningitidis* serogroup B, *Moraxella*

nonliquefaciens, and *Mannheimia (Pasteurella) haemolytica* A2 [47–52]. Polysialic acid with alternating α 2,8/ α 2,9-linkages has been identified in *E. coli* K92 [52, 54, strain Bos12 (O16:K92:H-)], whereas the capsule of *N. meningitidis* serogroup C is solely α 2,9-linked [49]. Polysialic acid can be modified by *O*-acetyl esters at free carbon positions of the Neu5Ac glycerol side chain [55–57].

The α 2,8-linked polysialic acid is also found in vertebrates, where it is primarily found as a posttranslational modification of NCAM [4, 58]. It was also described as a major glycan structure of the CD36 scavenger receptor in human milk, the α -subunit of the voltage-gated sodium channel in rat brain, neuropilin-2 on maturing human dendritic cells, and recently identified as posttranslational modification of SynCAM 1 in mouse brain [2, 3, 59–63]. Polysialic acid has been ascribed important roles in neural development, cell differentiation and plasticity [5]. It is also expressed in malignancies and has been associated with tumor growth and metastasis [6–8].

By mimicry of this host structure, polysialic acid encapsulated bacteria evade the human immune system [64]. In contrast to many other capsular polysaccharides, α 2,8-linked polysialic acid is a poor immunogen and protects the bacterium from complement-mediated lysis and opsonophagocytosis, and it has been suggested that it enables the bacterium to cross the blood–brain barrier [10, 39, 41, 65–67].

4 K1-Specific Bacteriophages and Endosialidases

The polysialic acid capsule of *E. coli* K1 is not only an essential virulence factor; it also provides an attachment site for specialized K1 bacteriophages. Almost 30 lytic enterobacterial phages specific for *E. coli* K1 have been isolated, mainly from sewage samples. Interestingly, they have been found to exhibit different morphologies, thus belonging to the three different families of the *Caudovirales* (Table 1).

Most of the characterized K1-specific phages belong to *Podoviridae*; five isolates were found to be *Siphoviridae* [77, 86], whereas the isolate phi92 contains a contractile tail – the characteristic feature of *Myoviridae* [74]. In 2005, the first temperate phage CUS-3 was identified to be integrated in the genome of the *E. coli* K1 strain RS218 [80, 81]. Since then, six further *E. coli* K1 lysogen genomes (i.e., bacterial genomes containing a prophage) have been found [82–85, 87].

As a common characteristic feature, all known K1-specific phages and prophages contain a gene encoding for a tailspike protein with polysialic acid degrading activity – an endosialidase [12, 13]. The phages require endosialidases for the recognition of the bacteria, the binding to and the degradation of their polysialic acid capsule. Many viruses that use carbohydrate as receptors contain receptor-destroying enzymes. For K1-specific bacteriophages the endosialidases represent the receptor-destroying enzymes [36, 88–90].

Table 1 Overview of isolated bacteriophages specific for *Escherichia coli* K1 carrying endosialidase activity

Name	Virus family	Host strain or lysogen	Phage isolation	Endosialidase gene cloned	Enzyme
K1A	<i>Podoviridae</i>	<i>E. coli</i> U9/41 (O2:K1:H4)	[68]	[69]	endoNA
K1B, K1C, and K1D	<i>Podoviridae</i>	<i>E. coli</i> U9/41 (O2:K1:H4)	[68]	–	–
K1E	<i>Podoviridae</i>	<i>E. coli</i> U9/41 (O2:K1:H4)	[68]	[70, 71]	endoNE
1.2	<i>Podoviridae</i>	<i>E. coli</i> K235 (O1:K1:H-)	[72]	–	–
B, C, D, F, G, K, L, P, and R	n.d.	<i>E. coli</i> K1	[73]	–	–
phi92	<i>Myoviridae</i>	<i>E. coli</i> Bos12 (O16:K92:H-), <i>E. coli</i> K1	[74]	–	–
K1F	<i>Podoviridae</i>	<i>E. coli</i> K1	[75]	[14, 76]	endoNF
63D	<i>Siphoviridae</i>	<i>E. coli</i> K1	[77]	[78]	endoN63D
3, 9, 63A, and 63E	<i>Podoviridae</i>	<i>E. coli</i> K1	[77]	–	–
a and d	<i>Siphoviridae</i>	<i>E. coli</i> K1	[77]	–	–
K1-5	<i>Podoviridae</i>	<i>E. coli</i> K1, <i>E. coli</i> K5	[79]	[79]	endoNK1-5
CUS-3	n.d. (prophage)	<i>E. coli</i> K1	[80, 81]	[81]	endoNK1
<i>E. coli</i> UTI89	n.d. (prophage)	<i>E. coli</i> K1	[82]	–	–
<i>E. coli</i> APEC O1	n.d. (prophage)	<i>E. coli</i> K1	[83]	–	–
<i>E. coli</i> IAI39 and S88	n.d. (prophage?)	<i>E. coli</i> K1	[84]	–	–
<i>E. coli</i> IHE3034	n.d. (prophage?)	<i>E. coli</i> IHE3034 (lysogen)	[85]	–	–
K1H = K1-dep(1) and K1G = K1-dep(4)	<i>Siphoviridae</i>	<i>E. coli</i> K1	[86]	–	–
<i>E. coli</i> UM146	n.d. (prophage?)	<i>E. coli</i> UM146 (lysogen)	[87]	–	–

4.1 Nomenclature of Exo- and Endosialidases

There are three different types of sialidases known, relating to the mode of cleavage reaction:

1. Exosialidases (exo- α -sialidase, neuraminidase) are hydrolases chopping off terminal residues from the non-reducing end of glycans
2. Anhydrosialidases are a special type of exosialidases which release 2,7-anhydro- α -N-acetylneuraminate in an elimination reaction from the non-reducing end [91]

Table 2 IUBMB enzyme nomenclature for endo- and exosialidase

EC	3.2.1.129	3.2.1.18	4.2.2.15 formerly: 3.2.1.138
IUBMB name	Endo- α -sialidase	Exo- α -sialidase	Anhydrosialidase
Other name(s)	<ul style="list-style-type: none"> • Endo-<i>N</i>-acetylneuraminidase • Endoneuraminidase • Endo-<i>N</i>-acetylneuraminidase • Poly(α2,8-sialosyl) endo-<i>N</i>-acetylneuraminidase • Poly(α2,8-sialoside) α2,8-sialosylhydrolase • Endosialidase • EndoN 	<ul style="list-style-type: none"> • Neuraminidase • Sialidase • α-Neuraminidase • <i>N</i>-Acetylneuraminidase • <i>Suggested: exoN</i> 	<ul style="list-style-type: none"> • Anhydroneuraminidase • Sialglycoconjugate <i>N</i>-acetylneuraminyhydrolase (2,7-cyclizing) • Sialidase L
Systematic name	Polysialoside α 2,8-sialosylhydrolase	Acetylneuraminyl hydrolase	Glycoconjugate sialyl-lyase (2,7-cyclizing)

3. Endosialidases (endo- α -sialidase, endo-*N*-acetylneuraminidase) cleave within a polysialic acid chain

The first exosialidase was discovered by 1942 [92]. After 35 years, endosialidases were discovered in a bacteriophage lysate to promote research in the field of polysialic acid [35, 93, 94]. Exo- α -sialidase has been listed in the IUBMB nomenclature since 1961, while endo- α -sialidase and anhydrosialidases have been incorporated much later, in 1990 [95] and 1992, respectively. The endo- α -sialidase (EC 3.2.1.129) is a glycosyl hydrolase which cleaves α 2,8-linkages of polysialic acid. Names and reactions of sialidases are summarized in Table 2.

Endosialidases have also been called “endo-*N*-acyl-neuraminidases” (in short endoN), since they can cleave polysialic acid composed of either 5-*N*-acetylneuraminic acid (Neu5Ac) or 5-*N*-glycolylneuraminic acid (Neu5Gc) – two derivatives of sialic acid – in α 2,8-linkage [96]. Even though Cabezas suggested later on to employ the term “endosialidases” to follow the recommendations made for other (exo-)sialidases [95], the abbreviation “endoN” has been maintained. For example, the endosialidase of coliphage K1F is termed “endosialidase F” or “endoNF.”

4.2 Specificity of Endosialidase

In contrast to exosialidases that have been found to release sialic acid from galactose or galactosamine in α 2,3 or α 2,6 linkages, or α 2,8-linked to sialic acid, respectively [97, 98], endosialidases specifically cleave α 2,8-linked polysialic acid. Further, exosialidases can be promiscuous in their substrate specificity like the sialidase from *Arthrobacter ureafaciens* releasing sialic acid from any glycoconjugate in any possible linkage with different activities [97]. Endosialidases are strictly limited to polysialic acid, cleaving solely α 2,8-linkages. Interestingly,

myovirus phi92 has been isolated from *E. coli* K92 that is encapsulated by polysialic acid in alternating α 2,8/ α 2,9-ketosidic linkages [74]. However, phi92 has also been found to infect *E. coli* K1 and further studies have revealed that the endosialidase of phi92 solely cleaves the α 2,8-linkages, since the polysialic acid of *N. meningitidis* serogroup C polysaccharide containing α 2,9 linked *N*-acetylneuraminic acid residues is not cleaved [74, 99]. Bacteriophages capable of infecting *N. meningitidis* serogroup C, encapsulated with an α 2,9-linked sialic acid polymer, are as yet undiscovered.

Polysialic acid can be *O*-acetylated, which renders it more immunogenic than the non-*O*-acetylated polymer [100]. Endosialidase E rapidly hydrolyses both the *O*-acetylated and non-*O*-acetylated forms of polysialic acid [101].

Endosialidases require a minimum chain length of eight sialic acid residues for cleavage under moderate digestion conditions [102]. The presence of a minimum of three residues is required on the non-reducing end from the cleavage site, and a minimum of five residues on the reducing-end moiety. Endosialidases from different sources show very similar cleavage patterns [103]. However, under in vitro conditions with more exhaustive digestion the minimum chain length of polysialic acid to be cleaved has ranged from trimer to octamer [103–105] (for an overview see [77]). Interestingly, endoNF is able to cleave oligosialic acid with a minimum of four residues whereas three residues are required at the non-reducing end [105]. In the crystal structure of the inactive mutant endoNF-H350A, a tri-sialic acid is bound to the active site. The β -anomeric state of the reducing end sialic acid residue clearly corroborates the fact that it resembles the terminal residue of the chain and thus the bound trimer is a product of the cleavage reaction (cf. Fig. 2c, Sect. 7; [106]).

5 The Endosialidase Gene and Its Modular Architecture

The different anti-K1 bacteriophages have acquired their K1-specificity by inserting an endosialidase gene into the tailspike or fiber locus of the respective precursor phage. For example, K1F is a derivative of a T7 phage in which the T7 tail fiber gene has been replaced by the gene encoding for the endosialidase (endoNF). Similarly, in K1E and K1-5 the endosialidase gene has been inserted into the tailspike locus of gene 47 of salmonella phage SP6 [81]. Finally, the prophage CUS-3 was identified as being emanated from the salmonella phage P22 by insertion of an endosialidase gene. The comparative genome analysis [81] has thus revealed that K1-specific bacteriophages do not have a common ancestor. Interestingly, the respective endosialidase gene is neatly fused to the coding sequence of a short N-terminal peptide sequence (K1E and K1-5) or a capsid binding domain (K1F) from the original fiber gene to allow proper binding of the newly acquired endosialidase to the phage particle [12, 13]. The capsid binding modules of respective progenitors are combined with horizontally transferred endosialidase genes. This kind of reconstruction allows bacteriophages to change host specificity without major structural rearrangements of the phage particle. Further, the phage K1-5 carries two enzymes, the endosialidase and the K5-lyase,

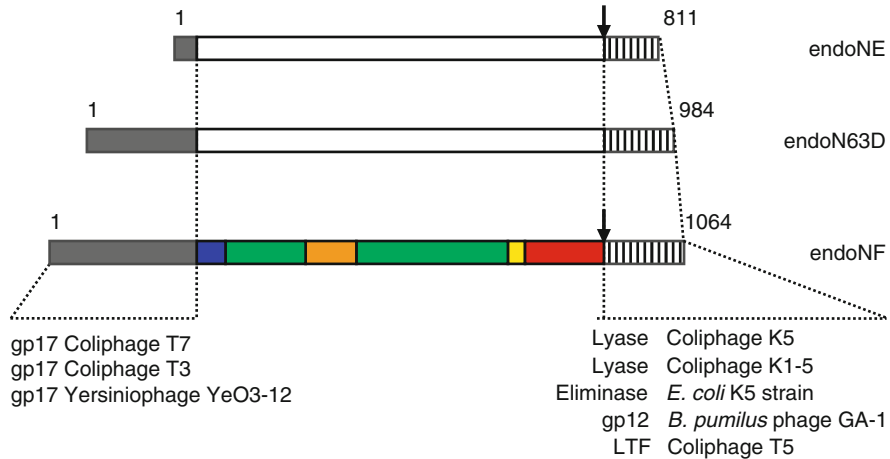


Fig. 1 Schematic representation of cloned endosialidases. The highly conserved catalytic part common to all endosialidases is depicted in white. Differing from this, the catalytic part of endoNF is subdivided into domains identified in the crystal structure [12] colored in the same code as in Fig. 2b (see legend of Fig. 2 for details). The variable N-terminal parts are shown in gray. Proteolytic cleavage site Thr ↓ Ser identified in Mühlhoff et al. [14] and Schulz et al. [109] are indicated by arrows whereas conserved stretches within the C-terminal part are schematically represented by black line. The N-terminal domains of endoN63D and endoNF show sequence similarities to the N-terminal part of the tail fiber protein gp17 of coliphages T3 and T7, and of yersiniophage YeO312. The C-terminal part exhibits sequence similarities to the C-terminal parts of K5-specific lyases of coliphages K1 and K1-5, the K5-eliminase of an *Escherichia coli* strain, the neck-appendage protein precursor of *Bacillus pumilus* phage GA-1, and the L-shaped tail-fiber protein (LTF) of coliphage T5 (modified from Schwarzer [110])

thereby being capable of infecting both *E. coli* K1 and K5, which have two chemically different capsule types [13, 79]. Interestingly, both tailspike proteins are linked to the same adapter molecule of the phage tail via a short N-terminal undecapeptide (endoN) or decapeptide (K5-lyase) sequence, respectively [12, 13].

It has been shown that the endosialidase is the only factor mediating K1-specificity since the T7 phage lacking the endosialidase gene is unable to infect *E. coli* K1 strains. However, after enzymatic removal of the K1 capsule by endosialidase, the naked *E. coli* K1 strain can be infected by T7 [38]. Vice versa, the K1F phage is unable to infect *E. coli* strains lacking a polysialic acid capsule [108].

So far, five endosialidase genes have been cloned and expressed as functional proteins from *E. coli* phages K1A, K1E, K1F, 63D, and K1-5 [69–71, 78, 79]. The tailspike proteins endoNA, endoNE, and endoNK1-5 are equal in length (811 amino acids), whereas endoN63D and endoNF comprise 984, and 1,064 amino acids, respectively. EndoNE and endoNK1-5 are nearly identical with about 98% amino acid sequence similarity, and the recently published sequence of endoNA exhibits 90% similarity to endoNE. Therefore endoNE will be considered in the following as representative of this group.

Endosialidases share a modular architecture (Fig. 1) which can be subdivided into three distinct parts:

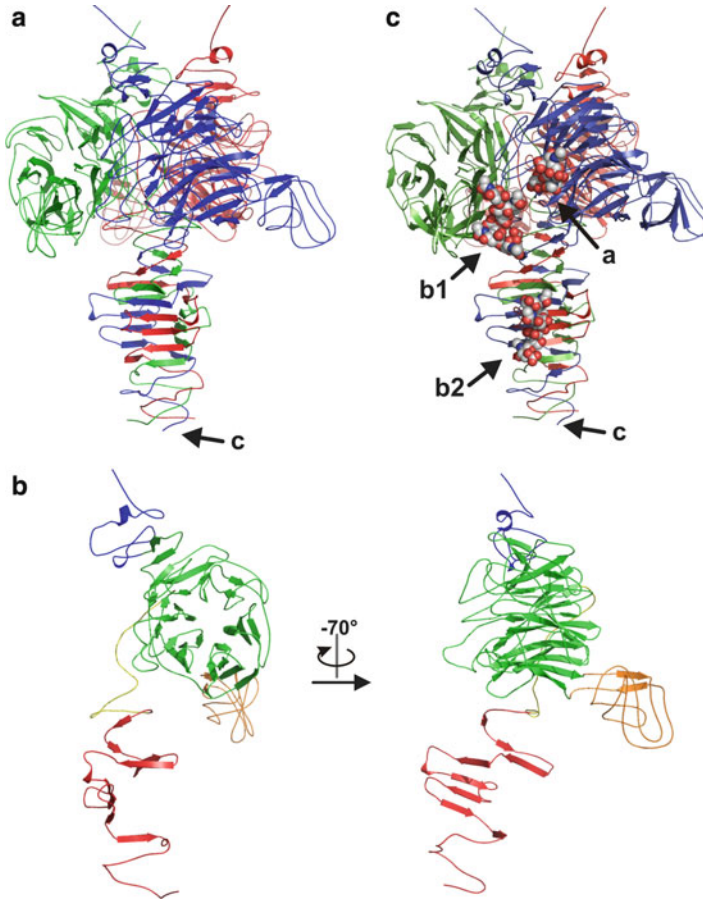


Fig. 2 Ribbon diagram of endosialidase F [106]. (a) Wild type endoNF (PDB ID: 3GVL; [106]). The three subunits of the homotrimeric complex are given in *red*, *green*, and *blue*. The homotrimer matures by releasing the intramolecular chaperone domain at the proteolytical cleavage site *c* (see Sect. 7). Bound sialic acids are hidden. (b) Isolated subunit of wild type endoNF (PDB ID: 3GVL). View along the pseudo sixfold symmetry axis of the β -propeller and rotated by 70° . The domains are indicated in different colors. A small N-terminal portion, remnant of the head-binding domain, is given in *blue*. The six-bladed β -propeller, the β -barrel domain, and the stalk domain are *green*, *orange*, and *red*, respectively. The *yellow* stretch links the β -propeller with the stalk domain. (c) The three subunits of the endoNF-R647A mutant (PDB ID: 3GVJ) are given in *red*, *green*, and *blue*. Spheres represent bound sialic acid oligomers in three distinct binding sites *a*, *b1*, and *b2* at the front face of the enzyme. The crystal structures of endoNF-H350A and endoNF-R647A were superposed (Coot SSM superpose [107]) to place the sialic acid trimer bound to the active site *a* of endoNF-H350A (PDB ID: 3GVK) in the crystal structure of endoNF-R647A. The overall backbone of the two superposed proteins is highly conserved. Modified from Schulz et al. [106]

1. The endosialidases differ in their N-terminal binding modules required for their attachment to the respective phage particle. EndoNF and endoN63D are N-terminally extended by 200 and 120 amino acids, respectively. These regions are a common feature of the tailspike or fiber proteins of most members of

the T7-type superfamily as the coliphages T3 and T7, and of yersiniophage YeO3-12 [111]. Via this domain the proteins are directly fixed to the phage particle [81]. By contrast, endoNE contains an N-terminal undecapeptide sequence that specifically interacts with an adapter protein (gp37) common to all phages of the SP6-type [81].

2. The central part of about 660 amino acids shows high sequence similarities and harbors the enzymatic activity [14, 112]. The first crystal structure of this catalytic part has been solved for endoNF in 2005 [12] and will be described in Sect. 6. The color code of the catalytic part of endoNF reflects the distinct domains depicted in Fig. 2b.
3. Endosialidases contain an intramolecular C-terminal chaperone domain (CTD). This part is required for proper folding of the catalytic part and shares sequence similarities with the C-terminal parts of a variety of tailspike and fiber proteins which have been identified to share a similar chaperone function [112]; for an overview of further similar proteins see Schulz et al. [109]. The intramolecular chaperone domain is proteolytically released after completion of its job and will be described in Sect. 9.

The first cloned endosialidase gene of phage K1F (named endo-N) has been expressed as insoluble protein. The molecular weight of the protein perfectly matched the endoNF protein purified from bacteriophage particles. Therefore, it has been suggested that it requires bacteriophage chaperones that were not co-expressed in this attempt [76]. However, the first endosialidase gene of K1F that can be expressed as soluble active enzyme (endoNF) was cloned in 2003. Sequence comparison of endo-N and endoNF revealed that the endo-N gene was cloned presumably based on a sequencing mistake which led to a premature stop of the cloned gene [14]. Endo-N by this lacks the crucial C-terminal intramolecular chaperone domain and therefore is not folded properly.

6 Crystal Structure of Endosialidase F

The crystal structure of the catalytic part of endoNF (amino acids 246–910) has been solved in three studies [12, 106, 113] (Table 3). First, the protein structure has been solved in its *apo* state and with bound substrate after soaking of sialic acid trimer. By this means, two oligosialic acid binding sites have been identified [12] (see Fig. 2c and Sect. 8). To allow identification of longer sialic acid oligomers bound to endoNF, two active site mutants have been created upon structural information from the first structure [106]. Finally, an endoNF structure has been solved at the high resolution of 0.98 Å [113].

Figure 2a shows the crystal structure of wild type endoNF (modified from Schulz et al. [106], PDB ID: 3GVL). The tailspike enzyme is a homotrimeric structure with a mushroom-like outline. The three subunits are depicted in *red*, *green*, and *blue* which each fold into three distinct domains (Fig. 2b). Apart from a short α -helix at

Table 3 Data statistics from the crystal structures of the catalytic part of endoNF (amino acids 246–910)

EndoNF variant	Wild type	Wild type	Wild type	R647A	H350A	Wild type
Published	[12]	[12]	[106]	[106]	[106]	[113]
PDB ID	1V0E	1V0F	3GVL	3GVJ	3GVK	3JU4
Space group	P222 ₁	P222 ₁	R3	R3	P2 ₁ 2 ₁ 2 ₁	R3
Resolution [Å]	30–1.9	30–2.55	19.6–1.4	11.9–1.5	15.0–1.8	19.36–0.98
Substrate [Å]	— DP3 ^a	DP3 ^b	DP5 ^b	DP5t	DP3 ^b	
<i>Observed sialic acid oligomer</i>						
Site <i>a</i>	—	—	—	—	DP3	—
Site <i>b1</i>	—	DP2	DP2	DP5	DP3	—
Site <i>b2</i>	—	DP1	DP1	DP4	DP2	DP1

^aOligosialic acid has been soaked into the endoNF crystal

^bCo-crystallization with oligosialic acid

the N-terminus (*blue*), the structure is dominated by β -structures. The N-terminal “cap” of the mushroom contains a sialidase domain which is composed of a six-bladed β -propeller (*green* in Fig. 2b) and a lectin-like β -barrel domain (*orange*) inserted in the second blade of the β -propeller. This dual domain structure is built up individually from each subunit followed by a linker stretch (*yellow*). The three sialidase domains interact by forming a hydrophobic core in the center of the “cap.”

As in exosialidases, the β -propeller (*green*) of endoNF contains the catalytic site (*a*, Fig. 2c) which is located in a wide cleft in its center (see Sect. 7) and was found to share a higher structural similarity with bacterial and eukaryotic rather than viral exosialidases [12].

Another feature common to all nonviral sialidases is the so-called Asp-box, a motif (S/T-X-D-[X]-G-X-T-W/F) that repeats one to five times along the sequences. Each Asp-box forms a clamp-like loop and they occur at topologically equivalent positions on the outside of the structure between the third and fourth β -strand of a propeller blade [12, 114, 115]. For example, in the sialidase NanI of *Clostridium perfringens* four Asp-boxes are located in the blades one to four (PDB ID: 2BF6 [115]). However, in endosialidases only two Asp-boxes have been found in the first and fourth blade of the propeller [12], with the sequences S-G-D-D-G-Q/K-T-W and S-X-D-X-G-X-X-W that are conserved in all so far known endosialidases. Interestingly, in endoNF the β -barrel domain is inserted between the third and fourth β -sheet of the second blade, thereby replacing a potential Asp-box. Since the Asp-boxes are located on the “back side” of the propeller and remote from the active site, any functions other than structural folds have not been found as yet for these motifs.

Endosialidases are not unique among sialidases in combining a β -propeller with other structures. Many other sialidases contain additional domains like β -barrels, jelly rolls, or immunoglobulin modules [116] which are placed N-terminally, C-terminally, or even inserted into the β -propeller. N-terminal lectin-like domains have been identified in leech sialidase L [117] and *Vibrio cholerae* sialidase [118]. Interestingly, in the *V. cholerae* sialidase a second β -structure domain is inserted

into the β -propeller. The trypanosome transsialidases contain a lectin-like domain fused C-terminally to the β -propeller [119, 120]. Often, the additional domains of sialidases harbor a carbohydrate binding site, which has been suggested to increase the catalytic efficiency [121]. In the β -barrel domain of endoNF an oligosialic acid binding site has been identified (*b1*; Fig. 2c) [12, 105, 106].

To compare structural similarities between endoNF and exosialidases, an HHpred/HHsearch analysis with the amino acid sequence of the endoNF sialidase domain (β -propeller and β -barrel domain, amino acids 293–753) was performed for this review, which revealed only bacterial sialidases as structural homologs. The best hits were the sialidase from *Salmonella typhimurium* LT2 [PDB ID: 3SIL] and the sialidase NanI from *C. perfringens* (PDB ID: 2BF6 [115]). Interestingly, like endoNF, NanI contains a β -barrel domain which is inserted in the second blade of the β -propeller. Although the identity of both sequences is only 13%, the common structural architecture suggests that the origin of the dual domain structure of endosialidases might have been jointly inserted from a bacterial sialidase into the tailfiber locus of a bacteriophage.

The stalk region of the endoNF trimer is built up by the intertwining tailspike domains (Fig. 2b, *red*). It folds into a left-handed triple β -helix that is interrupted by a triple β -prism domain [12]. Moreover, an oligosialic acid binding site has been identified in the triple β -prism domain (*b2*, Fig. 2c) which is important for proper binding and processive degradation of polysialic acid [12, 105, 106] (see Sects. 8 and 10). Triple β -structures have been described for other phage tailspike proteins: the triple β -helix domain was first described in the short tail fibers (gp12) followed by the cell-puncturing device (gp5), both of coliphage T4 [31, 32]. The triple β -prism was first identified to be part of a tailspike protein [30, 88] and later in the envelope-penetrating needle, both of the *Salmonella* phage P22 [122] as well as gp5 of coliphage T4 [32]. Interestingly, these bacteriophage structural proteins are generally thermostable and resistant to proteases as well as SDS denaturation [123, 124]. The endoNF trimer was also described to be highly stable even in the presence of SDS [12, 14, 112]. The unusual stability is not only created by the trimerization of the endoNF stalk but also by the fact that the stalk is stabilized by hydrophobic stacking interactions and by this more than 60% of the monomer's solvent accessible surface is buried [12].

The endoNF mushroom-like structure requires its intramolecular CTD to be properly folded (see Sect. 9). EndoNF constructs lacking the CTD or containing mutations of highly conserved amino acids within the CTD are prone to form insoluble aggregates [14, 76, 112].

In total, endoNF reveals a unique domain architecture that synergistically combines characteristics of the original elements and thus corroborates the concept of modular evolution of bacteriophages [125, 126]. The N-terminal head-binding domain is maintained from the progenitor phage. The sialidase β -propeller, and the lectin-like β -barrel domain, as well as the triple β -prism domain containing interaction sites *a*, *b1*, and *b2*, whose interplay is required for the specific binding and degradation of polysialic acid. Finally, the highly stable stalk domain typical of tailspike proteins requires its C-terminal intramolecular chaperone domain.

Concerning only one of the three faces of endoNF, as in Fig. 2c, one can conclude that one polysialic acid chain can interact with the binding site *b2* of the *red* subunit, the binding site *b1* of the *green* subunit, and the active site *a* of the *blue* subunit. By this unique arrangement, endosialidases are functional trimers. As stated above, many other sialidases contain additional domains that frequently contain carbohydrate binding modules which increase the catalytic efficiency. However, to allow proper catalysis the carbohydrate binding module and the active site have to be in close contact. In monomeric sialidases the additional domains are mostly inserted at the front side of the β -propeller like the β -barrel domain of NanI. In contrast, in endosialidases the β -barrel domain is inserted at the back side of the β -propeller with the binding site *b1* in close proximity to the active side of the adjacent subunit.

7 The Active Site of Endosialidase

A superposition (Coot SSM superpose [107]) of the β -propellers of endosialidases (endoNF) and bacterial exosialidases (represented by NanI) corroborates the highly conserved structural similarities between both enzymes. However, amino acids involved in catalysis differ dramatically (Fig. 3a). Only three amino acids are highly conserved by both sialidase types: two arginine (NanI: 555/615; endoNF: 596/647) and a glutamate (NanI: 539; endoNF: 581). In exosialidases, the two arginines together with Arg-266 form an arginine triad to bind the carboxyl group of the terminal sialic acid (Sia) residue. The third arginine is absent in endosialidases. In exosialidases the glycosidic linkage is attacked by a nucleophilic tyrosine (NanI: 655, *yellow* in Fig. 3b) forming a covalently linked Tyr-Sia intermediate. The Sia residue is released by a water molecule activated by a highly conserved aspartate (NanI: 291). These two important residues are absent in endosialidases, suggesting a different mechanism which will be discussed in Sect. 7.1. Further residues (*white* in Fig. 3b) involved in binding of the Sia residue are not conserved in endosialidases. In total, the active site of exosialidases is almost a narrow gorge allowing binding of only a Sia molecule and not a polysialic acid chain as in endosialidases. Additionally, all sialidase active sites contain a hydrophobic pocket to accommodate the N-acyl group of the Sia residue. Interestingly, a hydrophobic cap covers the hydrophobic pocket of NanI (*arrow* in Fig. 3b; [115]). In contrast, the active site of endoNF is a more accessible cleft and allows binding of the long polysialic acid. The three residues that are also conserved in exosialidases (Glu-581, Arg-596, Arg-647) have been mutated individually to alanine in endoNF, resulting in dramatic loss of activity below 2% of wild type activity [12, 106]. The combination of any two of these residues completely abolishes enzymatic activity [12, 105]. To identify further amino acids involved in binding sialic acid, the almost inactive mutant R647A has been co-crystallized with pentameric sialic acid [106]. Although pentameric (degree of polymerization 5; DP5) and tetrameric sialic acid (DP4) are bound to the binding sites *b1* and *b2*,

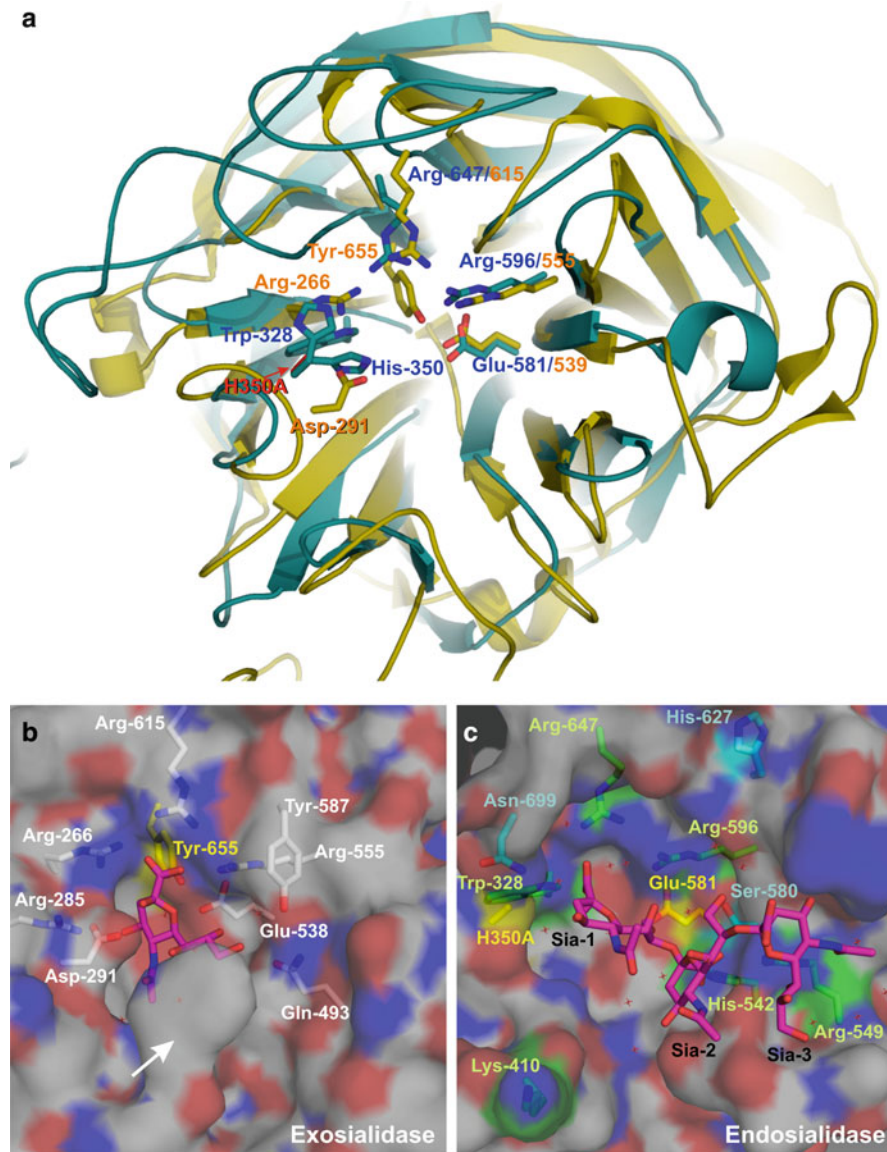


Fig. 3 Active sites of exo- and endosialidasas. **(a)** Superposition (Coot SSM superpose [107]) of the β -propellers of the exosialidase NanI from *Clostridium perfringens* (PDB ID: 2BF6) and of endoNF (wild type, PDB ID: 3GVL [106]). View along the pseudo-sixfold symmetry axis of the β -propeller. The active site residues that have been described to be involved in binding or cleavage of (poly-)sialic acid are shown as sticks and labeled with the position in the respective protein sequence. Remarkably, only three of these residues are conserved in both enzyme families. The two arginines (596 and 647 in endoNF) and the glutamate (581 in endoNF) were described to be involved in catalysis of exosialidasas. The third arginine (Arg-266) of the so-called arginine triad of exosialidase has no counterpart in endosialidasas. Instead, a tryptophan (Trp-328) and a

respectively (see Sect. 8), no sialic acid residues are identified in the active site. The residue Arg-647 is therefore important for proper coordination of a sialic acid residue of the chain and the mutation R647A already efficiently abolishes binding of oligosialic acid in the active site which is reflected in impaired enzymatic activity.

The amino acid His-350 is located in a position similar to the general acid/base aspartic acid of exosialidases (Fig. 3a; NanI: 291) and has been mutated to alanine for co-crystallization with DP5 [106]. In comparison to wild type endoNF, the alanine residue (*red* in Fig. 3a) H350A perfectly superposes with the C α and C β atoms of the histidine residue. Thus, the backbone of the β -propeller is not affected by this mutation. In the H350A mutant, a trimeric sialic acid (DP3) is bound in the active site coordinated by a network of water molecules and polar contacts (Fig. 3c). The geometry of the active site constrains the otherwise helical polysialic acid into an extended conformation (see Sect. 8 for helical epitopes).

Apart from water contacts, the reducing end residue (Sia-1) solely forms a hydrogen bond to Glu-581 of endoNF. The residue adopts an unusual conformation with its *N*-acetyl residue pointing from the protein and the carboxyl group pointing towards Trp-328 as well as the arginines 596 and 647. Compared to the sialic acid residue in exosialidase (cf. Fig. 3b) the Sia-1 residue of DP3 is turned by approximately 80° around the α 2,8-glycosidic linkage away from the two arginine residues. Sia-1 is found in its β -anomeric state. Since the β -anomer is the energetically favored conformation over the α -anomer with a ratio of 92:8 [128, 129], it has been concluded that this residue forms the terminal residue of the chain and therefore the DP3 is a product of catalysis. The other two sialic acid moieties of DP3 are bound by various polar contacts like His-542 and Arg-549. Apart from His-542, all residues involved in binding of the sialic acid trimer have been found to be critical for the enzymatic activity (*green* in Fig. 3c). Exchange to alanine leads to a dramatic loss of activity [106]. Lys-410 has been suggested as active-site gateway residue and to be important for delegating the polysialic acid chain from the binding site *b1* towards the active site, since exchange to alanine leads to reduction of

Fig. 3 (continued) histidine (His-350) residue are found in endoNF in this proximity. Further, the nucleophilic tyrosine (Tyr-655) that attacks the glycosidic bond and the general acid/base Asp-291 are absent in endosialidases. The residue His-350 has been exchanged into alanine to allow binding of short polysialic acid chains in the active site (PDB ID: 3GVK [106]). The alanine residue (H350A) of the H350A-mutant is shown in *red* (*arrow*) and perfectly superposes with the C α and C β atoms of the histidine residue of wild type endoNF. (b) Surface representation of the active site of the exosialidase NanI of *C. perfringens* (PDB ID: 2BF6 [115]) and bound sialic acid residue (*magenta*). Amino acids involved in Sia binding are given as *white sticks* and labeled. The catalytic tyrosine residue (Tyr-655) is represented as *yellow stick*. (c) Surface representation of the active site of the inactive mutant endoNF-H350A (PDB ID: 3GVK) with the bound sialic acid trimer (*magenta*). The sialic acid residue at the reducing end (*Sia-1*) is in the β -anomeric state. Amino acids that have been identified to be required for proper polysialic acid cleavage are given as *sticks*. *Yellow* residues have been described to be directly involved in catalysis [106, 127] whereas *green* amino acids have been mutated and identified as polysialic acid binding residues [12, 106]. *Blue* colored residues have been identified in inactive endoNA or endoNE mutants [15, 69]

Table 4 Active site residues of endosialidase

Active site residue	In endoNF	In endoNA/endoNE	Function
Trp	328 ^a (Arg)	118 ^b (Cys)	Binding
His	350 ^a (Ala, Asn, Gln)	140	General acid
Lys	410 ^a (Ala)	200	Active-site gateway
His	542 ^a (Ala)	332 ^b (Asn)	Binding
Arg	549 ^a (Ala)	339	Binding
Ser	580	370-Ala ^c (Glu)	Unknown/binding
Gly	581 ^a (Ala)	371	Binding/general base/ water activation
Arg	596 ^a (Ala)	386	Binding
His	627	417 ^b (Tyr)	Binding
Arg	647 ^a (Ala)	437	Binding
Asn	699	489 [@] (Asp)	Binding

^aResidue has been exchanged by the amino acid given in brackets [12, 106]

^bExchange of the residue by the amino acid given in brackets occurred naturally in endoNA by random mutation in combination with the amino acid exchange N489A ([@]) [69]

^cExchange of the residue by the amino acid given in brackets occurred naturally by random mutation in endoNE [69]

enzymatic activity [106]. Further residues were identified by natural mutation of the endosialidases from K1A and K1E to be important for catalysis (depicted in *blue* in Fig. 3c; [69], see Sect. 12.1). Table 4 summarizes the important active site residues identified in different endosialidases. Interestingly, the mutation N699D of endoNA (N489D) occurred only in combination with other mutations. The single mutant N489D shows a reduced activity but not complete loss of catalysis. Interestingly, in endoNF the amino acid His-627 is located outside the active site but exchange to tyrosine in endoNA (H417) already impairs activity by almost 60%. Like the mutant H542A of endoNF, the corresponding amino acid His-332 in endoNA mutated to Asn does not completely inactivate the enzyme but only in combination with the mutation N489D leads to a fully inactive endoNA-mutant. For endoNE the single mutation A370E results in strongly impaired activity [69]. However, in endoNF the corresponding residue is Ser-580 which is located in direct proximity to Glu-581 and forms a hydrogen bond via a water molecule to O9 of the Sia-1 residue [106]. The alanine at this position, as in endoNA and endoNE, might not interfere with catalysis, but the more space filling glutamate residue in the mutant A370E might be sufficient to block the active site residue for proper polysialic acid coordination.

7.1 Catalytic Mechanism of Endosialidase

As discussed in the previous section, exosialidases and endosialidases have different active site topologies. Although three residues have been identified to be highly

conserved in exosialidases and endosialidases (Glu-581, Arg-596, and Arg-647), the important catalytic nucleophile and the general acid/base (NanI: Tyr-655 and Asp-291, Fig. 3b) do not have a similar counterpart in endosialidases. Three studies have discussed a possible enzymatic mechanism of endosialidases. Based on the first crystal structure of endoNF [12] it has been suggested, that the enzyme accelerates a reaction which already occurs for polysialic acid without enzyme with lower velocity: under mild acidic conditions polysialic acid shows a significant intramolecular self-cleavage explained by the presence of the carboxyl groups in the chain [130]. Stummeyer et al. have suggested that this substrate-assisted reaction is enhanced by potentially introducing a conformational constraint into the polysialic acid chain upon binding to endoNF [12]. However, no sialic acid oligomers were found in the active site of the first structures to allow a validation of the mechanism.

In a second study, Glu-581, which is also conserved in exosialidases, has been suggested to act as catalytic residue since the mutation E581A results in complete loss of activity [127]. They could show that endoNF catalyzes polysialic acid cleavage in an inverting mechanism explained by the absence of a catalytic nucleophile. The activity optimum of the enzyme is around pH 5 [14, 127]. The pK_a of the carboxyl group of oligosialic acid (TFMU-DP3, trifluoromethylumbelliferyl linked to sialic acid trimer) has been determined to be 3.6–3.8 [127]. Based on this it has been stated that the carboxyl group of the sialic acid moiety in the active site must act as a general base in the catalytic reaction to deprotonate a water molecule attacking the glycosidic linkage. Therefore, the Glu-581 has been suggested as the general acid which protonates the α 2,8-ketosidic linkage. However, endoNF is also active at pH 7.4 and pH 8.0 [12, 105]. Under these conditions the glutamate is deprotonated and therefore cannot act as a general acid. Interestingly, in the region of the exosialidase tyrosine that acts as the catalytic nucleophile, enough space has been found that could harbor a water molecule during catalysis [106]. Like the carboxyl group of the Sia residue, Glu-581 could activate the water molecule and therefore act as a general base. Thereby another residue acting as general acid is required to protonate the glycosidic linkage upon attack of the activated water molecule.

This general acid function has been suggested to be provided by His-350 on the opposite site of the catalytic center [106]. The sialic acid trimer bound to the active site of endoNF-H350A is positioned with the reducing end close to the two active site residues Glu-581 and His-350 [106]. Histidines are known to function as proton switches between active site residues like serine and aspartates, as in catalytic triads. However, no such residues are found in endosialidases and, with a pK_a of 6.0, histidines only work at lower pH as general acids since they are mostly deprotonated above pH 7. Since endosialidases are also active at pH above 7, another residue is required to allow a proton transfer onto the glycosidic linkage. Tyr-325 buried in the hydrophobic core of the active site forms a hydrogen bond (2.9 Å) to His-350 in the wild-type endoNF structure. This tyrosine is highly conserved in all endosialidases and might change the electrostatics of His-350 to act as acid at higher pH. A proposed mechanism of polysialic acid cleavage at

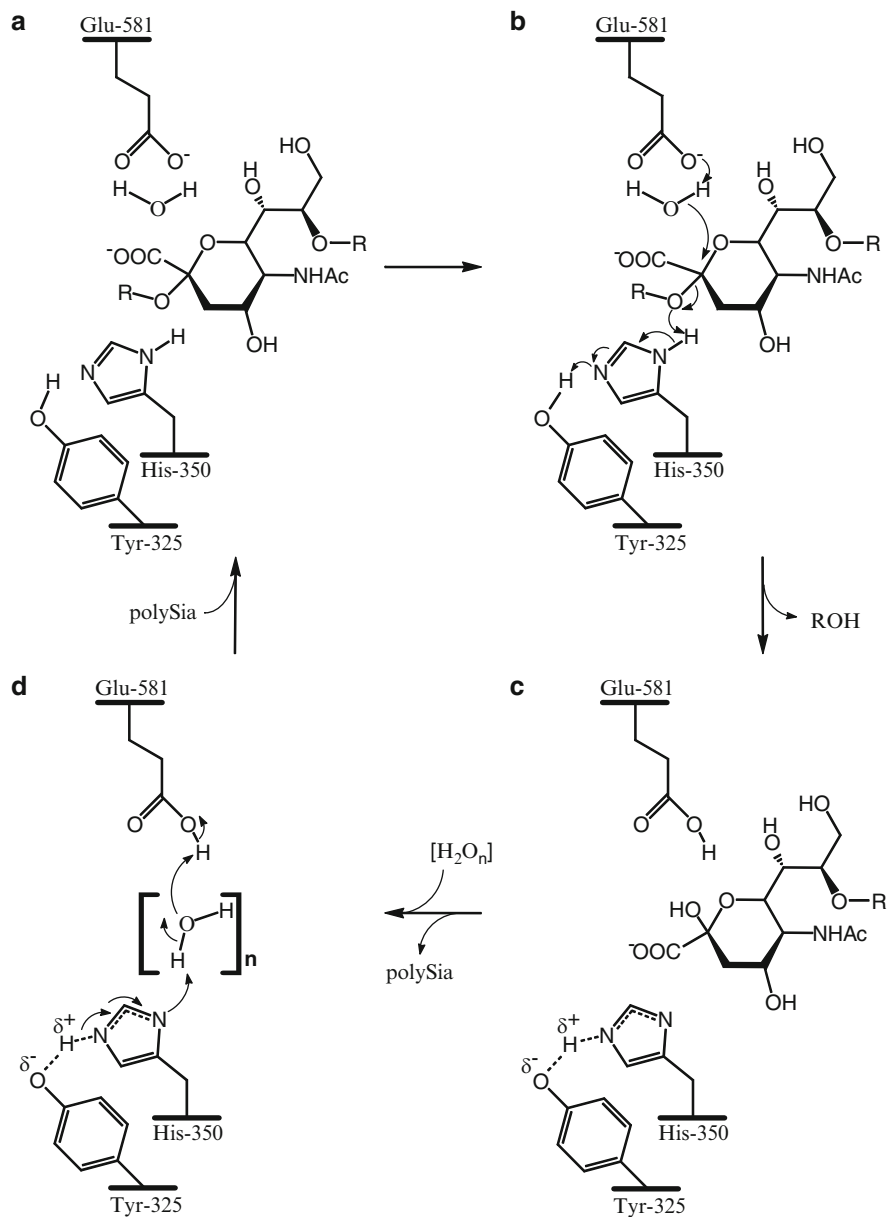


Fig. 4 Proposed mechanism for the cleavage of polysialic acid by an endosialidase at pH >7 based on the crystal structure of endoNF

higher pH is depicted in Fig. 4. After binding of the substrate (a), a water molecule is activated by the general base Glu-581 and attacks the glycosidic linkage. His-350 protonates the glycosidic linkages and the negative charge is delocalized towards

Tyr-325 (b). After release of the reducing end moiety (*ROH*, c) the initial state of the active site is regenerated by proton transfer via a water chain ($[\text{H}_2\text{O}]_n$, d) and re-binding of polysialic acid. In total, the structural data allow the conclusion that the endosialidase mechanism is an $\text{S}_{\text{N}}1$ -type reaction, i.e., the sialic acid stereochemistry is directly inverted into the β -anomer [106]. By contrast, exosialidases form a covalently bound intermediate which is released by a water molecule; thus the α 2,8-linked sialic acid residue is released as α -anomer underlying the mutarotation towards the energetically favored β -conformation. However, no crystal structure of a non-cleavable substrate is available as yet, which would allow an unambiguous elucidation of the endosialidase mechanism [106].

8 Oligosialic Acid Binding Sites

In soaking and co-crystallization experiments of endoNF variants with oligosialic acid of different length, two binding sites have been identified, *b1* and *b2*, that are located in the β -barrel domain and the β -prism of the stalk domain, respectively (Table 3). In the mutant endoNF-R647A, a pentamer and a tetramer of Sia are bound to *b1* and *b2*, respectively (*spheres* in Fig. 2c; [106]).

The sialic acids are coordinated by hydrogen bonds with side chains mainly of serine, arginine, threonine, and asparagine, as well as the backbone of alanine and glycine. In addition, hydrophobic interactions have been identified to valine, tyrosine, and threonine [12, 106]. Interestingly, the reducing end residue of the pentamer forms hydrogen bonds to the adjacent sialic acid residues and is not directly bound to *b1* [106]. In all cases the non-reducing end of oligosialic acid points towards the active site cleft of the adjacent subunit, emphasizing that endosialidases are functional trimers. Further, a lysine residue (Lys-410) from the β -propeller forms a hydrogen bond to the acetyl group of the non-reducing end sugar of DP3 in the endoNF-H350A mutant. Lys-410 has been suggested to bend the polysialic acid chain to allow access to the active site. Since lysines are highly flexible residues it is also possible that Lys-410 prevents dissociation of the chain during the processive degradation of polysialic acid (see Sect. 10). An exchange to alanine effectively reduces enzymatic activity to 20% and Lys-410 has been referred to as the active-site gateway [106].

In the first solved crystal structure, a sialic acid bound to site *b2* has been identified. It forms hydrogen bonds with the amino acids Arg-837, Ser-848, and Gln-853 [12]. Interestingly, although further residues have been identified to interact with dimeric and tetrameric sialic acid bound to *b2* of endoNF-H350A and endoNF-R647A, respectively [106], these three residues play a crucial role for proper binding of polysialic acid [105]. Except for the mutant endoNF-S848A, the two single mutations R837A and Q853A are sensitive to denaturation by SDS. A combination of the mutations R837A and Q853A further leads to completely insoluble protein. Potentially, these amino acids are involved in the folding and complex formation of the stalk domain (see Sect. 9).

Polysialic acid has been reported to undergo conformational changes in solution. Several left-handed helical epitopes with different pitches have been calculated and their relative occurrence in solution has been discussed [131, 132]. Polysialic acid bound to the cell surface adopts a similar average epitope as in solution [133]. A sialic acid decamer is bound in an extended helix to the monoclonal antibody 735 [134]. In contrast, a sialic acid pentamer has been described to bind endoNF in a compressed helical epitope [135] which has also been measured for polysialic acid in solution [132] and on the cell surface [133]. Therefore, polysialic acid selectively undergoes conformational changes upon binding to endosialidase or monoclonal antibody.

In the crystal structure of the mutant endoNF-R647A the two bound sialo-oligomers have been described to bind in two different left-handed helical conformations [106]. The pentamer in the *b1* site shows the compressed conformation with a helical pitch of 2.5 Å per monomer. In contrast, the tetramer in the *b2* site has been found in a more extended helical conformation with a helical pitch of 3.7 Å per monomer. The average helical parameters of the two epitopes perfectly fit those measured previously by NMR [131, 135]. The two different epitopes may reflect important requirements for a processive degradation of polysialic acid ([105, 106]; see Sect. 10).

9 Intramolecular Chaperone of Endosialidase

As described above (see Sect. 5, item 3), folding and trimerization of endosialidases crucially depends on an intramolecular CTD. The CTD has also been identified in other tailspike and fiber proteins of different bacteriophages [14, 105, 109]. The CTD fulfils the same function in all identified protein families regardless of the function of the N-terminal moiety. Moreover, the CTD can be exchanged between functionally unrelated proteins and fulfils its chaperone function leading to functional N-terminal proteins. However, an exchange of highly conserved amino acids within the CTD into alanine or complete deletion of the CTD results in insoluble proteins [76, 112]. In all investigated protein families with an endosialidase-like CTD the chaperone domains are proteolytically released at a highly conserved serine residue. An exchange of this serine (endoNF: Ser-911) into alanine completely abolishes proteolytic cleavage but does not inhibit the chaperone function [14, 112]. The requirements for the proteolysis will be discussed below.

Several trimeric β -structures similarly require C-terminal trimerization domains (for an overview see [136]). In all cases these domains prevent nonstoichiometric aggregation of the hydrophobic protein cores. In endosialidases the sialidase domain folds independently in the three subunits while the stalk domain is composed of an intertwined triple β -helix which requires a complex folding procedure [105, 136]. Recently, a crystal structure has been solved which comprises part of the stalk domain and the CTD of endoNF. The release of the CTD in this construct has been prevented by introducing the mutation S911A into the endoNF-stalk

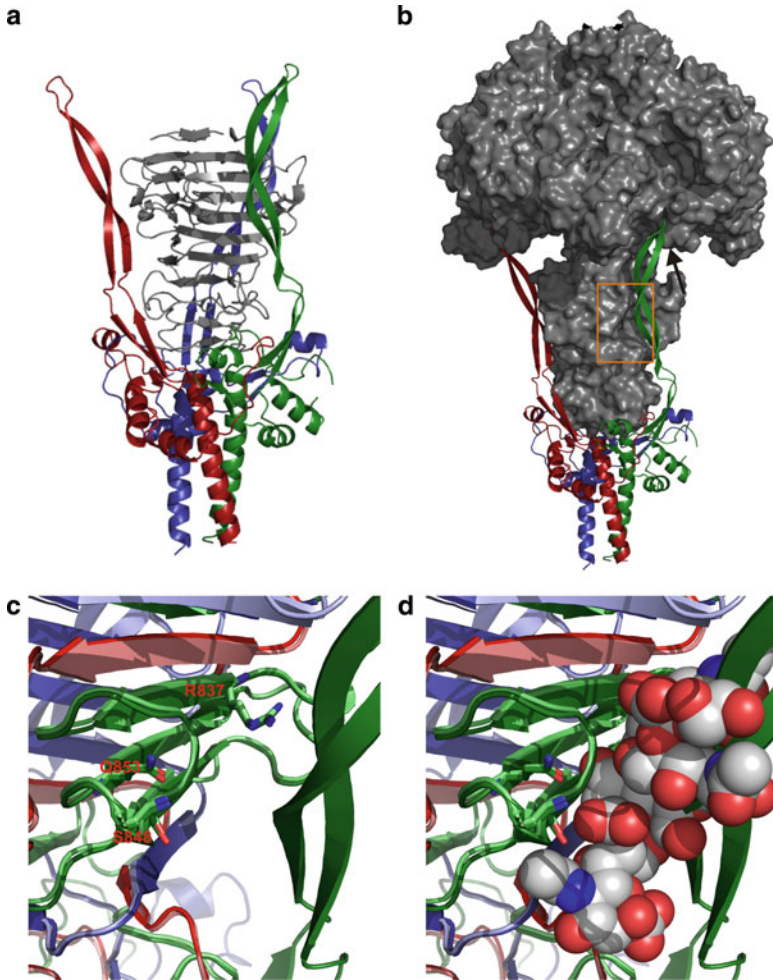


Fig. 5 The intramolecular chaperone domain [109]. **(a)** Crystal structure of the truncated non-cleavable endoNF (premature endoNF-stalk; S911A-mutant of amino acids 800–1,064). The stalk region is shown in *gray* and the three subunits of the chaperone domain in *blue*, *red*, and *green* in the cartoon representation. **(b)** Superposition (Coot SSM superpose [107]) of the crystal structures of the premature endoNF-stalk (*ribbon diagram*) and the mutant endoNF-R647A (*gray surface* representation). The *arrow* indicates “clashing” of the tentacle tip with the β -propeller in the superposed structures. The *orange rectangle* marks the area of the binding site *b2* enlarged in **c** and **d**. **(c, d)** Ribbon diagram of the sialic acid binding site of superposed premature endoNF-stalk and endoNF-R647A with sialic acid tetramer (*spheres*) bound to the binding site *b2* of endoNF-R647A (**d**). In **(c)**, the oligosialic acid is hidden. The subunits of premature endoNF-stalk are given in the same color as in **(a)**, whereas the mutant endoNF-R647A is represented in the respective *pale color*. The amino acids involved in binding oligosialic acid (R837, S848, Q853 [105]) are shown as *sticks*. The oligosialic acid is bound to endoNF at the same location as the tentacle of the chaperone domain of the premature endoNF-stalk (**d**)

(premature endoNF-stalk; [109]). The structure of the stalk domain is unaltered compared to the wild type endoNF and is given in *gray* in Fig. 5a. The chaperone domain is reminiscent of a jelly fish in an “upside down” orientation relative to the “mushroom” (cf. Fig. 5b). The subunits of the CTD trimer are given in *red*, *blue*, and *green*. The core of the CTD is predominantly α -helical, which is in clear contrast to the predominantly β -structure containing mushroom of endosialidases (cf. Fig. 2a). From each subunit, a tentacle of two twisted anti-parallel β -sheets protrudes towards the N-terminus of the stalk. In the superposition with the wild type mushroom, the tentacle tips clash with the β -propeller (*arrow* in Fig. 5b). Therefore, the conformation of the tentacles depicted in Fig. 5 is in part artificial compared to a full-length protein [106].

The tentacles of various proteins differ in length, whereas the core topology of the CTD is highly conserved [109]. For example, in a sequence alignment the tentacles of endoNF are substantially longer than in endoNE. A previous domain swap experiment revealed that the catalytic part of endoNE is properly folded into an active enzyme when fused to the longer CTD of endoNF (endoNE-NF), but proteolytic release has not been observed. However, a domain exchange in the other direction yields the insoluble construct endoNF-NE although the catalytic parts of both enzymes share 88% sequence similarity [14].

Based on biochemical data a folding process of endosialidases has been suggested [105] which has been adapted using newly acquired structural data [136]. In summary, the trimerization of the CTD most likely is initiated by the α -helical core of the CTD whereas the tentacles tightly interact with a shallow groove of the stalk domain. Further, the tentacle tip potentially interacts with the β -propeller (*arrow* in Fig. 5b) which might be the reason for the need for the long tentacles in endoNF for proper intertwining of the triple β -helix [109, 136]. After completion of the complex folding, the CTD is proteolytically released from the mature enzyme. Five amino acids have been identified to be required for proteolysis. First, Arg-897 forms hydrogen bonds with Asp-886 and Tyr-910. Since the arginine residue and the other two amino acids are from different subunits this interaction can only occur after proper folding. Arg-897 by this has been referred to as folding sensor [109]. Proteolysis occurs by an intramolecular Lys–Ser–dyad reaction. Lys-916 from the CTD functions as general base to activate the Ser-911 at the cleavage site of the same subunit. Since Lys-916 is buried in the interior of the protein, it can be assumed that it is in the deprotonated state to act as general base. The peptide backbone forms an unusual S-shape to allow proximity of the two residues. The activated serine residue acts as a nucleophile, attacking its peptide bond to Tyr-910. Additionally, Arg-897 has been suggested to form the oxanion pocket to stabilize the transition state [109]. A similar CTD from the neck appendage protein gp12 of *Bacillus* phage phi29 has been crystallized. The gp12 is a trimeric tailspike protein composed of β -helices. The CTD is predominantly composed of β -sheets, in contrast to the CTD of endosialidases. Proteolytic release uniquely requires an ATP molecule to provide the required energy [137].

Proteolytic maturation is a common theme among phage proteins involving either phage-derived proteases or autocatalytic processes [32, 138–141]. In contrast to pro-enzymes like mammalian digestive serine proteases, the function of

proteolytic maturation of phage proteins is not restricted to enzyme activation. In the case of the major head proteins of coliphage T4, cleavage is required for head expansion and DNA packaging [142]. For endoNF it has been shown that release of the CTD is not required for activation of the enzyme [14, 112]. However, certain restraints of the enzyme have been identified when the CTD remains part of the complex [105, 112]. The two following possible requirements for proteolytic release of the CTD have been discussed:

1. The non-cleavable S911A mutant of endoNF shows an unusual electrophoretic migration in the presence of SDS. This additional band represents a folding intermediate with lower electrophoretic mobility and appears at elevated temperatures before dissociation into monomers. The non-cleavable S911A mutant is thus more sensitive to SDS than the wild type. The catalytic part released from wild type requires higher temperatures for dissociation in the presence of detergent and has been suggested to build a kinetically stable complex by release of the CTD [112]. Since bacteriophage proteins are exposed to harsh extracellular conditions like high salt concentrations and variations in temperature and pH, their proteins require an unusual stability in these environments. Frequently, kinetically stable proteins are, like endosialidases, mainly composed of β -structures. These proteins are stable against protease digest and have unusual longevity [123, 143]. For endosialidases and proteins with a similar CTD, the proteolytic release of the CTD leads to the establishment of a kinetically stable trimer [112].
2. Another study has revealed that presence of the CTD in the mutant S911A severely impacts binding to polysialic acid and, as a consequence, leads to an impaired degradation of surface bound polysialic acid. However, S911A shows an increased activity on soluble polysialic acid since the reduced binding to polysialic acid allows faster dissociation of the substrate and re-association with a new polysialic acid chain [105]. By combinatorial mutagenesis of amino acids Arg-837, Ser-848, and Gln-853 (Fig. 5c; [12]) of the binding site *b2*, the effects observed for S911A can be reproduced to a similar extent. In total, it has been shown that the presence of the CTD or the mutation of *b2* drastically diminishes processive degradation of polysialic acid (see Sect. 10; [105]). The crystal structure of the CTD (cf. Fig. 5a) reveals that the tentacles are in close proximity to the binding site *b2* in the β -prism domain (Fig. 5b–d) and thereby prevent binding of polysialic acid [109]. Additionally, the structural data corroborate the importance of the three identified amino acids [12] of site *b2* for proper folding [105].

In total, the CTD is released after finishing its job to kinetically stabilize the trimeric complex and to allow proper binding to polysialic acid, which is an important prerequisite for processive degradation of polysialic acid.

10 Processivity of Endosialidase

Electron micrographs of phages attacking encapsulated bacteria have revealed that penetration of the capsule by the phage particle mostly occurs in a processive and unidirectional manner leading to the formation of a narrow tunnel [36, 37, 144].

It has been discussed whether this processive penetration is due to an interplay of the six copies of endosialidases attached to the phage particle [13] or whether the tailspike protein per se exhibits processivity [145]. In extensive biochemical experiments with endoNF it has been shown for the first time that a bacteriophage tailspike enzyme is processive by itself in vitro [105]. Anion exchange chromatography of polysialic acid and its cleavage products has been used to show that, after initial binding to the polymeric substrate, endoNF apparently migrates towards the reducing end of polysialic acid and predominantly releases short oligosialic acid units with a degree of polymerization of three or more while the average chain length of long-chain polysialic acid changes slowly. This is a characteristic feature of processive depolymerases like chitinases or cellulases [146–150]. In these enzymes, the substrate-binding clefts are lined with aromatic residues that are thought to facilitate sliding of the polymer chain through the cleft during a processive mode of action. As shown for chitinase B from *Serratia marcescens*, single alanine substitution of aromatic residues in the binding cleft reduces processivity and thereby the ability to degrade crystalline chitin, but increases the degradation rate for soluble substrate [151]. As stated above (Sect. 9), the non-cleavable mutant endoNF-S911A shows an impaired binding to polysialic acid due to covering the *b2* binding site by the tentacles of the CTD and, as a consequence, the enzymatic activity is influenced (Fig. 5c, d; [105, 109]). Like the mutants of chitinase B, the mutants of endoNF with a covered or mutated binding site *b2* have shown a drastically reduced activity on surface bound substrate but an increased activity with soluble polysialic acid. Remarkably, the impaired *b2* binding site also dramatically affected the processive degradation, emphasizing the importance of site *b2* for endosialidases and consequently for K1-specific bacteriophages to attach effectively the phage particle to the remnant polysialic acid chains on the bacterial cell surface [12, 105].

The exact mechanism of processive polysialic acid digestion is unknown. Since endoNF contains three polysialic acid interaction sites (*a*, *b1*, and *b2*, Fig. 6a, cf. Fig. 2c), of which *b2* is important for the processive degradation (cf. Fig. 5c, d; [105]), and *b1* and *a* are in direct proximity, allowing binding of one polysialic acid chain simultaneously, an interplay of these three interaction sites has been suggested to be the key to processive degradation [105, 106]. Further, the active site of inactive endoNF-H350A binds a sialic acid trimer [106] which also resembles the major cleavage product released from endoNF [105]. The two different helical epitopes identified at binding sites *b1* and *b2* of endoNF-R647A (see Sect. 8) might be essential for processive migration of the polysialic acid chain [106]. However, processivity per se is a unidirectional movement which can only occur in a cyclic process with an irreversible step.

The combination of recent data obtained from crystal structures [106, 109] and biochemical data [105] allows postulating a cyclic mechanism of endoNF processivity (Fig. 6b). The structural data (Fig. 6a) are simplified to ellipses representing the interaction sites and a *ball* representing one Sia molecule (Fig. 6b). For the processive endoNF wild type, four states are conceivable, in which polysialic acid is bound differently (1–4). Additionally, in state 0, all

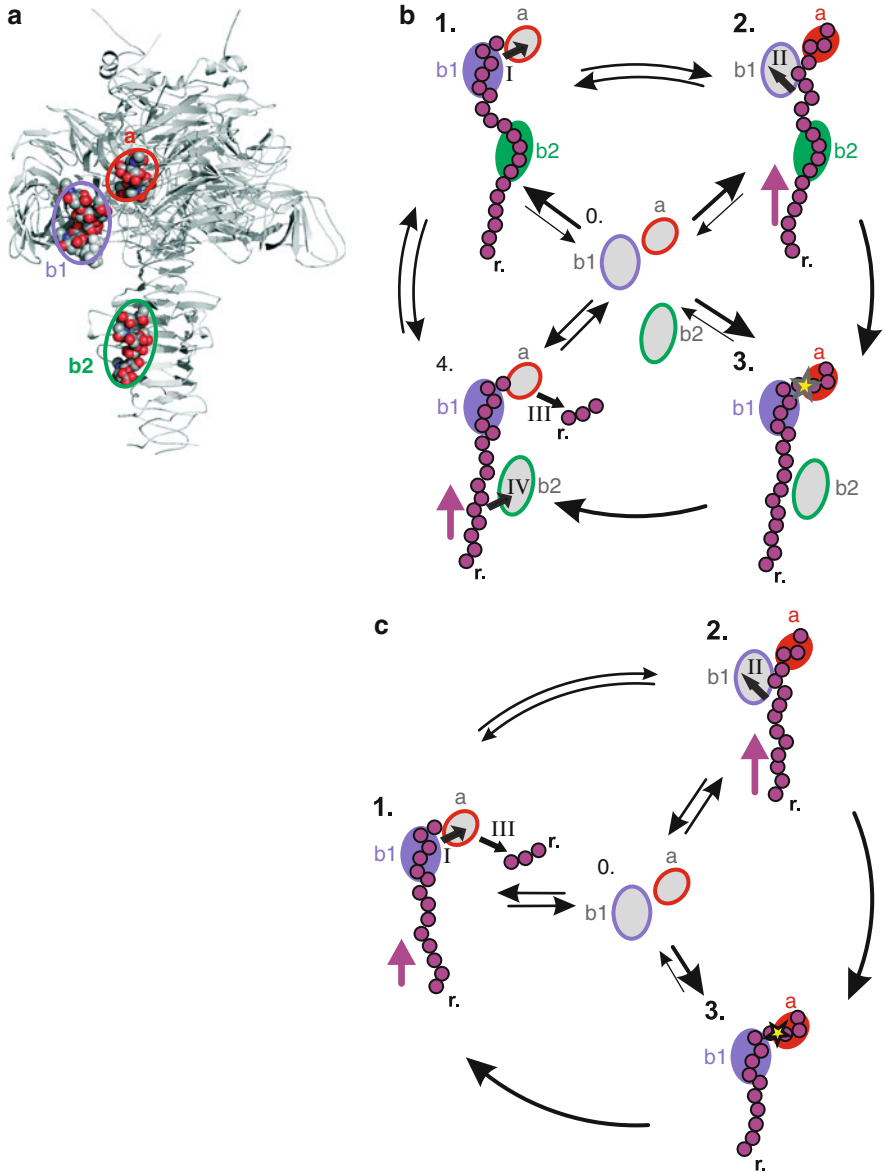


Fig. 6 Suggested processivity mechanism of endosialidase. (a) Ribbon diagram of endoNF-R647A (gray) with bound sialic acid oligomers (spheres, cf. Fig. 2c). Sialic acid trimer bound to the active site *a* (red ellipse) of endoNF-H350A, sialic acid tetramer and pentamer bound to the binding sites *b1* (blue ellipse) and *b2* (green ellipse) of endoNF-R647A, respectively. Oligosialic acid is given in element color code, whereas the terminal residue at the reducing end of DP5 in the *b1* site is given in gray, since it is not directly coordinated by endoNF. (b) Suggested cyclic processivity mechanism of wild-type endosialidases. The interaction sites from *A* are depicted as ellipses in the respective color. Sia residues are given as purple circles with the reducing end depicted as *r*. Unoccupied binding sites are filled in gray. (c) Cyclic non-processive mechanism for endosialidase mutants with an impaired binding site *b2*

polysialic acid interaction sites of endoNF are idle (*gray filling*). Spontaneous binding of a polysialic acid chain to the enzyme can lead to any of the four states. *Arrows* mark transitions between the states, whereas the thickness qualitatively resembles the possibility of the transition. A possible fifth (super-) state is not shown, in which all three sites are occupied simultaneously. The state with occupied sites *b1* and *b2* may mark a start of the cycle (*I*). By diffusion, the polysialic acid chain changes its binding state, depicted by *thick arrows* with *Roman numbers (I–IV)*. A net “upward” movement is depicted by *purple arrows* with different lengths. Note that the reverse movement from state 2 to 1 leads to a net “downward” movement (not shown). Movement *II* potentially leads to the conformation of polysialic acid required for catalysis (*yellow asterisk*). The cleavage reaction prevents the back movement to state 2 and therefore drives the cycle into one direction – this resembles the irreversible step written above. States 3 and 4 could be combined since movement *IV* could happen simultaneously with cleavage. However, in state 4 the dissociation of the small cleavage product (here a sialic acid trimer) is a second irreversible step. From all states (*I–4*), polysialic acid may dissociate from endoNF leading to step 0. However, when polysialic acid is bound to two different binding sites as illustrated in states 1–3, dissociation is less likely than in state 4, in which polysialic acid is bound to only one site. Since polysialic acid is a helical structure, the diffusion from one site to the other requires a reorientation of the chain if the movement does not occur by one turn (“in frame”). This can lead to torsion tensions that may facilitate release from another binding site. Finally, the compressed helical epitope [106] in site *b1* could expand to a more stretched conformation facilitating movement into the active site. This movement might be delegated by the active-site gateway residue Lys-410 [106].

In Fig. 6c a non-processive endoNF – caused by an impaired *b2* binding site [105] is shown. Deduced from Fig. 6b, the lack of binding site *b2* leads to the reduction of possible binding modes. State 1 of Fig. 6b with polysialic acid attached to two binding sites is not possible. In addition, since polysialic acid in states 1 and 2 of Fig. 6c is bound by only one binding site, dissociation is more likely in non-processive endoNF. Only in state 3 is polysialic acid bound by sites *b1* and *a* during the cleavage reaction. In total, a higher possibility for dissociation of polysialic acid may be the explanation for the loss of processivity in *b2*-defective mutants.

11 Applications of Active Endosialidase

11.1 Preparation and Properties of Catalytically Active Endosialidase

E. coli K1 specific bacteriophages have been tested for phage-typing of bacterial strains [152], but this application has not found much use. It was soon discovered that the bacteriophage-derived endosialidase is suitable for the study of eukaryotic

polysialic acid derived from neural tissue [75, 102]. Since then, endosialidase has been used in numerous studies related to different aspects of eukaryotic polysialic acid.

Crude bacteriophage lysate, purified phage particles [72, 74, 102], or endosialidase purified from bacterial lysate [75, 101, 153] have been used in early studies on polysialic acid. Originally endosialidase purification has involved ammonium sulfate precipitation as well as hydrophobic, gel filtration, and DEAE-ion-exchange chromatography [101, 104]. More conveniently, endosialidase is produced as a recombinant enzyme as His6-tagged, GST-tagged, and MBP-tagged fusion proteins [14, 71, 154].

The pH optimum of endoNE has been reported to be slightly above 5 [101]. Using a TFMU sialotrioside substrate (TFMU α -linked trifluoromethylumbelliferyl glycoside) and monitoring the release of the free coumarin endoNF has been found to have a pH optimum around 4.5 [127]. Determination of $k_{\text{cat}}/K_{\text{M}}$ at a series of pH values has revealed a dependence on a single protonated group of $\text{p}K_{\text{a}}$ 5. This observed $\text{p}K_{\text{a}}$ has been proposed to belong to an enzymic carboxylic acid, Glu-581, one of the putative catalytic residues. The enzyme is unstable below pH 4 and above pH 12.

Acetate is preferred over phosphate as the buffer anion, and the catalytic activity is reduced in the presence of 9 mM Ca^{2+} [101]. However, it should be noted that endosialidase is still active under physiological conditions in the presence of phosphate and calcium. Polyanions, like DNA, have been reported to inhibit the activity of endoNF [104].

The K_{M} values for long, \sim DP200, α 2,8-linked sialyl polymers are significantly lower (K_{M} 50–70 μM) than for short, \sim DP10–20, sialyl oligomers (K_{M} 1.2 mM) for K1F endosialidase [104]. A K_{M} value of 7.4 mM has been reported for endoNE, assuming a DP value of 175 [101]. The minimal substrate of endoNF is DP4 and shows a K_{M} of 0.85 mM and k_{cat} of 262.2 min^{-1} [105]. Using the artificial TFMU sialotrioside substrate, the values for K_{M} , 0.68 mM and k_{cat} , 77 min^{-1} , have been recorded [127]. Unexpectedly, the alternating α 2,8/ α 2,9-linked sialyl polymer with an apparent K_{M} of 6.6 μM appears to be a better substrate than the α 2,8-linked polysialic acid [104]. Only the α 2,8-linkages in this polymer, however, are hydrolyzed. The K_{M} for polysialic acid of the phage-bound and free form of endoNF had no significant differences.

The active endosialidase with its ability to specifically degrade polysialic acid represents a powerful tool for the study of the chemistry, metabolism, and biological roles of polysialic acid. The inactive endosialidase, on the other hand, able to recognize specifically and remain bound to polysialic acid, represents a sensitive and convenient tool for the specific detection of polysialic in different applications.

11.2 Applications of Active Endosialidase for the Study of NCAM

Identification of polysialic acid as posttranslational modification of NCAM has been based on the release of oligomers of sialic acid from it by incubation with

endosialidase, and the resulting dramatic increase in gel electrophoretic mobility [153]. Similarly, polysialic acid containing glycopeptides isolated by proteolytic degradation show reduced size in gel filtration [102]. The endosialidase-released oligomers of sialic acid have been detected with various methods including ion-exchange chromatography after radioactive labeling [75], thin-layer chromatography [102], polyacrylamide gel electrophoresis [155], high-performance liquid chromatography [156], and mass spectrometry [157].

Endosialidases have been a key towards studies aiming at deciphering the role of NCAM in biological interactions. NCAM is a glycoprotein that forms cell–cell contacts via direct interaction between NCAMs on each cell. Early studies that used endoNF *in vitro* and *in vivo* applications have provided the first evidence that the adhesion between living cells may be regulated by the polysialic acid chains on NCAM [153]. These observations have been substantiated in studies on knockout (KO) mice [158]. Endosialidase has been a helpful tool in studies on the polysialylation process of NCAM. The evaluation of the contribution of two polysialyltransferases – ST8SiaII (STX) and STSiaIV (PST) – to *in vivo* polysialic acid synthesis on NCAM has been carried out with brain samples obtained from newborn wild-type, ST8SiaII-deficient, and STSiaIV-deficient mutant mice. EndoNF and endoglycosidase PNGase F have been used to treat polysialic acid containing glycopeptides and reveal a comparable quality of polysialylation by ST8SiaII and ST8SiaIV and a distinct synergistic action of the two enzymes in the synthesis of long polysialic acid chains at N-glycosylation site 5 [159].

Immunohistochemistry with the specific polysialic acid recognizing antibodies, especially mAb 735 [160], and with polysialic acid α 2,8-cleaving endosialidase have allowed the detection and identification of polysialic acid on different cells and tissues and to estimate its function and role. Polysialic acid plays an important role in nervous system phenomena such as cell migration, axonal guidance, synapse formation, and functional plasticity [161]. As a long (steric) and highly negative charged molecule, polysialic acid can modify adhesion and binding properties of other molecules, receptors, and cells.

The stability of endosialidase and its activity under physiological pH and ion conditions have promoted its use under various experimental setups *in vivo*. In early studies it has already been shown that injection of endosialidase into the brain during development results in altered responses of developing axons to their environment.

EndoNF injection into the fourth ventricle of P0 rats results in the reduction of collateral branching of corticospinal axons [162]. Injection of endosialidase NF into the lateral ventricle of P1 mice results in rapid diffusion of the enzyme throughout the brain. The removal of polysialic acid induces aberrant mossy fiber innervation and ectopic synaptogenesis in the hippocampus [163].

Removal of polysialic acid by treatment with endosialidase has demonstrated the involvement of polysialic acid in dynamic cellular processes as varied as migration of neuronal precursor cells, axonal outgrowth, synaptogenesis, physiological and morphological synaptic plasticity, and control of circadian rhythm [63, 75, 153, 164–174]. One example on the application of endosialidase in cell culture studies is

the use of the enzyme in the study of the myelination process, especially timing and localization of myelination and the role of astrocytes in controlling the onset of myelination [175, 176]. The capacity of O-2A oligodendrocyte progenitors to migrate in cell culture is completely blocked by treatment with endosialidase [177]. On the other hand, astrocytes promote the adhesion of oligodendrocyte processes to axons, and endosialidase treatment closely mimics the effects of astrocytes [178].

EndoNF has been established as a unique tool to degrade specifically artificial polysialic acid hydrogels and other polysialic acid derivatives in a biocompatible and induced manner [179–182]. In a mouse femoral nerve regeneration model it has been found that if endosialidase is injected into the lesion site polysialic acid reexpression in the regenerating axons is abolished. At the same time, the regeneration accuracy is hampered, which indicates a role of polysialic acid in nerve regeneration [183]. In future practical applications of polysialic acid as scaffold of neural regeneration, endosialidase may become useful as an agent of specific degradation of the scaffold to avoid second surgery [180].

Although the vast majority of polysialic acid in the nervous tissue is carried by the NCAM molecule, it is not certain whether all the effects recorded by the use of endosialidase are specifically due to its action on NCAM or an action on other molecules like SynCAM 1 [63].

11.3 Other Applications of Active Endosialidase

In adult rat brain NCAM is not the only polysialylated molecule. The sodium channel glycoprotein has been reported to be polysialylated in its α subunit [59]. The significance of the negative charge on the sodium channel is not clear, but desialylation decreases the conductivity of the channel. α 2,8-linked polysialic acid chains are also carried by sodium channels on the electrical organ of the eel *Electrophorus electricus* [184].

By studying the brain of newborn NCAM knock-out mice, another polysialylated glycoprotein, SynCAM 1, has been identified [63]. This molecule belongs to the Ig superfamily and is in homophilic and heterophilic interactions involved in synapse formation. Polysialylation of SynCAM 1 abolishes its homophilic binding. The multifunctional glia population NG2 cells expresses the polysialylated form of SynCAM 1. These cells receive glutamatergic input and their integration into neural network might be modulated by polysialylation of SynCAM 1.

CD36 is a member of the class B scavenger receptor family of cell surface proteins. In humans, CD36 is present, for example, in milk and in platelets, but only the CD36 of milk is polysialylated, on O-linked glycan chain(s) [60]. The polysialic acid CD36 is secreted in milk during lactation. It has been suggested that this polysialylation is important for neonatal development in terms of protection and nutrition. EndoNF and monoclonal antibodies mAb 12E3 (IgM) recognizing oligo/

polyNeu5Ac with a DP ≥ 5 and mAb S2-566 (IgM) recognizing the Neu5Ac α 2,8 Neu5Ac α 2,3Gal sequence have been used for the detection and identification of CD36 polysialylation in milk.

The biological roles of polysialic acid outside the nervous system have been largely unexplored. Polysialic acid has been found on NCAM in human NK cells, mouse hematopoietic progenitors, and myeloid cells. Natural killer (NK) cells express an isoform of NCAM (CD56, Leu-19, NKH1), which is polysialylated [61, 185, 186]. Neural polysialic acid participates in migration, cytokine response, and cell contact-dependent differentiation. The immune system needs the same functions, and polysialic acid appears to be involved in the development of immunological response and host defense. Human NK cells modulate NCAM expression and the degree of polymerization of its polysialic acid according to their activation state. On human immune cells the size of polysialic acid chains is comparable to that of glycosaminoglycans, which have important functional and structural roles in the extracellular matrix. Transferase-deficient ST8SialIV^{-/-} mice (lacking polysialic acid expression in the immune compartment) demonstrate an increased contact hypersensitivity response and decreased tumor growth control.

Furthermore, glycans of neuropilin-2 carry O-linked polysialic acid [62]. Neuropilin-2 (NRP-2) is a receptor expressed on the surface of human dendritic cells. Two isoforms of neuropilin-2, NRP2a, and NRP2b, are both polysialylated [187]. The expression of NRP-2 is increased during dendritic cell maturation simultaneously with ST8SialIV expression. Removal of polysialic acid from the surface of dendritic cells promotes dendritic cell-induced activation and proliferation of T lymphocytes. Polysialic acid controls CCL21-directed migration of mature dendritic cells, but doesn't enhance chemotaxis towards CCL19 [188]. When dendritic cell polysialylation is prevented, the CCL21 activation of the JNK and Akt signaling pathways is decreased. The enhancing effect of polysialic acid on dendritic cell chemotaxis is dependent on the highly basic C-terminal region of the CCL21 molecule.

Immunohistochemical studies relate polysialic acid-NCAM expression to tumor malignancy in neuroblastoma, alveolar rhabdomyosarcoma [189], neuroendocrine lung tumors [190], Wilm's tumor [191], aggressive colon cancers [192], and pancreatic cancer [193]. In studies on the role of polysialic acid in tumor cell behavior, endosialidase has been used as a tool to remove polysialic acid from the surface of cells and tumors. In a metastasis model using human rhabdomyosarcoma TE671 cells, endoNA injected intraperitoneally reduces the expression of polysialic acid and the number of lung and liver metastases formed from intraperitoneal primary tumors [194]. For comparison, the effect was not seen with intramuscular primary tumors that cannot be reached by the intraperitoneally injected endosialidase.

Endosialidase is also of interest as therapeutic agent. For instance, it has been described as an efficient agent in the treatment of septicemia and meningitis in *E. coli* K1 infected rats [195]. Small quantities of recombinant endoNE injected intraperitoneally into newborn mouse pups to remove polysialic acid from the bacterial surface are able to prevent the development of bacteremia and death in an experimental

model [196]. The experiment demonstrates both the therapeutic and the prophylactic potential of the enzyme. The polysialic acid capsule of *E. coli* K1 is integral to the virulence of these bacteria. The low immunogenicity of the capsule varies due to random *O*-acetylation. *O*-Acetylated polysialic acid is more resistant to degradation by neuraminidases [100]. Degradation of *O*-acetylated polysialic acids by endoNE has been used as tool in the identification of bacterial genes that are involved in the modification of the capsule [55].

12 Catalytically Inactive Endosialidase

12.1 Generation and Properties of Inactive Endosialidase

Using *E. coli* mutants with a sparse polysialic acid capsule, host range mutants of bacteriophages K1A and K1E have been isolated [15, 197]. These mutants contain endosialidases which are catalytically partly or totally inactive, but still retain polysialic acid binding. The catalytically inactive enzymes have proven to be valuable, since they can still bind to the substrate without degradation, and can therefore be used as tools for polysialic acid identification [166, 168]. Further, by structure-based mutagenesis inactive variants of endoNF have been generated which bind polysialic acid with an apparent K_d of less than 2 nM [105].

The aligned sequences of the endosialidases of phages K1A and K1F (endoNF) have 83% identity [69], with the major domains having even higher identity (β -propeller sequences 93% and the β -barrels 100%). A homology-based structural model of the endosialidases predicts the spatial arrangement of the mutant amino acids close to each other at the active site of the enzyme [69] (Fig. 2b). The mutations affect the orientation of the residues and the shape of the active site cleft.

Four spontaneous point mutations (W118C, H332N, H417Y, N489D, see Sect. 7 and Table 4) implicated in reduced activity of three endoNA mutants are located on or near the active site. Each observed mutant has N489D, which alone does not cause complete inactivation [69]. Interestingly, in the endoNF structure the equivalent Asn-699 is spatially located next to the His-350. Natural mutation of endoNE, A370E (see Sect. 7), decreases the catalytic activity to one tenth of that of native endoNE [69].

Influenza neuraminidases have residues around the catalytic cleft that are invariant between different strains of influenza virus neuraminidases [198]. Eight of these residues are functional which suggests that they contact the substrate directly, but another ten residues stabilize the active site structure and are referred to as “framework residues” (cf. Fig. 3b). Analogously, the spontaneous point mutations of endoNA [15, 69] are part of a framework rather than being directly involved in

catalysis (Fig. 3c). While evolutionary pressure is strong on the catalytic residues, variation of the framework can still adjust the specificity of the enzyme.

12.2 Use of Inactive Endosialidase for the Detection of Polysialic Acid

Purified whole phage particles with catalytically inactive endosialidase have been used for the specific detection of polysialic acid on dot blots of bacterial cultures, Western blots of tissue extracts, as well as fluorescence and light microscopy of cells and tissue sections [166]. The phage mutants recognize polysialic acid, and the bound phages have been detected with an anti-phage antibody followed by a peroxidase-conjugated or fluorescently labeled secondary antibody [166]. This method can thus be used for the specific detection of both bacterial and eukaryotic polysialic acid, but it includes multiple steps and needs time-consuming propagation of phages.

Endosialidases have been engineered to bind specifically α 2,8-linked polysialic acid as inactive variants for fluorescence microscopy [69, 105, 168]. The endosialidase gene of the mutant phage K1A2 has been fused to green fluorescent protein (GFP). A repeat part of the *Yersinia enterocolitica* adhesin (yadA) stalk [199] was added between the endosialidase and GFP to increase the yield of soluble protein. A fusion protein containing the catalytically active form of the endosialidase prepared for control was negative [168].

The binding affinity (K_d) of widely used molecular markers usually ranges from millimolar to micromolar. Several anti-carbohydrate antibodies have been determined to have K_d in the range from micromolar to 10 nM [200]. The best-known polysialic acid antibody, mAb 735, has K_d of 7 nM [201]. The apparent K_d of inactive endoNF (R596A/R647A) has been determined to be 1.9 nM [105]. The fusion protein of endoNA has been determined to have K_d of 19 nM [168]. Therefore, the endoNA fusion protein has a slightly lower binding affinity than the uncoupled inactive endoNF, but it still has an excellent binding affinity for polysialic acid.

One example of the applications of the inactive endosialidase for polysialic acid detection is the staining of neuroblastoma cells (Fig. 7). The fusion protein with an inactivated endosialidase has also been used for the detection of polysialic acid in cells, bacteria, paraffin-embedded tissue sections, frozen sections, fluorescence counting, immunoblots, and flow cytometry [168, 203, 204]. The fusion protein represents a sensitive and specific detection method of polysialic acid. A further advantage is the convenience of use as a single-step reagent. A potential cross reaction concerns the rare *E. coli* K92 polysaccharide with the alternating α 2,8/ α 2,9 sialic acid linkages. Such polymers have not been found in eukaryotic samples so far.

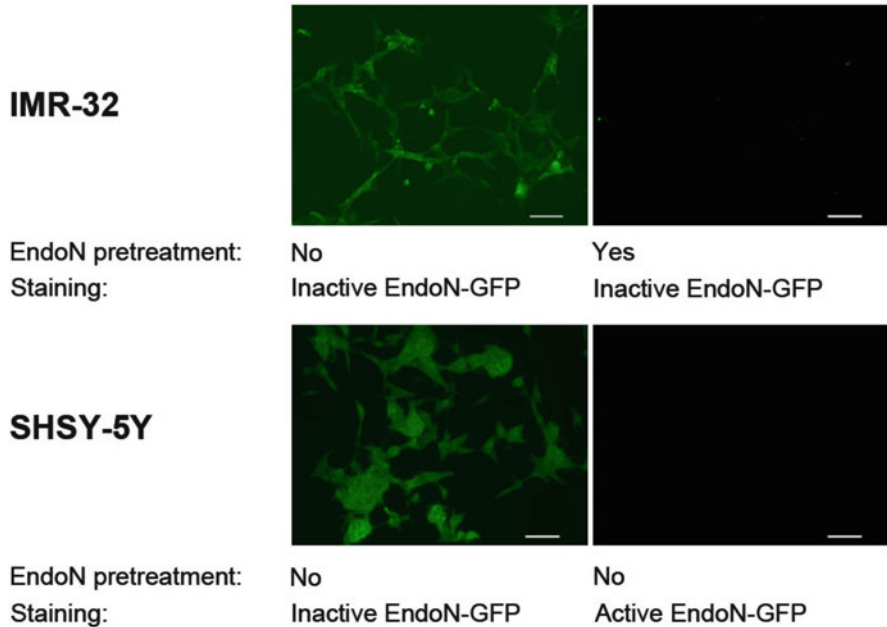


Fig. 7 Detection of polysialic acid in neuroblastoma cells with a fusion protein of endosialidase and GFP. Paraformaldehyde-fixed IMR-32 (ATCC-CCL-127) and SH-SY5Y cells (European Collection of Cell Cultures ECACC 94030304; [202]) were stained with a fusion protein containing inactive or active endosialidase without pretreatment or after pretreatment with active endosialidase as indicated, and detected by fluorescence microscopy (scale bar 20 μ m)

References

1. Barry GT (1958) Colominic acid, a polymer of N-acetylneuraminic acid. *J Exp Med* 107:507–521
2. Reglero A, Rodríguez-Aparicio LB, Luengo JM (1993) Polysialic acids. *Int J Biochem* 25:1517–1527
3. Drake PM, Nathan JK, Stock CM, Chang PV, Muench MO, Nakata D, Reader JR, Gip P, Golden KPK, Weinhold B, Gerardy-Schahn R, Troy FA, Bertozzi CR (2008) Polysialic acid, a glycan with highly restricted expression, is found on human and murine leukocytes and modulates immune responses. *J Immunol* 181:6850–6858
4. Finne J, Finne U, Deagostini-Bazin H, Goridis C (1983) Occurrence of alpha 2–8 linked polysialosyl units in a neural cell adhesion molecule. *Biochem Biophys Res Commun* 112:482–487
5. Rutishauser U (1998) Polysialic acid at the cell surface: biophysics in service of cell interactions and tissue plasticity. *J Cell Biochem* 70:304–312
6. Tanaka F, Otake Y, Nakagawa T, Kawano Y, Miyahara R, Li M, Yanagihara K, Nakayama J, Fujimoto I, Ikenaka K, Wada H (2000) Expression of polysialic acid and STX, a human polysialyltransferase, is correlated with tumor progression in non-small cell lung cancer. *Cancer Res* 60:3072–3080

7. Seidenfaden R, Krauter A, Schertzinger F, Gerardy-Schahn R, Hildebrandt H (2003) Polysialic acid directs tumor cell growth by controlling heterophilic neural cell adhesion molecule interactions. *Mol Cell Biol* 23:5908–5918
8. Cheung IY, Vickers A, Cheung NKV (2006) Sialyltransferase STX (ST8SialII): a novel molecular marker of metastatic neuroblastoma. *Int J Cancer* 119:152–156
9. Jann K, Jann B (1987) Polysaccharide antigens of *Escherichia coli*. *Rev Infect Dis* 9 (Suppl 5):S517–S526
10. Wyle FA, Artenstein MS, Brandt BL, Tramont EC, Kasper DL, Altieri PL, Berman SL, Lowenthal JP (1972) Immunologic response of man to group B meningococcal polysaccharide vaccines. *J Infect Dis* 126:514–521
11. Sundar MM, Nagananda GS, Das A, Bhattacharya S, Suryan S (2009) Isolation of host-specific bacteriophages from sewage against human pathogens. *Asian J Biotechnol* 1:163–170
12. Stummeyer K, Dickmanns A, Mühlenhoff M, Gerardy-Schahn R, Ficner R (2005) Crystal structure of the polysialic acid-degrading endosialidase of bacteriophage K1F. *Nat Struct Mol Biol* 12:90–96
13. Leiman PG, Battisti AJ, Bowman VD, Stummeyer K, Mühlenhoff M, Gerardy-Schahn R, Scholl D, Molineux IJ (2007) The structures of bacteriophages K1E and K1-5 explain processive degradation of polysaccharide capsules and evolution of new host specificities. *J Mol Biol* 371:836–849
14. Mühlenhoff M, Stummeyer K, Grove M, Sauerborn M, Gerardy-Schahn R (2003) Proteolytic processing and oligomerization of bacteriophage-derived endosialidasas. *J Biol Chem* 278:12634–12644
15. Pelkonen S, Aalto J, Finne J (1992) Differential activities of bacteriophage depolymerase on bacterial polysaccharide: binding is essential but degradation is inhibitory in phage infection of K1-defective *Escherichia coli*. *J Bacteriol* 174:7757–7761
16. Twort FW (1915) An investigation on the nature of ultramicroscopic viruses. *Lancet* 186:1241–1243
17. d’Herelle F (1917) Sur un microbe invisible antagoniste des bacilles dysentériques. *Comptes Rendus de l’ Académie des Sciences (Paris)* 165:373–375
18. Bergh O, Børshheim KY, Bratbak G, Haldal M (1989) High abundance of viruses found in aquatic environments. *Nature* 340:467–468
19. Chibani-Chennoufi S, Canchaya C, Bruttin A, Brüssow H (2004) Comparative genomics of the T4-like *Escherichia coli* phage JS98: implications for the evolution of T4 phages. *J Bacteriol* 186:8276–8286
20. Brüssow H, Canchaya C, Hardt WD (2004) Phages and the evolution of bacterial pathogens: from genomic rearrangements to lysogenic conversion. *Microbiol Mol Biol Rev* 68:560–602
21. Ackermann HW (2007) 5500 Phages examined in the electron microscope. *Arch Virol* 152:227–243
22. Ackermann HW (2006) Classification of bacteriophages. In: Calendar R, Abedon ST (eds) *The bacteriophages*, 2nd edn. Oxford University Press, New York, pp 8–16
23. Bradley DE (1967) Ultrastructure of bacteriophage and bacteriocins. *Bacteriol Rev* 31:230–314
24. Ackermann HW (1998) Tailed bacteriophages: the order *Caudovirales*. *Adv Virus Res* 51:135–201
25. Leiman PG, Kanamaru S, Mesyanzhinov VV, Arisaka F, Rossmann MG (2003) Structure and morphogenesis of bacteriophage T4. *Cell Mol Life Sci* 60:2356–2370
26. Mesyanzhinov VV, Leiman PG, Kostyuchenko VA, Kurochkina LP, Miroshnikov KA, Sykilinda NN, Shneider MM (2004) Molecular architecture of bacteriophage T4. *Biochemistry (Mosc)* 69:1190–1202
27. Rossmann MG, Mesyanzhinov VV, Arisaka F, Leiman PG (2004) The bacteriophage T4 DNA injection machine. *Curr Opin Struct Biol* 14:171–180

28. Sayers JR (2006) Bacteriophage T5. In: Calendar R, Abedon ST (eds) *The bacteriophages*, 2nd edn. Oxford University Press, New York, pp 268–276
29. Molineux IJ (2006) The T7 group. In: Calendar R, Abedon ST (eds) *The bacteriophages*, 2nd edn. Oxford University Press, New York, pp 277–301
30. Steinbacher S, Seckler R, Miller S, Steipe B, Huber R, Reinemer P (1994) Crystal structure of P22 tailspike protein: interdigitated subunits in a thermostable trimer. *Science* 265:383–386
31. van Raaij MJ, Schoehn G, Burda MR, Miller S (2001) Crystal structure of a heat and protease-stable part of the bacteriophage T4 short tail fibre. *J Mol Biol* 314:1137–1146
32. Kanamaru S, Leiman PG, Kostyuchenko VA, Chipman PR, Mesyanzhinov VV, Arisaka F, Rossmann MG (2002) Structure of the cell-puncturing device of bacteriophage T4. *Nature* 415:553–557
33. Freiberg A, Morona R, den Bosch LV, Jung C, Behlke J, Carlin N, Seckler R, Baxa U (2003) The tailspike protein of Shigella phage Sf6. A structural homolog of salmonella phage P22 tailspike protein without sequence similarity in the beta-helix domain. *J Biol Chem* 278:1542–1548
34. Weigele PR, Scanlon E, King J (2003) Homotrimeric, beta-stranded viral adhesins and tail proteins. *J Bacteriol* 185:4022–4030
35. Stirm S, Freund-Mölbert E (1971) *Escherichia coli* capsule bacteriophages. II. Morphology. *J Virol* 8:330–342
36. Lindberg AA (1977) Bacterial surface carbohydrates and bacteriophage adsorption. In: Sutherland IW (ed) *Surface carbohydrates of the procaryotic cell*. Academic, New York, pp 289–356
37. Sutherland IW (1977) Enzymes acting on bacterial surface carbohydrates. In: Sutherland IW (ed) *Surface carbohydrates of the procaryotic cell*. Academic, New York, pp 209–245
38. Scholl D, Adhya S, Merrill C (2005) *Escherichia coli* K1's capsule is a barrier to bacteriophage T7. *Appl Environ Microbiol* 71:4872–4874
39. Cross AS (1990) The biologic significance of bacterial encapsulation. *Curr Top Microbiol Immunol* 150:87–95
40. Moxon ER, Kroll JS (1990) The role of bacterial polysaccharide capsules as virulence factors. *Curr Top Microbiol Immunol* 150:65–85
41. Taylor CM, Roberts IS (2005) Capsular polysaccharides and their role in virulence. *Contrib Microbiol* 12:55–66
42. Johnson JR (1991) Virulence factors in *Escherichia coli* urinary tract infection. *Clin Microbiol Rev* 4:80–128
43. Goller CC, Seed PC (2010) Revisiting the *Escherichia coli* polysaccharide capsule as a virulence factor during urinary tract infection: contribution to intracellular biofilm development. *Virulence* 1:333–337
44. Robbins JB, McCracken JGH, Gotschlich EC, Orskov F, Orskov I, Hanson LA (1974) *Escherichia coli* K1 capsular polysaccharide associated with neonatal meningitis. *N Engl J Med* 290:1216–1220
45. Sarff LD, McCracken GH, Schiffer MS, Glode MP, Robbins JB, Orskov I, Orskov F (1975) Epidemiology of *Escherichia coli* K1 in healthy and diseased newborns. *Lancet* 1:1099–1104
46. Varki NM, Varki A (2007) Diversity in cell surface sialic acid presentations: implications for biology and disease. *Lab Invest* 87:851–857
47. Barry GT, Goebel WF (1957) Colominic acid, a substance of bacterial origin related to sialic acid. *Nature* 179:206
48. McGuire EJ, Binkley SB (1964) The structure and chemistry of colominic acid. *Biochemistry* 3:247–251
49. Bhattacharjee AK, Jennings HJ, Kenny CP, Martin A, Smith IC (1975) Structural determination of the sialic acid polysaccharide antigens of *Neisseria meningitidis* serogroups B and C with carbon 13 nuclear magnetic resonance. *J Biol Chem* 250:1926–1932

50. Devi SJN, Schneerson R, Egan W, Vann WF, Robbins JB, Shiloach J (1991) Identity between polysaccharide antigens of *Moraxella nonliquefaciens*, group B *Neisseria meningitidis*, and *Escherichia coli* K1 (non-O acetylated). *Infect Immun* 59:732–736
51. Adlam C, Knights JM, Mugridge A, Williams JM, Lindon JC (1987) Production of colominic acid by *Pasteurella haemolytica* serotype A2 organism. *FEMS Microbiol Lett* 42:23–26
52. Furowicz AJ, Ørskov F (1972) Two new *Escherichia coli* O antigens, O150 and O157, and one new K antigen, K92, in strains isolated from veterinary diseases. *Acta Pathol Microbiol Scand B Microbiol Immunol* 80:441–444
53. Zhang Y, Inoue Y, Inoue S, Lee YC (1997) Separation of oligo/polymers of 5-N-acetylneuraminic acid, 5-N-glycolylneuraminic acid, and 2-keto-3-deoxy-D-glycero-D-galacto-nononic acid by high-performance anion-exchange chromatography with pulsed amperometric detector. *Anal Biochem* 250:245–251
54. Egan W, Liu TY, Dorow D, Cohen JS, Robbins JD, Gotschlich EC, Robbins JB (1977) Structural studies on the sialic acid polysaccharide antigen of *Escherichia coli* strain Bos-12. *Biochemistry* 16:3687–3692
55. Steenbergen SM, Lee YC, Vann WF, Vionnet J, Wright LF, Vimr ER (2006) Separate pathways for O acetylation of polymeric and monomeric sialic acids and identification of sialyl O-acetyl esterase in *Escherichia coli* K1. *J Bacteriol* 188:6195–6206
56. Bergfeld AK, Claus H, Vogel U, Mühlenhoff M (2007) Biochemical characterization of the polysialic acid-specific O-acetyltransferase NeuO of *Escherichia coli* K1. *J Biol Chem* 282:22217–22227
57. Bergfeld AK, Claus H, Lorenzen NK, Spielmann F, Vogel U, Mühlenhoff M (2009) The polysialic acid-specific O-acetyltransferase OatC from *Neisseria meningitidis* serogroup C evolved apart from other bacterial sialate O-acetyltransferases. *J Biol Chem* 284:6–16
58. Finne J (1982) Occurrence of unique polysialosyl carbohydrate units in glycoproteins of developing brain. *J Biol Chem* 257:11966–11970
59. Zuber C, Lackie PM, Catterall WA, Roth J (1992) Polysialic acid is associated with sodium channels and the neural cell adhesion molecule N-CAM in adult rat brain. *J Biol Chem* 267:9965–9971
60. Yabe U, Sato C, Matsuda T, Kitajima K (2003) Polysialic acid in human milk. CD36 is a new member of mammalian polysialic acid-containing glycoprotein. *J Biol Chem* 278:13875–13880
61. Moebius JM, Widera D, Schmitz J, Kaltschmidt C, Piechaczek C (2007) Impact of polysialylated CD56 on natural killer cell cytotoxicity. *BMC Immunol* 8:13
62. Curreli S, Arany Z, Gerardy-Schahn R, Mann D, Stamatou NM (2007) Polysialylated neuropilin-2 is expressed on the surface of human dendritic cells and modulates dendritic cell-T lymphocyte interactions. *J Biol Chem* 282:30346–30356
63. Galuska SP, Rollenhagen M, Kaup M, Eggers K, Oltmann-Norden I, Schiff M, Hartmann M, Weinhold B, Hildebrandt H, Geyer R, Mühlenhoff M, Geyer H (2010) Synaptic cell adhesion molecule SynCAM 1 is a target for polysialylation in postnatal mouse brain. *Proc Natl Acad Sci USA* 107:10250–10255
64. Finne J, Leinonen M, Mäkelä PH (1983) Antigenic similarities between brain components and bacteria causing meningitis. Implications for vaccine development and pathogenesis. *Lancet* 322:355–357
65. Frosch M, Roberts I, Görgen I, Metzger S, Boulnois GJ, Bitter-Suermann D (1987) Serotyping and genotyping of encapsulated *Escherichia coli* K1 sepsis isolates with a monoclonal IgG anti K1 antibody and K1 gene probes. *Microb Pathog* 2:319–326
66. Prasadarao NV, Wass CA, Weiser JN, Stins MF, Huang SH, Kim KS (1996) Outer membrane protein A of *Escherichia coli* contributes to invasion of brain microvascular endothelial cells. *Infect Immun* 64:146–153
67. Jódar L, Feavers IM, Salisbury D, Granoff DM (2002) Development of vaccines against meningococcal disease. *Lancet* 359:1499–1508

68. Gross RJ, Cheasty T, Rowe B (1977) Isolation of bacteriophages specific for the K1 polysaccharide antigen of *Escherichia coli*. *J Clin Microbiol* 6:548–550
69. Jakobsson E, Jokilampi A, Aalto J, Ollikka P, Lehtonen JV, Hirvonen H, Finne J (2007) Identification of amino acid residues at the active site of endosialidase that dissociate the polysialic acid binding and cleaving activities in *Escherichia coli* K1 bacteriophages. *Biochem J* 405:465–472
70. Long GS, Bryant JM, Taylor PW, Luzio JP (1995) Complete nucleotide sequence of the gene encoding bacteriophage E endosialidase: implication for K1E endosialidase structure and function. *Biochem J* 309:543–550
71. Gerardy-Schahn R, Bethe A, Brennecke T, Mühlenhoff M, Eckhardt M, Ziesing S, Lottspeich F, Frosch M (1995) Molecular cloning and functional expression of bacteriophage PK1E-encoded endoneuraminidase Endo NE. *Mol Microbiol* 16:441–450
72. Kwiatkowski B, Boschek B, Thiele H, Stirn S (1982) Endo-N-acetylneuraminidase associated with bacteriophage particles. *J Virol* 43:697–704
73. Smith HW, Huggins MB (1982) Successful treatment of experimental *Escherichia coli* infections in mice using phage: its general superiority over antibiotics. *J Gen Microbiol* 128:307–318
74. Kwiatkowski B, Boschek B, Thiele H, Stirn S (1983) Substrate specificity of two bacteriophage-associated endo-N-acetylneuraminidases. *J Virol* 45:367–374
75. Vimr ER, McCoy RD, Vollger HF, Wilkison NC, Troy FA (1984) Use of prokaryotic-derived probes to identify poly(sialic acid) in neonatal neuronal membranes. *Proc Natl Acad Sci USA* 81:1971–1975
76. Petter JG, Vimr ER (1993) Complete nucleotide sequence of the bacteriophage K1F tail gene encoding endo-N-acylneuraminidase (endo-N) and comparison to an endo-N homolog in bacteriophage PK1E. *J Bacteriol* 175:4354–4363
77. Miyake K, Muraki T, Hattori K, Machida Y, Watanabe M, Kawase M, Yoshida Y, Iijima S (1997) Screening of bacteriophages producing endo-N-acetylneuraminidase. *J Ferment Bioeng* 84:90–93
78. Machida Y, Hattori K, Miyake K, Kawase Y, Kawase M, Iijima S (2000) Molecular cloning and characterization of a novel bacteriophage-associated sialidase. *J Biosci Bioeng* 90:62–68
79. Scholl D, Rogers S, Adhya S, Merrill CR (2001) Bacteriophage K1-5 encodes two different tail fiber proteins, allowing it to infect and replicate on both K1 and K5 strains of *Escherichia coli*. *J Virol* 75:2509–2515
80. Deszo EL, Steenbergen SM, Freedberg DI, Vimr ER (2005) *Escherichia coli* K1 polysialic acid O-acetyltransferase gene, neuO, and the mechanism of capsule form variation involving a mobile contingency locus. *Proc Natl Acad Sci USA* 102:5564–5569
81. Stummeyer K, Schwarzer D, Claus H, Vogel U, Gerardy-Schahn R, Mühlenhoff M (2006) Evolution of bacteriophages infecting encapsulated bacteria: lessons from *Escherichia coli* K1-specific phages. *Mol Microbiol* 60:1123–1135
82. Chen SL, Hung CS, Xu J, Reigstad CS, Magrini V, Sabo A, Blasiar D, Bieri T, Meyer RR, Ozersky P, Armstrong JR, Fulton RS, Latreille JP, Spieth J, Hooton TM, Mardis ER, Hultgren SJ, Gordon JI (2006) Identification of genes subject to positive selection in uropathogenic strains of *Escherichia coli*: a comparative genomics approach. *Proc Natl Acad Sci USA* 103:5977–5982
83. Johnson TJ, Kariyawasam S, Wannemuehler Y, Mangiamale P, Johnson SJ, Doetkott C, Skyberg JA, Lynne AM, Johnson JR, Nolan LK (2007) The genome sequence of avian pathogenic *Escherichia coli* strain O1:K1:H7 shares strong similarities with human extraintestinal pathogenic *E. coli* genomes. *J Bacteriol* 189:3228–3236
84. Touchon M, Hoede C, Tenaillon O, Barbe V, Baeriswyl S, Bidet P, Bingen E, Bonacorsi S, Bouchier C, Bouvet O, Calteau A, Chiapello H, Clermont O, Cruveiller S, Danchin A, Diard M, Dossat C, Karoui ME, Frapy E, Garry L, Ghigo JM, Gilles AM, Johnson J, Bouguéné C, Lescat M, Mangenot S, Martinez-Jéhanne V, Matic I, Nassif X, Oztas S, Petit MA, Pichon C, Rouy Z, Ruf CS, Schneider D, Tournet J, Vacherie B, Vallenet D, Médigue C, Rocha EPC,

- Denamur E (2009) Organised genome dynamics in the *Escherichia coli* species results in highly diverse adaptive paths. *PLoS Genet* 5:e1000344
85. Moriel DG, Bertoldi I, Spagnuolo A, Marchi S, Rosini R, Nesta B, Pastorello I, Corea VAM, Torricelli G, Cartocci E, Savino S, Scarselli M, Dobrindt U, Hacker J, Tettelin H, Tallon LJ, Sullivan S, Wieler LH, Ewers C, Pickard D, Dougan G, Fontana MR, Rappuoli R, Pizza M, Serino L (2010) Identification of protective and broadly conserved vaccine antigens from the genome of extraintestinal pathogenic *Escherichia coli*. *Proc Natl Acad Sci USA* 107:9072–9077
 86. Bull JJ, Vimr ER, Molineux IJ (2010) A tale of tails: sialidase is key to success in a model of phage therapy against K1-capsulated *Escherichia coli*. *Virology* 398:79–86
 87. Krause DO, Little AC, Dowd SE, Bernstein CN (2011) Complete genome sequence of adherent invasive *Escherichia coli* UM146 isolated from ileal Crohn's disease biopsy tissue. *J Bacteriol* 193:583
 88. Steinbacher S, Baxa U, Miller S, Weintraub A, Seckler R, Huber R (1996) Crystal structure of phage P22 tailspike protein complexed with *Salmonella* sp. O-antigen receptors. *Proc Natl Acad Sci USA* 93:10584–10588
 89. Müller JJ, Barbirz S, Heinle K, Freiberg A, Seckler R, Heinemann U (2008) An intersubunit active site between supercoiled parallel beta helices in the trimeric tailspike endorhamnosidase of *Shigella flexneri* Phage Sf6. *Structure* 16:766–775
 90. Barbirz S, Müller JJ, Uetrecht C, Clark AJ, Heinemann U, Seckler R (2008) Crystal structure of *Escherichia coli* phage HK620 tailspike: podoviral tailspike endoglycosidase modules are evolutionarily related. *Mol Microbiol* 69:303–316
 91. Li YT, Nakagawa H, Ross SA, Hansson GC, Li SC (1990) A novel sialidase which releases 2,7-anhydro-alpha-N-acetylneuraminic acid from sialoglycoconjugates. *J Biol Chem* 265:21629–21633
 92. Hirst GK (1942) Adsorption of influenza hemagglutinins and virus by red blood cells. *J Exp Med* 76:195–209
 93. Stirm S (1968) *Escherichia coli* K bacteriophages. I. Isolation and introductory characterization of five *Escherichia coli* K bacteriophages. *J Virol* 2:1107–1114
 94. Stirm S, Bessler W, Fehmel F, Freund-Mölbert E (1971) Bacteriophage particles with endoglycosidase activity. *J Virol* 8:343–346
 95. Cabezas JA (1991) Some questions and suggestions on the type references of the official nomenclature (IUB) for sialidase (s) and endosialidase. *Biochem J* 278(Pt 1):311–312
 96. Kitajima K, Inoue S, Inoue Y, Troy FA (1988) Use of a bacteriophage-derived endo-N-acetylneuraminidase and an equine antipolysialyl antibody to characterize the polysialyl residues in salmonid fish egg polysialoglycoproteins. Substrate and immunospecificity studies. *J Biol Chem* 263:18269–18276
 97. Uchida Y, Tsukada Y, Sugimori T (1979) Enzymatic properties of neuraminidases from *Arthrobacter ureafaciens*. *J Biochem* 86:1573–1585
 98. Cheng MC, Lin CH, Lin HJ, Yu YP, Wu SH (2004) Hydrolysis, lactonization, and identification of alpha(2 ->8)/alpha(2 ->9) alternatively linked tri-, tetra-, and polysialic acids. *Glycobiology* 14:147–155
 99. Kwiatkowski B, Stirm S (1987) Polysialic acid depolymerase. *Methods Enzymol* 138:786–792
 100. Orskov F, Orskov I, Sutton A, Schneerson R, Lin W, Egan W, Hoff GE, Robbins JB (1979) Form variation in *Escherichia coli* K1: determined by O-acetylation of the capsular polysaccharide. *J Exp Med* 149:669–685
 101. Tomlinson S, Taylor PW (1985) Neuraminidase associated with coliphage E that specifically depolymerizes the *Escherichia coli* K1 capsular polysaccharide. *J Virol* 55:374–378
 102. Finne J, Mäkelä PH (1985) Cleavage of the polysialosyl units of brain glycoproteins by a bacteriophage endosialidase. Involvement of a long oligosaccharide segment in molecular interactions of polysialic acid. *J Biol Chem* 260:1265–1270

103. Pelkonen S, Pelkonen J, Finne J (1989) Common cleavage pattern of polysialic acid by bacteriophage endosialidases of different properties and origins. *J Virol* 63:4409–4416
104. Hallenbeck PC, Vimr ER, Yu F, Bassler B, Troy FA (1987) Purification and properties of a bacteriophage-induced endo-N-acetylneuraminidase specific for poly-alpha-2,8-sialosyl carbohydrate units. *J Biol Chem* 262:3553–3561
105. Schwarzer D, Stummeyer K, Haselhorst T, Freiburger F, Rode B, Grove M, Scheper T, von Itzstein M, Mühlenhoff M, Gerardy-Schahn R (2009) Proteolytic release of the intramolecular chaperone domain confers processivity to endosialidase F. *J Biol Chem* 284:9465–9474
106. Schulz EC, Schwarzer D, Frank M, Stummeyer K, Mühlenhoff M, Dickmanns A, Gerardy-Schahn R, Ficner R (2010) Structural basis for the recognition and cleavage of polysialic acid by the bacteriophage K1F tailspike protein EndoNF. *J Mol Biol* 397:341–351
107. Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60:2126–2132
108. Vimr ER, Aaronson W, Silver RP (1989) Genetic analysis of chromosomal mutations in the polysialic acid gene cluster of *Escherichia coli* K1. *J Bacteriol* 171:1106–1117
109. Schulz EC, Dickmanns A, Urlaub H, Schmitt A, Mühlenhoff M, Stummeyer K, Schwarzer D, Gerardy-Schahn R, Ficner R (2010) Crystal structure of an intramolecular chaperone mediating triple-beta-helix folding. *Nat Struct Mol Biol* 17:210–215
110. Schwarzer D (2008) Characterisation of bacteriophage-derived tailspike and tail fibre proteins. Ph.D. thesis. Leibniz University Hannover. Hannover. (<http://edok01.tib.uni-hannover.de/edoks/e01dh08/564826944.pdf>)
111. Pajunen M, Kiljunen S, Skurnik M (2000) Bacteriophage phiYeO3-12, specific for *Yersinia enterocolitica* serotype O:3, is related to coliphages T3 and T7. *J Bacteriol* 182:5114–5120
112. Schwarzer D, Stummeyer K, Gerardy-Schahn R, Mühlenhoff M (2007) Characterization of a novel intramolecular chaperone domain conserved in endosialidases and other bacteriophage tail spike and fiber proteins. *J Biol Chem* 282:2821–2831
113. Schulz EC, Neumann P, Gerardy-Schahn R, Sheldrick GM, Ficner R (2010) Structure analysis of endosialidase NF at 0.98 Å resolution. *Acta Crystallogr D Biol Crystallogr* 66:176–180
114. Crennell SJ, Garman EF, Laver WG, Vimr ER, Taylor GL (1993) Crystal structure of a bacterial sialidase (from *Salmonella typhimurium* LT2) shows the same fold as an influenza virus neuraminidase. *Proc Natl Acad Sci USA* 90:9852–9856
115. Newstead SL, Potter JA, Wilson JC, Xu G, Chien CH, Watts AG, Withers SG, Taylor GL (2008) The structure of *Clostridium perfringens* NanI sialidase and its catalytic intermediates. *J Biol Chem* 283:9080–9088
116. Gaskell A, Crennell S, Taylor G (1995) The three domains of a bacterial sialidase: a beta-propeller, an immunoglobulin module and a galactose-binding jelly-roll. *Structure* 3:1197–1205
117. Luo Y, Li SC, Chou MY, Li YT, Luo M (1998) The crystal structure of an intramolecular trans-sialidase with a NeuAc alpha2->3 Gal specificity. *Structure* 6:521–530
118. Moustafa I, Connaris H, Taylor M, Zaitsev V, Wilson JC, Kiefel MJ, von Itzstein M, Taylor G (2004) Sialic acid recognition by *Vibrio cholerae* neuraminidase. *J Biol Chem* 279:40819–40826
119. Buschiazzo A, Tavares GA, Campetella O, Spinelli S, Cremona ML, Paris G, Amaya MF, Frasch AC, Alzari PM (2000) Structural basis of sialyltransferase activity in trypanosomal sialidases. *EMBO J* 19:16–24
120. Buschiazzo A, Amaya MF, Cremona ML, Frasch AC, Alzari PM (2002) The crystal structure and mode of action of trans-sialidase, a key enzyme in *Trypanosoma cruzi* pathogenesis. *Mol Cell* 10:757–768
121. Thobhani S, Ember B, Siriwardena A, Boons GJ (2003) Multivalency and the mode of action of bacterial sialidases. *J Am Chem Soc* 125:7154–7155
122. Olia AS, Casjens S, Cingolani G (2007) Structure of phage P22 cell envelope-penetrating needle. *Nat Struct Mol Biol* 14:1221–1226

123. Manning M, Colón W (2004) Structural basis of protein kinetic stability: resistance to sodium dodecyl sulfate suggests a central role for rigidity and a bias toward beta-sheet structure. *Biochemistry* 43:11248–11254
124. Weigele PR, Haase-Pettingell C, Campbell PG, Gossard DC, King J (2005) Stalled folding mutants in the triple beta-helix domain of the phage P22 tailspike adhesion. *J Mol Biol* 354:1103–1117
125. Botstein D (1980) A theory of modular evolution for bacteriophages. *Ann N Y Acad Sci* 354:484–490
126. Casjens SR (2005) Comparative genomics and evolution of the tailed-bacteriophages. *Curr Opin Microbiol* 8:451–458
127. Morley TJ, Willis LM, Whitfield C, Wakarchuk WW, Withers SG (2009) A new sialidase mechanism: bacteriophage K1F endo-sialidase is an inverting glycosidase. *J Biol Chem* 284:17404–17410
128. Friebolin H, Brossmer R, Keilich G, Ziegler D, Supp M (1980) 1H-NMR-Spectroscopic evidence for the release of N-acetyl-alpha-D-neuraminic acid as the first product of neuraminidase action (author's transl). *Hoppe Seylers Z Physiol Chem* 361:697–702
129. Todeschini AR, Mendonça-Previato L, Previato JO, Varki A, van Halbeek H (2000) Trans-sialidase from *Trypanosoma cruzi* catalyzes sialoside hydrolysis with retention of configuration. *Glycobiology* 10:213–221
130. Manzi AE, Higa HH, Diaz S, Varki A (1994) Intramolecular self-cleavage of polysialic acid. *J Biol Chem* 269:23617–23624
131. Yamasaki R, Bacon B (1991) Three-dimensional structural analysis of the group B polysaccharide of *Neisseria meningitidis* 6275 by two-dimensional NMR: the polysaccharide is suggested to exist in helical conformations in solution. *Biochemistry* 30:851–857
132. Brisson JR, Baumann H, Imberty A, Pérez S, Jennings HJ (1992) Helical epitope of the group B meningococcal alpha(2–8)-linked sialic acid polysaccharide. *Biochemistry* 31:4996–5004
133. Azurmendi HF, Vionnet J, Wrightson L, Trinh LB, Shiloach J, Freedberg DI (2007) Extracellular structure of polysialic acid explored by on cell solution NMR. *Proc Natl Acad Sci USA* 104:11557–11561
134. Michon F, Brisson JR, Jennings HJ (1987) Conformational differences between linear alpha (2–8)-linked homosialooligosaccharides and the epitope of the group B meningococcal polysaccharide. *Biochemistry* 26:8399–8405
135. Haselhorst T, Stummeyer K, Mühlhoff M, Schaper W, Gerardy-Schahn R, von Itzstein M (2006) Endosialidase NF appears to bind polySia DP5 in a helical conformation. *Chembiochem* 7:1875–1877
136. Schulz EC, Ficner R (2011) Knitting and snipping: chaperones in β -helix folding. *Curr Opin Struct Biol* 21:232–239
137. Xiang Y, Leiman PG, Li L, Grimes S, Anderson DL, Rossmann MG (2009) Crystallographic insights into the autocatalytic assembly mechanism of a bacteriophage tail spike. *Mol Cell* 34:375–386
138. Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227:680–685
139. Marvik OJ, Jacobsen E, Dokland T, Lindqvist BH (1994) Bacteriophage P2 and P4 morphogenesis: assembly precedes proteolytic processing of the capsid proteins. *Virology* 205:51–65
140. Conway JF, Duda RL, Cheng N, Hendrix RW, Steven AC (1995) Proteolytic and conformational control of virus capsid maturation: the bacteriophage HK97 system. *J Mol Biol* 253:86–99
141. Wang S, Chandramouli P, Butcher S, Dokland T (2003) Cleavage leads to expansion of bacteriophage P4 procapsids in vitro. *Virology* 314:1–8
142. Miller ES, Kutter E, Mosig G, Arisaka F, Kunisawa T, Rüger W (2003) Bacteriophage T4 genome. *Microbiol Mol Biol Rev* 67:86–156, table of contents
143. Cunningham EL, Jaswal SS, Sohl JL, Agard DA (1999) Kinetic stability as a mechanism for protease longevity. *Proc Natl Acad Sci USA* 96:11008–11014

144. Bayer ME, Thurow H, Bayer MH (1979) Penetration of the polysaccharide capsule of *Escherichia coli* (B1161/42) by bacteriophage K29. *Virology* 94:95–118
145. Leiman PG, Molineux IJ (2008) Evolution of a new enzyme activity from the same motif fold. *Mol Microbiol* 69:287–290
146. Koivula A, Kinnari T, Harjunpää V, Ruohonen L, Teleman A, Drakenberg T, Rouvinen J, Jones TA, Teeri TT (1998) Tryptophan 272: an essential determinant of crystalline cellulose degradation by *Trichoderma reesei* cellobiohydrolase Cel6A. *FEBS Lett* 429:341–346
147. Zhang S, Irwin DC, Wilson DB (2000) Site-directed mutation of noncatalytic residues of *Thermobifida fusca* exocellulase Ce16B. *Eur J Biochem* 267:3101–3115
148. Watanabe T, Ariga Y, Sato U, Toratani T, Hashimoto M, Nikaidou N, Kezuka Y, Nonaka T, Sugiyama J (2003) Aromatic residues within the substrate-binding cleft of *Bacillus circulans* chitinase A1 are essential for hydrolysis of crystalline chitin. *Biochem J* 376:237–244
149. Katouno F, Taguchi M, Sakurai K, Uchiyama T, Nikaidou N, Nonaka T, Sugiyama J, Watanabe T (2004) Importance of exposed aromatic residues in chitinase B from *Serratia marcescens* 2170 for crystalline chitin hydrolysis. *J Biochem* 136:163–168
150. Eijsink VGH, Vaaje-Kolstad G, Vårum KM, Horn SJ (2008) Towards new enzymes for biofuels: lessons from chitinase research. *Trends Biotechnol* 26:228–235
151. Horn SJ, Sikorski P, Cederkvist JB, Vaaje-Kolstad G, Sørli M, Synstad B, Vriend G, Vårum KM, Eijsink VGH (2006) Costs and benefits of processivity in enzymatic degradation of recalcitrant polysaccharides. *Proc Natl Acad Sci USA* 103:18089–18094
152. Nimmich W, Zingler G (1984) Biochemical characteristics, phage patterns, and O1 factor analysis of *Escherichia coli* O1:K1:H7:F11 and O1:K1:H:F9 strains isolated from patients with urinary tract infections. *Med Microbiol Immunol* 173:75–85
153. Rutishauser U, Watanabe M, Silver J, Troy FA, Vimr ER (1985) Specific alteration of NCAM-mediated cell adhesion by an endoneuraminidase. *J Cell Biol* 101:1842–1849
154. Leggate DR, Bryant JM, Redpath MB, Head D, Taylor PW, Luzio JP (2002) Expression, mutagenesis and kinetic analysis of recombinant K1E endosialidase to define the site of proteolytic processing and requirements for catalysis. *Mol Microbiol* 44:749–760
155. Pelkonen S, Häyrynen J, Finne J (1988) Polyacrylamide gel electrophoresis of the capsular polysaccharides of *Escherichia coli* K1 and other bacteria. *J Bacteriol* 170:2646–2653
156. Sato C, Inoue S, Matsuda T, Kitajima K (1998) Development of a highly sensitive chemical method for detecting alpha2->8-linked oligo/polysialic acid residues in glycoproteins blotted on the membrane. *Anal Biochem* 261:191–197
157. Liedtke S, Geyer H, Wuhrer M, Geyer R, Frank G, Gerardy-Schahn R, Zähringer U, Schachner M (2001) Characterization of N-glycans from mouse brain neural cell adhesion molecule. *Glycobiology* 11:373–384
158. Calandreau L, Márquez C, Bisaz R, Fantin M, Sandi C (2010) Differential impact of polysialyltransferase ST8SiaII and ST8SiaIV knockout on social interaction and aggression. *Genes Brain Behav* 9:958–967
159. Galuska SP, Geyer R, Gerardy-Schahn R, Mühlenhoff M, Geyer H (2008) Enzyme-dependent variations in the polysialylation of the neural cell adhesion molecule (NCAM) in vivo. *J Biol Chem* 283:17–28
160. Frosch M, Görden I, Boulnois GJ, Timmis KN, Bitter-Suermann D (1985) NZB mouse system for production of monoclonal antibodies to weak bacterial antigens: isolation of an IgG antibody to the polysaccharide capsules of *Escherichia coli* K1 and group B meningococci. *Proc Natl Acad Sci USA* 82:1194–1198
161. Brusés JL, Rutishauser U (2000) Polysialic acid in neural cell development: roles, regulation and mechanism. In: Fukuda M, Hindsgaul O (eds) *Molecular and cellular glycobiology*. Oxford University Press, New York, pp 116–132
162. Daston MM, Bastmeyer M, Rutishauser U, O’Leary DD (1996) Spatially restricted increase in polysialic acid enhances corticospinal axon branching related to target recognition and innervation. *J Neurosci* 16:5488–5497

163. Seki T, Rutishauser U (1998) Removal of polysialic acid-neural cell adhesion molecule induces aberrant mossy fiber innervation and ectopic synaptogenesis in the hippocampus. *J Neurosci* 18:3757–3766
164. Becker CG, Artola A, Gerardy-Schahn R, Becker T, Welzl H, Schachner M (1996) The polysialic acid modification of the neural cell adhesion molecule is involved in spatial learning and hippocampal long-term potentiation. *J Neurosci Res* 45:143–152
165. Dityatev A, Dityateva G, Schachner M (2000) Synaptic strength as a function of post- versus presynaptic expression of the neural cell adhesion molecule NCAM. *Neuron* 26:207–217
166. Aalto J, Pelkonen S, Kalimo H, Finne J (2001) Mutant bacteriophage with non-catalytic endosialidase binds to both bacterial and eukaryotic polysialic acid and can be used as probe for its detection. *Glycoconj J* 18:751–758
167. Durbec P, Cremer H (2001) Revisiting the function of PSA-NCAM in the nervous system. *Mol Neurobiol* 24:53–64
168. Jokilampi A, Ollikka P, Korja M, Jakobsson E, Loimaranta V, Haataja S, Hirvonen H, Finne J (2004) Construction of antibody mimics from a noncatalytic enzyme-detection of polysialic acid. *J Immunol Methods* 295:149–160
169. Dityatev A, Dityateva G, Sytnyk V, Delling M, Toni N, Nikonenko I, Muller D, Schachner M (2004) Polysialylated neural cell adhesion molecule promotes remodeling and formation of hippocampal synapses. *J Neurosci* 24:9372–9382
170. Weinhold B, Seidenfaden R, Röckle I, Mühlenhoff M, Schertzing F, Conzelmann S, Marth JD, Gerardy-Schahn R, Hildebrandt H (2005) Genetic ablation of polysialic acid causes severe neurodevelopmental defects rescued by deletion of the neural cell adhesion molecule. *J Biol Chem* 280:42971–42977
171. Burgess A, Weng YQ, Ypsilanti AR, Cui X, Caines G, Aubert I (2007) Polysialic acid limits septal neurite outgrowth on laminin. *Brain Res* 1144:52–58
172. Freiburger F, Claus H, Günzel A, Oltmann-Norden I, Vionnet J, Mühlenhoff M, Vogel U, Vann WF, Gerardy-Schahn R, Stummeyer K (2007) Bio-chemical characterization of a *Neisseria meningitidis* polysialyltransferase reveals novel functional motifs in bacterial sialyltransferases. *Mol Microbiol* 65:1258–1275
173. Oltmann-Norden I, Galuska SP, Hildebrandt H, Geyer R, Gerardy-Schahn R, Geyer H, Mühlenhoff M (2008) Impact of the polysialyltransferases ST8SiaII and ST8SiaIV on polysialic acid synthesis during postnatal mouse brain development. *J Biol Chem* 283:1463–1471
174. Hildebrandt H, Mühlenhoff M, Gerardy-Schahn R (2010) Polysialylation of NCAM. *Adv Exp Med Biol* 663:95–109
175. Kiss JZ, Wang C, Olive S, Rougon G, Lang J, Baetens D, Harry D, Pralong WF (1994) Activity-dependent mobilization of the adhesion molecule polysialic NCAM to the cell surface of neurons and endocrine cells. *EMBO J* 13:5284–5292
176. Garcia-Segura LM, Cañas B, Parducz A, Rougon G, Theodosis D, Naftolin F, Torres-Aleman I (1995) Estradiol promotion of changes in the morphology of astroglia growing in culture depends on the expression of polysialic acid of neural membranes. *Glia* 13:209–216
177. Wang C, Rougon G, Kiss JZ (1994) Requirement of polysialic acid for the migration of the O-2A glial progenitor cell from neurohypophyseal explants. *J Neurosci* 14:4446–4457
178. Meyer-Franke A, Shen S, Barres BA (1999) Astrocytes induce oligodendrocyte processes to align with and adhere to axons. *Mol Cell Neurosci* 14:385–397
179. Berski S, van Bergeijk J, Schwarzer D, Stark Y, Kasper C, Scheper T, Grothe C, Gerardy-Schahn R, Kirschning A, Dräger G (2008) Synthesis and biological evaluation of a polysialic acid-based hydrogel as enzymatically degradable scaffold material for tissue engineering. *Biomacromolecules* 9:2353–2359
180. Haile Y, Berski S, Dräger G, Nobre A, Stummeyer K, Gerardy-Schahn R, Grothe C (2008) The effect of modified polysialic acid based hydrogels on the adhesion and viability of primary neurons and glial cells. *Biomaterials* 29:1880–1891

181. Haastert-Talini K, Schaper-Rinkel J, Schmitte R, Bastian R, Mühlhoff M, Schwarzer D, Draeger G, Su Y, Scheper T, Gerardy-Schahn R, Grothe C (2010) In vivo evaluation of polysialic acid as part of tissue-engineered nerve transplants. *Tissue Eng Part A* 16:3085–3098
182. Steinhaus S, Stark Y, Bruns S, Haile Y, Scheper T, Grothe C, Behrens P (2010) Polysialic acid immobilized on silanized glass surfaces: a test case for its use as a biomaterial for nerve regeneration. *J Mater Sci Mater Med* 21:1371–1378
183. Franz CK, Rutishauser U, Rafuse VF (2008) Intrinsic neuronal properties control selective targeting of regenerating motoneurons. *Brain* 131:1492–1505
184. James WM, Agnew WS (1987) Multiple oligosaccharide chains in the voltage-sensitive Na channel from *Electrophorus electricus*: evidence for alpha-2,8-linked polysialic acid. *Biochem Biophys Res Commun* 148:817–826
185. Lanier LL, Testi R, Bindl J, Phillips JH (1989) Identity of Leu-19 (CD56) leukocyte differentiation antigen and neural cell adhesion molecule. *J Exp Med* 169:2233–2238
186. Avril T, North SJ, Haslam SM, Willison HJ, Crocker PR (2006) Probing the cis interactions of the inhibitory receptor Siglec-7 with alpha2,8-disialylated ligands on natural killer cells and other leukocytes using glycan-specific antibodies and by analysis of alpha2,8-sialyltransferase gene expression. *J Leukoc Biol* 80:787–796
187. Rey-Gallardo A, Delgado-Martín C, Gerardy-Schahn R, Rodríguez-Fernández JL, Vega MA (2011) Polysialic acid is required for neuropilin2a/b-mediated control of CCL21-driven chemotaxis of mature dendritic cells and for their migration in vivo. *Glycobiology* 21:655–662
188. Rey-Gallardo A, Escribano C, Delgado-Martín C, Rodríguez-Fernández JL, Gerardy-Schahn R, Rutishauser U, Corbi AL, Vega MA (2010) Polysialylated neuropilin-2 enhances human dendritic cell migration through the basic C-terminal region of CCL21. *Glycobiology* 20:1139–1146
189. Hildebrandt H, Becker C, Müräu M, Gerardy-Schahn R, Rahmann H (1998) Heterogeneous expression of the polysialyltransferases ST8Sia II and ST8Sia IV during postnatal rat brain development. *J Neurochem* 71:2339–2348
190. Lantuejoul S, Moro D, Michalides RJ, Brambilla C, Brambilla E (1998) Neural cell adhesion molecules (NCAM) and NCAM-PSA expression in neuroendocrine lung tumors. *Am J Surg Pathol* 22:1267–1276
191. Roth J, Zuber C, Wagner P, Taatjes DJ, Weisgerber C, Heitz PU, Goridis C, Bitter-Suermann D (1988) Reexpression of poly(sialic acid) units of the neural cell adhesion molecule in Wilms tumor. *Proc Natl Acad Sci USA* 85:2999–3003
192. Roesler J, Srivatsan E, Moetamed F, Peters J, Livingston EH (1997) Tumor suppressor activity of neural cell adhesion molecule in colon carcinoma. *Am J Surg* 174:251–257
193. Tezel E, Kawase Y, Takeda S, Oshima K, Nakao A (2001) Expression of neural cell adhesion molecule in pancreatic cancer. *Pancreas* 22:122–125
194. Daniel L, Durbec P, Gautherot E, Rouvier E, Rougon G, Figarella-Branger D (2001) A nude mice model of human rhabdomyosarcoma lung metastases for evaluating the role of polysialic acids in the metastatic process. *Oncogene* 20:997–1004
195. Mushtaq N, Redpath MB, Luzio JP, Taylor PW (2005) Treatment of experimental *Escherichia coli* infection with recombinant bacteriophage-derived capsule depolymerase. *J Antimicrob Chemother* 56:160–165
196. Mushtaq N, Redpath MB, Jp L, Taylor PW (2004) Prevention and cure of systemic *Escherichia coli* K1 infection by modification of the bacterial phenotype. *Antimicrob Agents Chemother* 48:1503–1508
197. Pelkonen S (1990) Capsular sialyl chains of *Escherichia coli* K1 mutants resistant to K1 phage. *Curr Microbiol* 21:23–28
198. Colman PM, Hoyne PA, Lawrence MC (1993) Sequence and structure alignment of paramyxovirus hemagglutinin-neuraminidase with influenza virus neuraminidase. *J Virol* 67:2972–2980

199. Hoiczky E, Roggenkamp A, Reichenbecher M, Lupas A, Heesemann J (2000) Structure and sequence analysis of *Yersinia* YadA and *Moraxella* UspAs reveal a novel class of adhesins. *EMBO J* 19:5989–5999
200. Pellequer JL, Van Regenmortel MH (1993) Affinity of monoclonal antibodies to large multivalent antigens: influence of steric hindrance on antibody affinity constants calculated from Scatchard plots. *Mol Immunol* 30:955–958
201. Häyrynen J, Haseley S, Talaga P, Mühlhoff M, Finne J, Vliegthart JFG (2002) High affinity binding of long-chain polysialic acid to antibody, and modulation by divalent cations and polyamines. *Mol Immunol* 39:399–411
202. Romano P, Manniello A, Aresu O, Armento M, Cesaro M, Parodi B (2009) Cell Line Data Base: structure and recent improvements towards molecular authentication of human cell lines. *Nucleic Acids Res* 37:D925–D932
203. Korja M, Jokilampi A, Salmi TT, Kalimo H, Pelliniemi TT, Isola J, Rantala I, Haapasalo H, Finne J (2009) Absence of polysialylated NCAM is an unfavorable prognostic phenotype for advanced stage neuroblastoma. *BMC Cancer* 9:57
204. Zelmer A, Bowen M, Jokilampi A, Finne J, Luzio JP, Taylor PW (2008) Differential expression of the polysialyl capsule during blood-to-brain transit of neuropathogenic *Escherichia coli* K1. *Microbiology* 154:2522–2532

Advanced Technologies in Sialic Acid and Sialoglycoconjugate Analysis

Ken Kitajima, Nissi Varki, and Chihiro Sato

Abstract Although the structural diversity of sialic acid (Sia) is rapidly expanding, understanding of its biological significance has lagged behind. Advanced technologies to detect and probe diverse structures of Sia are absolutely necessary not only to understand further biological significance but also to pursue medicinal and industrial applications. Here we describe analytical methods for detection of Sia that have recently been developed or improved, with a special focus on 9-*O*-acetylated *N*-acetylneuraminic acid (Neu5,9Ac), *N*-glycolylneuraminic acid (Neu5Gc), deaminoneuraminic acid (Kdn), *O*-sulfated Sia (SiaS), and di-, oligo-, and polysialic acid (diSia/oligoSia/polySia) in glycoproteins and glycolipids. Much more attention has been paid to these Sia and sialoglycoconjugates during the last decade, in terms of regulation of the immune system, neural development and function, tumorigenesis, and aging.

Keywords Chemical analysis · Deaminoneuraminic acid · Disialic acid · DMB derivatization · Immunochemical analysis · Immunohistochemistry · *N*-Glycolylneuraminic acid · *O*-Acetylated sialic acid · Oligosialic acid · *O*-Sulfated sialic acid · Polysialic acid

Contents

1	Introduction	76
1.1	<i>O</i> -Acetylated Sialic Acids (SiaAc)	77
1.2	<i>N</i> -Glycolylneuraminic Acid (Neu5Gc)	78
1.3	Deaminoneuraminic Acid (Kdn)	78
1.4	<i>O</i> -Sulfated Sialic Acids (SiaS)	78
1.5	Di-, Oligo-, and Polysialic Acids (diSia/oligoSia/polySia)	79

K. Kitajima (✉) and C. Sato
Bioscience and Biotechnology Center, Nagoya University, Nagoya, Japan
e-mail: kitajima@agr.nagoya-u.ac.jp

N. Varki
Department of Pathology, University of California, San Diego, CA, USA

2	Chemical Analyses of Sialic Acid and Sialoglycoconjugates	80
2.1	Detection of Modified Sialic Acids	80
2.2	Detection of Oligo/Polysialic Acids	82
3	Immunochemical Analyses of Sialic Acid and Sialoglycoconjugates	86
3.1	Immunochemical Detection of Kdn and <i>O</i> -Sulfated Sialic Acids	86
3.2	Immunochemical Detection of Oligo- and Polysialic Acids by Immunoblotting and ELISA	87
4	Immunohistochemistry for Specific Sialic Acids	90
4.1	Detection of 9- <i>O</i> -Acetylated Sialic Acids in Mammalian Tissues	92
4.2	Detection of Neu5Gc in Human Tissues	95
5	Conclusion and Perspectives	96
	References	97

1 Introduction

All cells are covered with glycoproteins and glycolipids, the biosynthesis of which takes place in the endoplasmic reticulum and Golgi compartments, and involves a variety of enzymes. The expression of some of these enzymes has been shown to be important in embryogenesis, cancer, pathogen recognition, and inflammation. *N*- and *O*-glycans and glycosphingolipids are often terminated with sialic acid (Sia), a family of nine carbon carboxylated monosaccharides, with high structural diversity. Sia consists of *N*-acetylneuraminic acid (Neu5Ac), *N*-glycolylneuraminic acid (Neu5Gc), deaminoneuraminic acid (Kdn), and their derivatives with modifications, such as acetylation, lactylation, methylation, and sulfation at the 4, 7, 8, and 9 positions (Fig. 1) [1–3]. These modifications of Sia have been implicated in embryonic development, protection from microbes and viruses, and modulation of complement activation [1, 2].

Sia exists as either free or bound sugar linked to galactose, *N*-acetylgalactosamine, and other sugar residues in various positions. Furthermore, Sia links to Sia itself to form dimeric, oligomeric, and polymeric structures (diSia/oligoSia/polySia). Thus, Sia shows extremely high structural diversity due to monosaccharide species as well as linkage modes. No other monosaccharide exhibits this structural diversity. Although there are some reports showing the biological significance of Neu5Gc and Neu5,9Ac [1, 4–6], for the majority of modified Sia, biological roles have not been well elucidated. For this purpose, specific methods or probes to detect specifically modified Sia are absolutely necessary. In this chapter we summarize current chemical and immunochemical methods to detect modified Sia species, with a special focus on *O*-acetylated Neu5Ac, Neu5Gc, Kdn, *O*-sulfated Sia, as well as diSia, oligoSia, and polySia. Because other chapters in this volume specifically focus on *O*-acetylated Neu5Ac [7], Neu5Gc [8], and polySia [9], we restrict ourselves in this chapter to a brief introduction of the different Sia species.

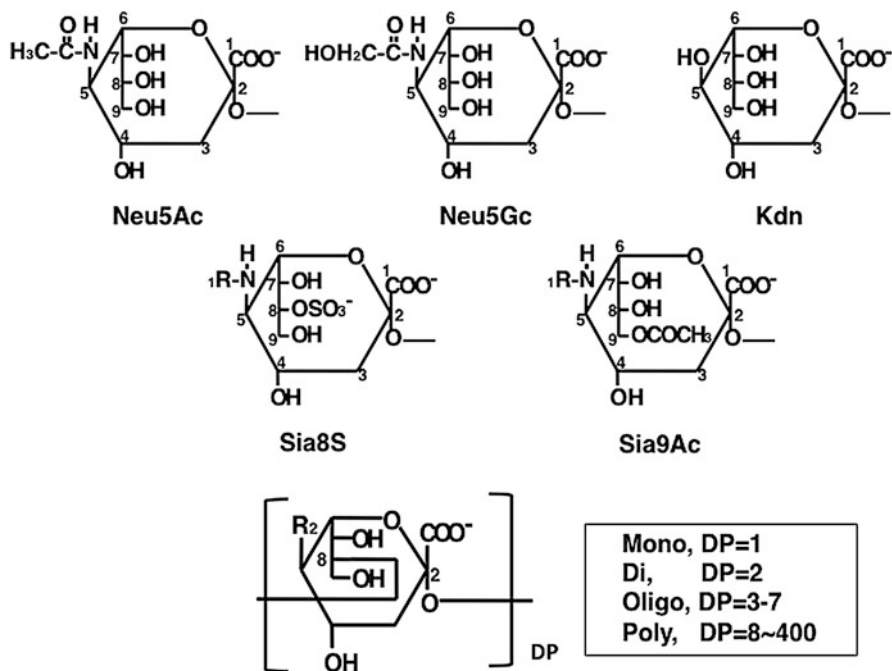


Fig. 1 Structures of sialic acid and sialoglycans described in this chapter. R₁: $\text{CH}_3\text{CO}-$, Neu5Ac; $\text{HOCH}_2\text{CO}-$, Neu5Gc. R₂: $\text{CH}_3\text{CONH}-$, Neu5Ac; $\text{HOCH}_2\text{CONH}-$, Neu5Gc; $\text{HO}-$, Kdn

1.1 O-Acetylated Sialic Acids (SiaAc)

The *O*-acetylated Sia (SiaAc) at the 4, 7, 8, and 9 positions frequently occurs in glycoproteins and glycolipids. SiaAc is expressed in cell type-specific and developmental stage-specific manners, and is involved in various biological phenomena, such as immune reactions, virus-cell adhesion, malignancy, and others [1, 10]. Sialate:*O*-acetyl-transferases [1, 11] and sialate:*O*-acetyl-esterases [12–14] are important enzymes for acetylation and deacetylation of Sia. The importance of 9-*O*-acetylated Sia during early development was evidenced by the transgenic expression of Influenza C hemagglutinin esterase in mice, leading to mice deficient in 9-*O*-acetylated Sia [4]. Recent analyses on sialate:*O*-acetyl-esterase deficient mice showed impairment of B cell-related functions that could be related to the inhibitory activity of siglec-2 ([5]; siglecs are discussed by Schwardt et al. [15]).

1.2 *N-Glycolylneuraminic Acid (Neu5Gc)*

Neu5Gc is broadly expressed in mammals. Neu5Gc is synthesized by the CMP-Neu5Ac hydroxylase (CMAH), a converting enzyme of CMP-Neu5Ac to CMP-Neu5Gc [16, 17]. Due to a genetic defect in *CMAH*, humans do not express a functional CMAH and thus do not endogenously produce Neu5Gc. However, Neu5Gc can be metabolically incorporated from dietary sources (particularly red meat and milk) into glycoproteins and glycolipids. Human fetuses, tumors, and some normal tissues were shown by mass spectrometry to incorporate Neu5Gc. Moreover, Neu5Gc has been demonstrated to be metabolically incorporated and covalently expressed on cultured human cell surfaces. Metabolically incorporated Neu5Gc is immunogenic in humans [18], and the presence of human anti-Neu5Gc antibodies has been associated with tumor progression [19] and vascular inflammation [20]. For additional details see Shilova et al. [21] and Davies and Varki [8].

1.3 *Deaminoneuraminic Acid (Kdn)*

Kdn contains a hydroxyl group at the C-5 position, instead of the acetamide-group in Neu5Ac (Fig. 1). Kdn was discovered in glycoproteins from rainbow trout eggs [22, 23]. In mammalian tissues, Kdn occurs at a very low level and amounts to <1% of total Sia. However, expression of Kdn increases in certain ovarian cancers, and the amount is proportional to the degree of malignancy [24]. Under hypoxic conditions free Kdn increases in mammalian tumor cells. As hypoxia-resistance is a frequent property found in tumors, it is likely that Kdn supports this state of mammalian cells [25]. In rat liver, cytosolic Kdn increases in an age-dependent manner. The level of free Kdn was reported to be high in erythrocytes of the umbilical vein. These findings demonstrate that attention has to be paid to the physiological significance of Kdn in malignancy and in aging [23]. Of interest in this context, Kdn residues, in contrast to Neu5Ac and Neu5Gc, are resistant to known bacterial and mammalian sialidases [22, 23]. In salmonid fish, Kdn caps polySia chains in egg polysialoglycoproteins, thus stopping chain elongation and possibly protecting polySia from bacterial sialidase attacks [22, 23]. Biological roles of Kdn in mammals remain to be elucidated, but increase of the Kdn level in cells and tissues may be a biomarker for some cancers and in aging cells.

1.4 *O-Sulfated Sialic Acids (SiaS)*

Sias are sometimes esterified with sulfate group to give *O*-sulfated Sia (SiaS). Methods for the detection of SiaS are not so well developed, and occurrence of SiaS has been only demonstrated in a limited number of animal species so far

[26–34]. In sea urchin, 8-*O*-sulfated Neu5Ac (Neu5Ac8S) and Neu5Gc8S are present in glycolipids [27–29], in glycoproteins from the vitelline layer polysialylated glycoprotein [30], and in the sperm flagellar glycoprotein, flagelliasialin [31, 32]. In mammals, Neu5Ac8S and Neu5Gc8S residues are found in gangliosides from bovine gastric mucosa [33, 34], and Neu5Ac8S was detected in tissues from human, rat, and mouse [35, 36]. The glycoconjugates carrying SiaS in mammals have remained undetermined with the exception of the bovine gastric gangliosides [33, 34]. Similarly, the biological significance of the SiaS in animal tissues is not known. It is interesting, however, that chemically synthesized SiaS-containing compounds have been applied as artificial inhibitors for fertilization [37] and bacterial infections [38], irrespective of the very limited knowledge on the natural occurrence of SiaS. These facts point to the potential importance of the SiaS residues in various biologic processes.

1.5 Di-, Oligo-, and Polysialic Acids (diSia/oligoSia/polySia)

In most cases, Sias are present as α 2,3- or α 2,6-linked monosialyl residues at the non-reducing terminal positions of glycan chains on glycoproteins and glycolipids. In two cases Sia has been identified as an internal sugar between neutral sugars: Fuc α 1 \rightarrow O_{glycolyl}Neu5Gc α 2-4Neu5Ac α 2-6Glc1- in echinoderms [39] and a polymer of \rightarrow 4Neu5Ac α 2 \rightarrow 6Glc/Gal1 disaccharide in the capsular polysaccharides of *Neisseria meningitidis* serogroups Y and W-135, respectively [40]. Moreover, Sias can be linked to each other to form diSia, oligoSia, and polySia (by definition DP > 8 is termed polySia). First identified in the glycocalyx of neuroinvasive bacteria, polySia has so far been demonstrated to be a posttranslational modification on six glycoproteins, the fish egg polysialoglycoprotein (PSGP), the neural cell adhesion molecule (NCAM), a voltage-gated sodium channel in eel, CD36 in human milk, NRP2 in human lymphocytes, and SynCAM-1 in mouse brain (for recent reviews see [41–43]). Polysialylation occurs in some tumors (probably on NCAM) and is involved in metastasis. By virtue of its net negative charge at physiological pH and exclusive volume, polySia serves as a mediator of ligand–receptor and cell–cell interactions via an anti-adhesive effect [44]. In addition, polySia has been demonstrated to function as a reservoir molecule for BDNF, dopamine, and FGF-2 [41, 43, 45–49] and to be involved in the regulation of ion transport via interactions with channels [32, 50–52]. Compared to what is known about polySia, the information on oligoSia is limited [43, 53, 54]. DiSia and oligoSia are common glyco-epitopes between glycolipids and glycoproteins. The functions of diSia on the glycolipids (GD3 and GT1b) have been well studied, while far less knowledge about the functions of diSia and oligoSia on the glycoproteins has been reported [55–57]. Interestingly, oligoSia and polySia with the degree of polymerization up to 16 have been recently found in glycolipids of sea urchin sperm [58], although their functions remain to be elucidated.

2 Chemical Analyses of Sialic Acid and Sialoglycoconjugates

2.1 Detection of Modified Sialic Acids

2.1.1 Colorimetric Analyses

To quantitate the amount of Sia at 0.1–100 μg , colorimetric analyses are most common with simple methods, and include the thiobarbituric acid [59, 60] and resorcinol methods [61]. In the resorcinol method, Neu5Ac and Neu5Gc in either free or bound form can be equally detected, while Kdn gives no color. In the thiobarbituric acid, Neu5Ac and Neu5Gc are detected only in free form, while Kdn can be detected in both free and bound forms. To quantitate the amount of Sia below 0.1 μg , a highly sensitive fluorometric high-performance liquid chromatography (HPLC)-based method is usually used. Here we focus on the detailed description of fluorometric HPLC technology.

2.1.2 Fluorometric HPLC Analysis

Fluorometric HPLC analysis is currently the most sensitive and reliable method for quantitative detection of Sia in the pmol range. In this analysis, a free form of various Sia species is labeled with the fluorescent dye 1,2-diamino-4,5-methylene-dioxybenzene (DMB), and fluorometrically analyzed on HPLC [62, 63]. DMB is a reagent which specifically reacts with α -keto acids, and detects not only Sia but also typical α -keto acids such as pyruvate and α -ketoglutarate. Fluorometric HPLC analysis consists of the following steps: step 1, acid hydrolysis of sialoglycoconjugates to release free Sia; step 2, fluorescent labeling of released Sia with DMB; step 3, separation and quantification of DMB-labeled Sia on HPLC.

Fluorometric HPLC analysis involves two acidic conditions. One is a hydrolysis, 0.1 N trifluoroacetic acid at 80°C for 2 h to release free Sia from sialoglycoconjugates at step 1. The second is the DMB labeling which takes place in 0.01 N trifluoroacetic acid at 50°C for 3 h (step 2). Attention must therefore be given to the acid lability of some Sia substituents. No significant degradation of Neu5Ac, Neu5Gc, and Kdn has been reported. For SiaS, the glycosidic bonds of SiaS were completely hydrolyzed under the hydrolysis conditions, while their sulfate esters are hydrolyzed by, at most, 4% [64]. No apparent desulfation occurred under the labeling conditions. The *O*-acetyl groups on Sia are very labile to basic conditions and more stable under acidic conditions. Although no quantitative data are available for the lability of *O*-acetyl group during the DMB-derivatization procedures, quantitative analysis is basically possible as long as authentic *O*-acetylated Sia compounds are used as a standard. It is of interest to note that the 8-*O*-acetyl ester is

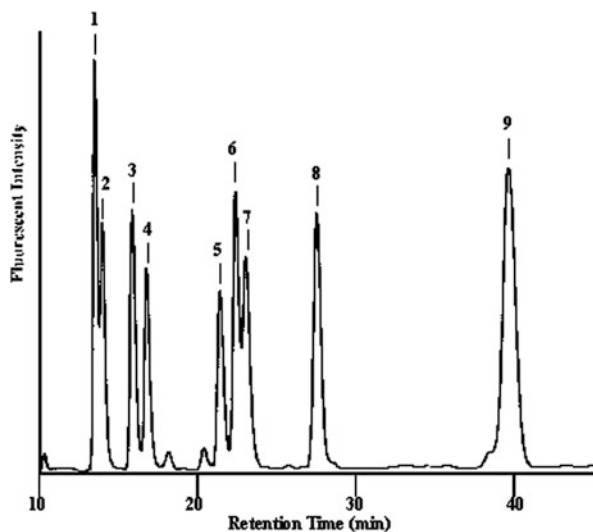


Fig. 2 Fluorometric HPLC of sialic acid. DMB-derivatives of Neu5Ac, Neu5Gc, Kdn, and their 8- or 9-*O*-sulfated forms are shown. DMB-derivatives were applied to an ODS column, eluted with acetonitrile/methanol/0.05% trifluoroacetic acid (4: 6: 90 by volume), and detected fluorometrically at excitation/emission 373/448 nm. The retention times (min) for the DMB-derivatives are: 13.9, DMB-Kdn9S (1); 14.4, DMB-Kdn8S (2); 16.3, DMB-Neu5Gc9S (3); 17.2, DMB-Neu5Gc8S (4); 21.8, DMB-Neu5Ac9S (5); 22.8, DMB-Kdn (6); 23.4, DMB-Neu5Ac8S (7); 27.9, DMB-Neu5Gc (8); 40.0, DMB-Neu5Ac (9)

more labile than a 9-*O*-acetyl ester, and acetyl-groups tend to migrate to the C-9 position at the end, even under neutral conditions at room temperature [65]. However, no such migration takes place with the 8-*O*-sulfate ester of Sia.

In Step 3, DMB derivatives of Neu5Ac, Neu5Gc, and Kdn are separately quantitated by HPLC on an octadecylsilyl (ODS) column using acetonitrile/methanol/water (9:7:84 by volume) as the elution solution [24]. The *O*-acetylated Sia are eluted more slowly than the corresponding non-*O*-acetylated Sia. To separate the various SiaS from each other, an acidic solution of acetonitrile/methanol/0.05% trifluoroacetic acid (9:7:84 by volume) [64] is recommended. This solution will enable the separation of DMB derivatives of Neu5Ac8S, Neu5Ac9S, Neu5Gc8S, and Neu5Gc9S. For the separation of DMB derivatives of Kdn8S and Kdn9S, a solution with higher polarity, acetonitrile/methanol/0.05% trifluoroacetic acid (4:6:90 by volume), is appropriate [64]. The described protocol enabled the separation of DMB derivatives of nine different sulfated and non-sulfated Neu5Ac, Neu5Gc, and Kdn (Fig. 2). The relative peak area represents the proportion of the peak area of each DMB-Sia to that of DMB-Neu5Ac, and can be used as an index for relative detection efficiency. Typical relative peak areas to that of DMB-Neu5Ac are shown in Table 1.

Table 1 Typical relative peak areas for the DMB-derivatives to that for DMB-Neu5Ac

DMB derivative of	Peak area/pmol ^a	Relative peak area ^b
Neu5Ac	5.0×10^6	1.0
Neu5Ac8S	2.8×10^6	0.65 ± 0.09
Neu5Ac9S	2.3×10^6	0.48 ± 0.02
Neu5Gc	3.8×10^6	0.72 ± 0.05
Neu5Gc8S	3.0×10^6	0.63 ± 0.04
Neu5Gc9S	3.3×10^6	0.64 ± 0.04
Kdn	2.3×10^6	0.46 ± 0.01
Kdn8S	0.9×10^6	0.21 ± 0.04
Kdn9S	2.0×10^6	0.40 ± 0.02

^aThe fluorescent peak area is divided by the amounts (pmol) of the indicated Sia and SiaS glycosides

^bRelative peak areas of the Sia and SiaS derivatives are shown as peak area/pmol values relative to that of Neu5Ac glycoside. The deviations from the average are shown for two independent experiments

2.2 Detection of Oligo/Polysialic Acids

When samples containing diSia, oligoSia, and polySia structures at 10–100 μg are analyzed, conventional methods including methylation analysis [66], NMR (nuclear magnetic resonance) [67], and mild acid hydrolysis-TLC (thin layer chromatography) [68] can be applied. However, the amount of these types of glycoproteins is often too small to be analyzed by conventional methods. The fact that the di-, oligo-, and polySia-modification of glycoproteins has not often been reported suggests that these species in an organism is rare. As highly sensitive chemical methods to analyze minute amounts of di-, oligo-, and polySia were developed using highly sensitive fluorescent reagents [62, 69–72], the number of studies identifying di-, oligo-, and polySia increased gradually [43]. Therefore, in this chapter we are focusing on the chemical and immunochemical detection of small amounts (picogram to nanogram amounts) of sialic acid which may serve as powerful tools to reveal the importance of di-, oligo-, and polySia in nature.

The polySia glycotope exhibits structural diversity in the sialic acid components (Neu5Ac, Neu5Gc, and Kdn) in the inter-sialyl linkages ($\alpha 2 \rightarrow 5O_{\text{glycolyl}}$, $\alpha 2 \rightarrow 8$, $\alpha 2 \rightarrow 9$, and $\alpha 2 \rightarrow 8/9$) and in the degree of polymerization (DP) [43, 68]. With a highly specific reagent for α -keto acid [62], a specific labeling of di-, oligo-, and polySia became possible and products could be detected by anion exchange chromatography. This broadly used method was first developed by us [70] and was further improved in later studies [73]. With this method, di-, oligo-, and polySia on glycoproteins were frequently demonstrated [43, 53, 54].

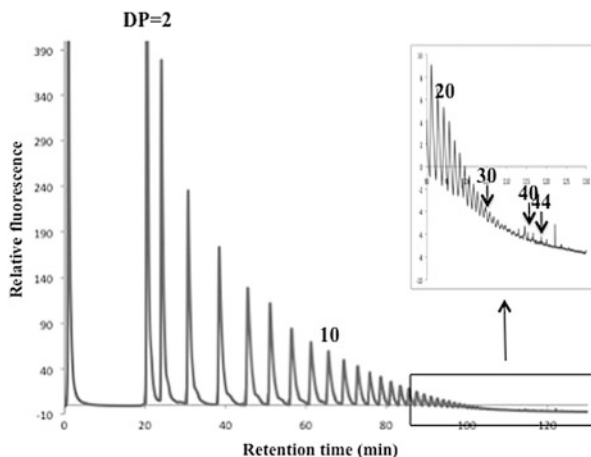


Fig. 3 Mild acid hydrolysis-fluorometric HPLC analysis. Mini Q anion exchange chromatography of $\alpha 2 \rightarrow 8$ -linked di/oligo/polyNeu5Ac-DMB. $\alpha 2 \rightarrow 8$ -Linked oligo/polyNeu5Ac was labeled with DMB and applied to a mini Q HR5/5 anion exchange column (1 mL, Cl^- -form). The column was eluted with 5 mM Tris-HCl (pH 8.0) with a gradient from 0 to 0.3 M NaCl for 75 min and 0.3 M NaCl to 0.4 M NaCl for 120 min after 15 min wash. The elution was monitored by a fluorescence detector (set at wavelength of 373 nm excitation and 448 nm emission). Each peak is assigned from the order of elution, based on DP

2.2.1 Mild Acid Hydrolysis-Fluorometric HPLC Analysis

Gangliosides, glycoproteins in solution or blotted on polyvinylidene difluoride (PVDF) membrane (Immobilon P, Millipore, U.S.A) are hydrolyzed with 0.01 N trifluoroacetic acid (TFA) at 50°C for 1 h to release di-, oligo-, and polySia units from the glycoconjugates. The samples were then freeze dried. To release di-, oligo-, and polySia, 20 μL of 0.01 N TFA and 20 μL of 7 mM DMB solution in 5.0 mM TFA containing 1 M 2-mercaptoethanol and 18 mM sodium hydrosulfite, are added to the samples. These samples are incubated at 50°C for 1–2 h or at 4°C overnight. The DMB-labeled samples are applied to an HPLC analysis. HPLC equipped with a Mono or Mini Q HR5/5 (0.5 \times 5 cm, GE, Uppsala, Sweden), Resource Q (1 mL, GE, Uppsala, Sweden), CarbopacPA-100, or DNapak1 PA-120 (4 \times 250 mm, Dionex) anion exchange column, and a fluorescence detector (FP-2025, JASCO). After equilibrating the column with 5 mM to 20 mM Tris-HCl (pH 8.0) at 26°C , samples are applied and the DMB-labeled di-, oligo-, and polySia are eluted at a flow rate of 0.5–1.0 mL/min with a linear gradient of NaCl (0 \rightarrow 0.4 M) after 15–30 min wash with 5–20 mM Tris-HCl (pH 8.0) (flow through). The fluorescence of the DMB-labeled samples is detected with a fluorescence detector at excitation 373 nm and emission 448 nm. The separation is dependent on the anion exchange column and buffer conditions. Mono Q and Resource Q are suitable for shorter oligo- and polymers. Mini Q and Carbopac PA-100 have almost the same ability to separate polymers with DPs ranging from 2 to DP 50–90 (Fig. 3).

Additionally, this method can be applied to determine the glycosidic linkage ($\alpha 2 \rightarrow 8$, $\alpha 2 \rightarrow 9$, and $\alpha 2 \rightarrow 5$) of diSia units (Neu5Ac \rightarrow Neu5Ac, Neu5Ac \rightarrow Neu5Gc, Neu5Gc \rightarrow Neu5Ac, Neu5Gc \rightarrow Neu5Gc, Kdn \rightarrow Kdn, and so on) and their component Sia species [70] according to their elution time. If the sample is pure sialyloligo or -polymer, it can be detected on a UV detector or a pulse amperometric detector (PAD) [71], instead of the DMB-derivatization/fluorometric detection method.

To increase the sensitivity, especially in higher DP, because the labeling site is one per chain, it is recommended to separate the polySia on an anion exchange column and to label each fraction with DMB labeling after hydrolysis of the sample with 0.1 N TFA at 80°C for 1–2 h [72]. The sensitivity increases depending on their DP.

2.2.2 Fluorometric C₇/C₉ Analysis

The periodate treatment of sialic acid residing in non-reducing terminal ends is a well-known method to modify the terminal sialic acid [74]. To observe and quantify the internal sialic acids of $\alpha 2 \rightarrow 8$ -linked di-, oligo-, and polySia-containing glycoconjugates, sensitive chemical methods were developed with highly sensitive fluorescent labeling reagents (DMB) as described above [69].

To perform the fluorometric C₇/C₉ analysis, several reagents are prepared. Solutions A–G are prepared as follows: solution A, 40 mM sodium acetate buffer (pH 5.5); solution B, 0.25 M periodate; solution C, 3% ethyleneglycol; solution D, 0.2 M sodium borohydride in 0.2 M sodium borate buffer (pH 8.0); solution E, 0.2 M trifluoroacetic acid (TFA); solution F, 0.01 M trifluoroacetic acid; solution G, 7 mM 1,2-diamino-4,5-methylenedioxybenzene (DMB) in 5 mM trifluoroacetic acid containing 1 M 2-mercaptoethanol and 18 mM sodium hydrosulfite. For glycoproteins or oligosaccharides in solution, samples (0.25–1,000 ng as Sia) are dissolved in 25 μ L of solution A and 2 μ L of solution B are added. After being left at 0°C for 3 h in the dark, 5 μ L of solution C and 32 μ L of solution D are added successively and allowed to stand at 0°C overnight. To the resultant mixture is added a pre-determined amount of Kdo (for example, as an internal standard for quantitating the resultant sialic acids) and the volume is set to 100 μ L with water. Following further addition of 100 μ L of solution E to adjust to 0.1 M TFA, the mixture is hydrolyzed at 80°C for 2–4 h. The hydrolysate is lyophilized by a Speed Vac. After addition of 20 μ L of solution F to the samples, 20 μ L of solution G is added and the samples are incubated at 50°C for 2 h (it depends on the DP). The resulting supernatants are applied to an HPLC for analysis. This method can also be applied to the samples blotted on the PVDF membrane.

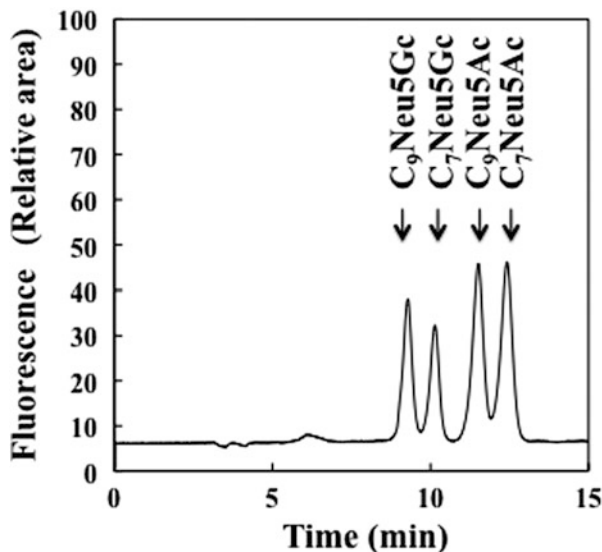


Fig. 4 HPLC profile in the fluorometric C_7/C_9 analysis. A typical elution profile of DMB derivatives of C_7 -analogues and authentic sialic acids (C_9) on the fluorometric HPLC. Disialyl sialitols, 12.5 ng each of Neu5Ac α 2 \rightarrow 8Neu5Ac α 2 \rightarrow 8-Neu5Ac-ol and Neu5Gc α 2 \rightarrow 8Neu5Gc α 2 \rightarrow 8-Neu5Gc-ol, were subjected to the periodate oxidation/reduction/hydrolysis, DMB derivatization, and fluorometric HPLC on a TSK-gel ODS-120T column (250 \times 4.6 mm i.d.). The column was eluted with acetonitrile/methanol/water (9:7:84 by volume) at 1.0 mL/min at 26°C. Elution profile was monitored by measurement of fluorescence: excitation, 373 nm; emission, 448 nm

The DMB-labeled sialic acids are applied to the HPLC equipped with a TSK-gel ODS-120T column (250 \times 4.6 mm i.d., Tosoh), and a fluorescence detector (FP-2025, JASCO). The column is equilibrated using acetonitrile/methanol/water (9:7:84 by volume) at 26°C. Then 2–20 μ L of the supernatants are applied to HPLC analysis for an isocratically flow at 1.0 mL/min and the DMB-labeled sialic acid is detected with a fluorescence detector at an excitation of 373 nm and an emission at 448 nm (Fig. 4). The estimate of the ratios of the quantity of internal sialic acid residues (C_9 (Sia)) to that of total sialic acid residues (C_7 (Sia) + C_9 (Sia)) can then be obtained.

The following limitations should be noted with respect to this method. First, this method is applicable to only α 2 \rightarrow 8-linked oligo/polymer of *N*-acylneuraminic acid, and cannot be used for DP analyses of α 2 \rightarrow 9, α 2 \rightarrow 8/ α 2 \rightarrow 9-mixed linkage polymers, or α 2 \rightarrow 5*O*_{glycolyl}-linkage. Second, the C_9 -derivatives formed do not always arise from α 2 \rightarrow 8-linkages, because 8-*O*-substituted Neu5Acyl residues may also give the same C_9 -derivatives. Therefore, mild alkali treatment of samples is usually carried out prior to periodate oxidation. Third, the molar proportion of

C9-derivatives to C7-derivatives does not directly represent the DP unless it is a linear polySia chain. Thus, the method does not yield the DP for multiply sialylated chains present in the same sample. In general, glycoproteins have more than one glycan chain and even one glycan chain bears two to four non-reducing terminal residues that may be terminated by monoSia or oligoSia residues.

3 Immunochemical Analyses of Sialic Acid and Sialoglycoconjugates

3.1 Immunochemical Detection of Kdn and O-Sulfated Sialic Acids

3.1.1 Immunodetection of Kdn Using Kdn-Specific Antibodies and Kdn'ase Sm

Kdn residues are resistant to various bacterial sialidases that are often used as reagents. However, a bacterial sialidase from *Sphingobacterium multivorum*, named Kdn'ase Sm, has a unique property in that it cannot cleave the glycosidic linkages of Neu5Ac and Neu5Gc in any free glycans and glycoconjugates but can specifically cleave the Kdn glycosidic linkages in free glycans, glycolipids and glycoproteins [75]. Similar Kdn'ases are also known in hepatopancreas of oyster. These enzymes strictly discriminate Kdn from Neu5Ac or Neu5Gc, and thus provide evidence for Kdn-specific recognition phenomena in these organisms.

Two monoclonal antibodies recognizing Kdn-containing epitopes have been developed. Kdn8Kdn, a mouse IgM, recognizes α 2,8-linked oligo/polyKdn with more than two residues [76]. mAb.kdn3G, a mouse IgG, recognizes the Kdn α 2,3Gal-structure [77]. In immunohistochemistry, enzyme-linked immunosorbent assay (ELISA), and Western blotting using these antibodies, the results would be more reliable, if the loss of antigenicity by digestion with Kdn'ase Sm, a bacterial Kdn-specific sialidase (see below), was confirmed. Of the known lectins that recognize Neu5Ac, *Sambucus sieboldiana* lectin (SSA) can recognize the Kdn α 2,6Gal-linkage better than Neu5Ac α 2,6Gal-linkage [78]. In a combination of SSA and Kdn'ase Sm, Kdn α 2,6Gal-structure can be detected.

Kdn can be detected in free glycans, glycolipids, and glycoproteins in various cells and tissues from animals that express Neu5Ac. The expression level is very low in mammals.

3.1.2 Immunodetection of SiaS Using Anti-SiaS Antibodies

Two monoclonal antibodies, mAb.3G9 and mAb.2C4, specifically recognizing SiaS, have been reported. The mAb.3G9 is a mouse IgM, and is highly specific for Neu5Ac8S. The antibody was generated using sea urchin sperm as an immunogen

[79]. The mAb.2C4 is also a mouse IgM and recognizes Neu5Ac8S and Neu5Gc8S, which can compensate for the binding specificity of mAb.3G9. This antibody was prepared using sea urchin egg low density detergent-insoluble membrane as an immunogen [64]. Detection limits of mAb.3G9 (1 $\mu\text{g}/\text{mL}$) and mAb.2C4 (culture supernatant) for Neu5Ac8S $\alpha 2 \rightarrow 8$ Neu5Ac $\alpha 2 \rightarrow 6$ Glc-Cer in ELISA were 0.94 and 1.9 $\mu\text{g}/\text{well}$, respectively [64]. As in the case with Kdn'ase Sm in immunodetection of Kdn, an enzyme specifically destroying the SiaS epitopes would be a powerful tool for reliable detection. Commercially available arylsulfatases, which can remove sulfate group from sulfatide containing 3-*O*-sulfated Gal residue, have no activities on sulfate group on Sia. There are no known sulfatases that can act on the sulfate ester on SiaS. Therefore, immunochemical detection should be confirmed at least by chemical evidence to obtain reliable conclusions of the localization of SiaS.

Western blotting of sperm lysate of *Hemicentrotus pulcherrimus* using mAb.3G9 detects unique smear at 40–80 kDa, which corresponds to flagelliasialin containing Neu5Ac8S-capped $\alpha 2,9$ -Neu5Ac-linked polymers [31, 32].

3.2 Immunochemical Detection of Oligo- and Polysialic Acids by Immunoblotting and ELISA

3.2.1 Immunodetection of Antibodies Recognizing di/oligo/polySia

Antibodies are powerful tools for studying the structure and function of glycotopes. However, it is very important that the immunospecificities toward linkage, DP, and component are determined in detail before use.

For $\alpha 2 \rightarrow 8$ -linked polySia glycotopes, several antibodies have been developed and used for the past 2 decades. Among the “so-called anti-polySia antibodies,” the horse polyclonal antibodies H.46 [80] and mouse mAb 735 [81] had been the only two whose immunospecificities were specifically determined by inhibition assays [82, 83]. The immunospecificity of the majority of other “anti-polySia antibodies” remained unknown until a comprehensive examination of the immunospecificity of these “anti-polySia” antibodies was carried out using an ELISA-based method with phosphatidylethanolamine-conjugated oligo- and polySia chains as antigens [84]. It was demonstrated that the “anti-polySia antibodies” discriminate in terms of the species of Sia residues and chain length. Our group has also developed a new anti-diSia antibody using copolymers of $\alpha 2 \rightarrow 8$ -linked *N*-acetylneuraminyl *p*-vinylbenzylamide and acrylamide as an immunogen [85]. Thus, a large number of antibodies recognizing di-, oligo-, and/or polySia structures with defined specificities now exist and are summarized in Table 2.

Table 2 Antigenic specificities and class of anti-oligo/polySia antibodies

Name of antibody	Animal origin ^a and immunoglobulin-type ^b	Sia in oligo/polySia recognized	Specificity on DP
<Group I > <i>anti-polySia antibody</i>			
H.46	ho, poly, IgM	Neu5Ac	DP \geq 8
735	mo, mono, IgG2a	Neu5Ac	DP \geq 11
<Group II > <i>anti-oligo + polySia antibody</i>			
12E3	mo, mono, IgM	Neu5Ac	DP \geq 5
5A5	mo, mono, IgM	Neu5Ac	DP \geq 3
2-2B	mo, mono, IgM	Neu5Ac	DP \geq 4
OL.28	mo, mono, IgM	Neu5Ac	DP \geq 4
2-4B	mo, mono, IgM	Neu5Gc	DP \geq 2
kdn8kdn	mo, mono, IgM	KDN	DP \geq 2
4F7	mo, mono, IgG	Neu5Ac ^c	n.d. ^d
<Group III > <i>anti-oligoSia antibody</i>			
S2-566	mo, mono, IgM	Neu5Ac	DP = 2 ^e
1E6	mo, mono, IgM	Neu5Ac	DP = 2
A2B5	mo, mono, IgM	Neu5Ac	DP = 3
AC1	mo, mono, IgG3	Neu5Gc	DP = 2–4

^aho horse, mo mouse

^bpoly polyclonal, mono monoclonal

^c α 2 \rightarrow 9-linkage specific

^dn.d. not determined

^eNeu5Ac α 2 \rightarrow 8Neu5Ac α 2 \rightarrow 3Gal. Gal residue is required

Interestingly, the anti-oligo/polySia antibodies can be classified into three groups, based on the immunospecificity for chain length and the involvement of the non-reducing terminus in the antibody recognition site. Group I consists of the “anti-polySia antibodies” that recognize chains of α 2 \rightarrow 8-linked Sia with a DP \geq 8, including fully extended polySia chains with a DP \sim 200–400. These antibodies recognize the helical conformation formed by Sia residues within the internal region of the polySia chains, but not the non-reducing terminal residues. Group II antibodies, designated as “anti-oligo + polySia antibodies,” recognize both oligoSia with DP 2–7 and polySia chains. These antibodies recognize the distal portion of oligo/polySia chains, including the non-reducing termini. Group III antibodies, designated as “anti-oligoSia antibodies,” recognize specific conformations of di- and oligoSia with a DP 2–4, but do not bind polySia. Group III antibodies can be further classified into two subgroups. One group is composed of “anti-diSia antibodies” that recognize the diNeu5Ac structure (S2-566, 1E6), and the other includes the “anti-oligoSia antibodies” (AC1). Group II and III antibodies would be useful for detecting and determining di- and oligoSia structures, in combination with treatment with exo- and endo-sialidases, as described below (Table 3). The antibodies can be applied to Western blotting (Fig. 5a), immunocytochemistry (Fig. 5b), and FACS analysis (Fig. 5c) [86].

Table 3 Reactivity of di-, oligo-, and polySia chains toward biochemical probes

Biochemical probes	diSia	oligoSia	polySia
	DP = 2	DP = 3-7	DP ≥ 8
Group I antibody	–	–	+
Group II antibody	–	+	+
Group III antibody	+	+ or –	–
Endo-sialidase (endoN)	–	– or + ^a	+
Endosialidase ^b	–	+	+
α2,3-Sialidase/α2,6-sialidase	–	–	–
α2,3-, α2,6, α2,8-Sialidase	+	+	+

+ reactive or sensitive, – unreactive or insensitive

^a+ in case of DP = 6,7

^bRefer to [91]

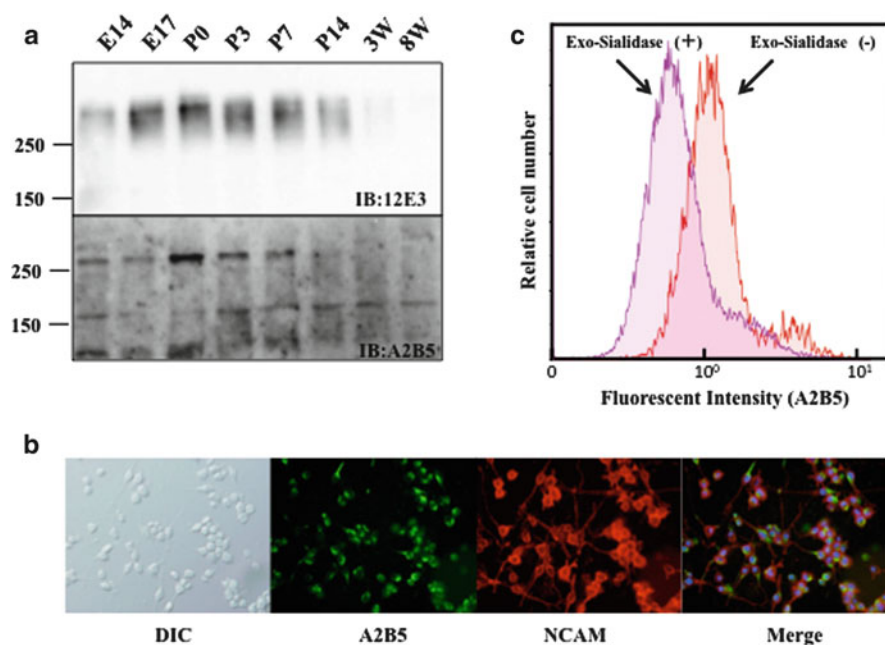


Fig. 5 TriNeu5Ac epitopes in mouse brain homogenates and Neuro2A cells. (a) Western blots of mouse brain homogenates with anti-trSia antibody (A2B5) and anti-oligo/polySia antibody (12E3). The photos are cited from [85]. (b) Cell surface staining with A2B5 using fluorescent microscopy. The mouse neuroblastoma cells were incubated with retinoic acid at 20 μ M in 2% FBS in DMEM and differentiated into neurons. At day 3, the cell surfaces were immunostained with anti-triSia antibody (A2B5, Green), and anti-NCAM antibody (Red). (c) Cell surface immunostaining with anti-triSia antibody using flowcytometer. Neuro2A cells were transfected with ST8Sia III gene although the cells have its enzyme endogenously. The stably ST8SiaIII expressing cells were immunostained with anti-triSia antibody before, and after exosialidase treatment

Surprisingly, anti-diSia antibodies obtained after immunization with gangliosides, developed, and used as anti-ganglioside antibodies can recognize cytosolic proteins such as carbonic anhydrase II [87] although they recognize diSia epitope on glycoproteins. These phenomena might explain the cytosolic immunostaining with those anti-ganglioside antibodies in part. The protein mimicry should be considered and confirm the epitope with those antibodies.

3.2.2 Use of Endo-Sialidases and Exo-Sialidases for the Detection of OligoSia and PolySia

Endo-sialidase (for a detailed description see Jakobsson et al. [88]) can serve as a specific molecular probe to detect and selectively modify $\alpha 2 \rightarrow 8$ -linked polySia chains. The soluble enzyme derived from bacteriophage K1F, designated endoN, catalyzes the depolymerization of polySia chains as follows: $(\rightarrow 8\text{Neu5Acyl}\alpha 2 \rightarrow)_n - X$ ($n \geq 5$) $\rightarrow (\rightarrow 8\text{Neu5Acyl}\alpha 2 \rightarrow)_{2-4} + (\rightarrow 8\text{Neu5Acyl}\alpha 2 \rightarrow)_2 - X$. Two other types of endo-sialidases whose substrate specificities are different from the endoN of bacteriophage K1F [89] have been isolated: endoNE [90] and an endosialidase [91] from a bacteriophage. The minimum chain length required for cleavage is $DP \geq 11$ and $DP \geq 3$, respectively. Exo-sialidases have been isolated that cleave specific linkages, for example, $\alpha 2,3$ -sialidase (NANase I), $\alpha 2,3$ -, $\alpha 2,6$ -sialidase (NANase II), $\alpha 2,3$ -, $\alpha 2,6$ -, $\alpha 2,8$ -sialidase (NANase III), and $\alpha 2,3$ -, $\alpha 2,6$ -, $\alpha 2,8$ -, and $\alpha 2,9$ -sialidase. As the endoN is insensitive toward di- and oligoSia structures and they should be cleavable by these exosialidases, it is possible to confirm the length of given di-, oligo-, and polySia chains in conjunction with endo- and exo-sialidase treatments before and after immunostaining with anti-diSia, oligoSia, and polySia antibodies (See Table 3) [43, 53, 54].

Finne et al. established a specific probe for detection of polySia from endoNE that inactivates its enzymatic activity but not its ability to bind to polySia [92]. They succeeded in the detection of polySia-NCAM with this probe.

4 Immunohistochemistry for Specific Sialic Acids

When tissues are removed from the body during surgery, or during an autopsy, the onset of autolysis occurs promptly. This can thwart efforts to isolate nucleotides or certain enzymes and proteins, or to perform high quality histology. Thus tissue ideally has to be flash-frozen for extracts, or frozen in cryoprotective agents, or fixed, using different procedure-dependent fixatives, for analysis using histopathology. Immunohistochemistry is the process whereby tissue sections are probed with

antibodies to determine which cells contain epitopes of interest. The bound antibody is then detected using secondary and/or tertiary reagents, using fluorescent tags, or enzyme labels. If using enzyme labels, detection is facilitated using a number of different substrates, which produce precipitates, which are visible under the microscope. The nuclei are counterstained so that the morphology of the tissue can then be recognized during the analysis of the slides. In all immunohistochemistry experiments, to determine specificity, it is important to use specific blocking or inhibitory reagents. Processing of tissues into paraffin destroys many antigenic epitopes and additional unmasking methods often have to be used to detect antigens that are left after the processing. However, some changes are irreversible, such as the extraction of glycolipids during the paraffin embedding process. Immunohistochemistry remains an exact science, and appropriate use of controls and blocking agents are critical while analyzing results.

Lectins are carbohydrate-binding proteins, usually derived from plants, and have been used as tools to detect changes in glycan branching presentation on different cells. Mice that have been genetically altered to be deficient for specific sialyltransferases are then useful as important controls to demonstrate specificity of some lectin binding patterns. This was demonstrated nicely in a report [93] describing mice deficient in the sialyltransferases ST6Gal-I (transfers α 2-6-linked Sia to Gal β 1-4GlcNAc units, mostly on *N*-glycans), or ST3Gal-I (transfers α 2-3-linked Sia to Gal β 1-3GalNAc units, mostly on *O*-glycans). *Sambucus nigra* agglutinin (SNA) staining was lost in essentially all tissues of adult ST6Gal-I null mice, indicating that this is the only enzyme generating the Sia α 2-6Gal β 1-4GlcNAc sequence. Lectin histochemistry with the α 2-3-Sia-specific *Maackia amurensis* Hemagglutinin (MAH or MAL-II) showed loss of binding in almost all tissues of ST3Gal-I null mice. However, use of the other isolectins of the *Maackia amurensis* seeds, the MAL-I lectin, or the MAA lectin, showed continued binding to many tissues from the ST3Gal-I null animals (Fig. 6).

Different isolectins have been isolated from *Maackia amurensis* seeds [94, 95] and have been used as probes for sialic acids that are α 2-3-linked to the penultimate galactose residues. The two most commonly used isolectins are the leucoagglutinin (MAL) and the hemagglutinin (MAH). Some commercial vendors use the term MAA to denote probes for sialic acids that are α 2-3-linked, but in fact MAA is a mixture of MAL and MAH [96, 97]. In both MAL and MAH the glycine and asparagines involved in sugar binding were substituted by lysine and aspartic acid, respectively, [94, 98] and shown to be important in sialic acid binding. Investigators, using biochemical methods, have described the carbohydrate binding specificity of these lectins, and some have also described the lectin histochemistry profile. However, many of the histochemistry studies use MAL or MAA, and only rarely has MAH been used to ascribe the binding being used to identify α 2-3-linked sialic acids. Most recently, it was shown that effective glycoanalysis with *Maackia amurensis* lectins requires a clear understanding of their binding specificities [99].

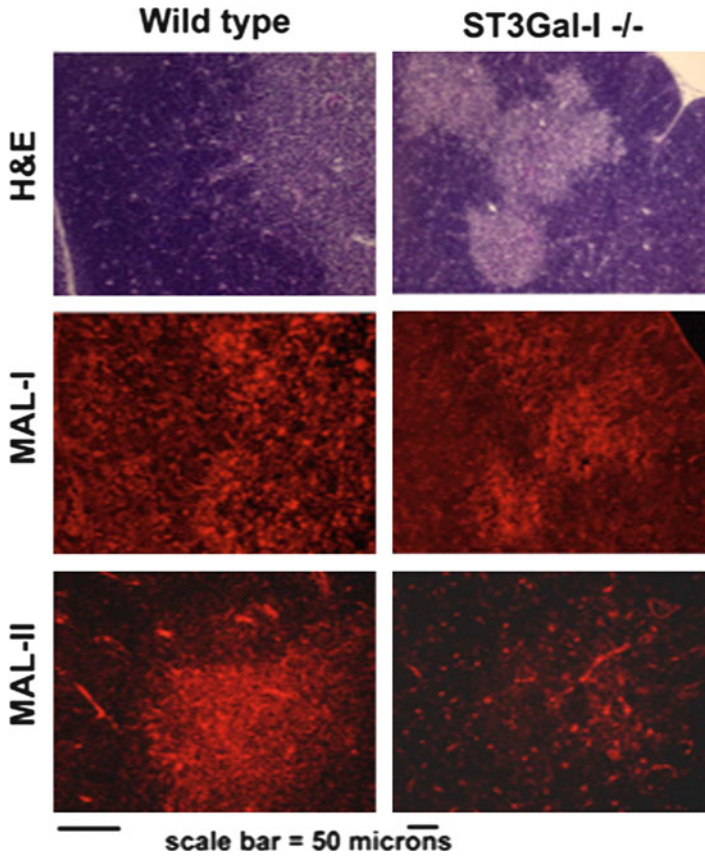


Fig. 6 Differences in binding to frozen sections of mouse thymus, from wild type animal (*left column*) and from ST3Gal-I homozygous mutant (*right column*) observed with MAL-I and MAL-II, detected using the Cy3 fluorescent streptavidin label. *Top row* shows hematoxylin and eosin stained sections of the thymus, with dark blue cortex and lighter staining medulla. *Middle row* shows binding of MAL-I to thymus from both wild type (*left*) and ST3Gal-I null animals (*right*). *Bottom row* shows binding of MAL-II (MAH) lectin to the thymic medulla of the wild type mouse (*left*) and only minimally to structures in the thymic medulla from ST3Gal-I null mouse (*right*). Scale bar = 500 μ m

4.1 Detection of 9-O-Acetylated Sialic Acids in Mammalian Tissues

In mammalian tissues, 9-O-acetylation is often associated with gut mucins, neural gangliosides, and erythrocytes. Direct demonstration of the presence of 9-O-acetylated Sias in mammalian tissues was indirect (modifications of the PAS

technique) until the development of a chimeric dual-functional probe derived from influenza-C hemagglutinin esterase (InfCHE). This probe was constructed using a soluble and versatile form of the InfCHE that retains both the hemagglutinin and esterase activity and also has the binding properties of IgG Fc. The esterase activity is very similar to that of the native protein, and pH conditions can be adjusted to remove either 9-*O*-acetyl groups alone or both 9- and 7-*O*-acetyl groups completely (by causing migration of the latter to the 9-carbon position). DFP (di-isopropyl fluorophosphates) inactivation stabilizes and “unmasks” the hemagglutinin activity, resulting in a probe that specifically detects 9-*O*-acetylated sialic acids at ambient temperatures. Since CHE-FcD and CHE-Fc differ only by a single di-isopropyl group, each can be used as control for the other. Thus, when the CHE-Fc is used to remove 9-*O*-acetyl esters, the CHE-FcD can be used as an inactive control. Conversely, when CHE-FcD is used to probe for 9-*O*-acetylated Sias, the CHE-Fc can be used as a control for nonspecific binding. The two forms can even be used sequentially; i.e., treatment with CHE-Fc can be used to remove the 9-*O*-acetyl esters prior to probing with CHE-FcD, leaving only nonspecific staining, if any. It should be noted that the CHE-Fc does have a masked hemagglutinin activity, which can give a weak positive response at low temperatures, and even at ambient temperatures if the density of *O*-acetylated Sias is very high [6, 100–105].

Tissue expression of 9-*O*-acetylated Sias was best analyzed using frozen sections, using the probe CHE-FcD pre-complexed to the secondary antibody. There were differences in expression of 9-*O*-acetylated Sias in tissues isolated from rats and those from mice. As shown Fig. 7, there was abundant expression in rat liver, rat lung, in the glomeruli of rat kidney, and gray matter of rat brain. However, although mouse red blood cells and blood vessels showed expression of 9-*O*-acetylated Sias with the CHE-FcD probe, expression was not robust in mouse liver and mouse lung, but was instead present in the T-cell areas of lymphoid follicles in spleen, in the mature lymphoid medulla of the thymus, gray matter of the brain, and adrenal medulla.

Since the submaxillary gland mucins of many species have been reported to have high concentrations of 9-*O*-acetylated Sias, it was surprising to find no significant staining in the parenchyma of the rat or mouse submaxillary gland. Lipid overlays and protein blots confirmed that this is not due to inaccessibility of molecules on the tissue section. It is possible that the sialic acids of rat submaxillary mucins are 4-*O*- rather than 9-*O*-acetylated [106].

The hemagglutinin esterases of influenza C viruses and certain nidoviruses, including group 2 coronaviruses, recognize 9-*O*-acetylated Sia-containing glycoconjugates on the surface of host cells, and some can remove the 9-*O*-acetyl moieties [5, 107, 108]. Recombinant soluble forms of the bovine coronavirus HE with and without inactivation of the esterase active site have been reported, and these molecules may be more stable than the InfCHE reagents. However,

Frozen sections probed with pre-complexed FcD

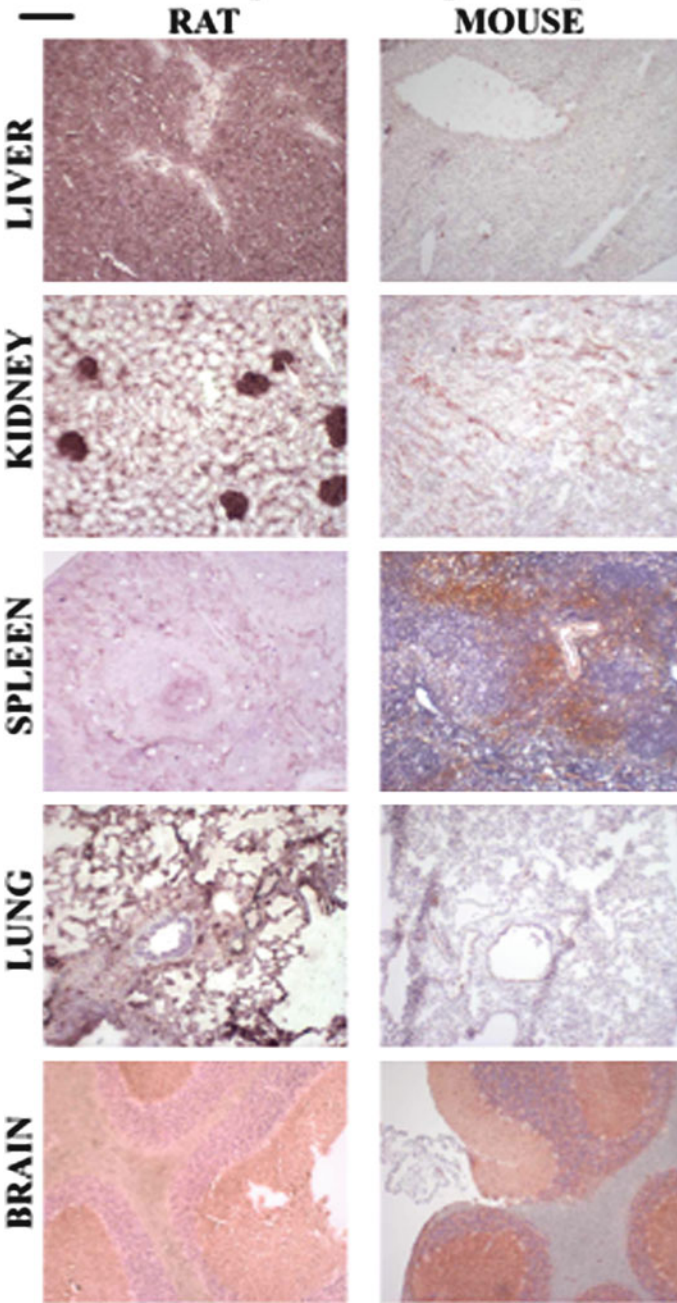


Fig. 7 Differences in binding seen with the probe for 9-O-acetylated sialic acids, CHE-FcD, to mouse and to rat sections. As shown, 9-O-acetylated sialic acids are abundantly expressed in rat liver but not in mouse liver, although it is present on mouse red blood cells (*right*). It is also expressed on rat kidney glomeruli but not in the mouse kidney where it is present only on some

Table 4 Expression of Neu5Gc in human tissues

All organs	Colon	Kidney	Epithelial carcinomas	Lymphomas, melanomas
Blood vessels	Luminal edge of epithelial cells	Glomeruli	About 30% of breast, ovary, prostate	Only on blood vessels

localization to tissue sections using immunohistochemistry has not yet been reported with the bovine coronaviruses reagents.

4.2 Detection of Neu5Gc in Human Tissues

Humans are genetically defective in synthesizing the common mammalian sialic acid Neu5Gc. It has been shown that Neu5Gc can be metabolically incorporated and covalently expressed on cultured human cell surfaces. Thus humans can also metabolically incorporate Neu5Gc into glycoproteins and glycolipids of human tumors, fetuses, and some normal tissues, likely from dietary sources (particularly red meats). Mass spectrometry confirmed the presence of small amounts of Neu5Gc in these human tissues.

As birds also do not synthesize Neu5Gc [109], it is possible to raise good anti-Neu5Gc antibodies in chickens. A polyclonal chicken anti-Neu5Gc was purified using a novel affinity method, utilizing sequential columns of immobilized human and chimpanzee serum sialo-glycoproteins, followed by specific elution from the latter column by free Neu5Gc. This purified anti-Neu5Gc antibody was used to characterize the expression of Neu5Gc in normal human and malignant tissues (Table 4).

Examination of frozen tissue sections showed expression in blood vessels, some epithelia, and in certain epithelial carcinomas [20, 110, 111]. Specificity of binding is confirmed by using a control IgY, and by inhibition using Neu5Gc rich chimpanzee serum. About 30% of breast, ovarian, and prostate carcinomas from humans showed expression of Neu5Gc (Fig. 8).

Other than expression within blood vessels within the malignant tissue, no expression was observed in melanomas and lymphomas using this three-step immunohistochemistry method. Paraffin sections showed a marked loss of expression, as did pre-incubation of frozen sections with methanol, indicating that the Neu5Gc epitopes are likely to be present mostly on glycolipids.



Fig. 7 (continued) medullary vessels. There is a different distribution in spleen from rat and mouse, and is present in rat lungs but not in mouse lungs; and diffusely in rat brain (*left*) and only in gray matter of mouse brain (*right*). Scale bar = 50 μ m

Frozen section immiunohistochemistry on carcinomas
anti-Neu5Gc **+10% chimp serum**

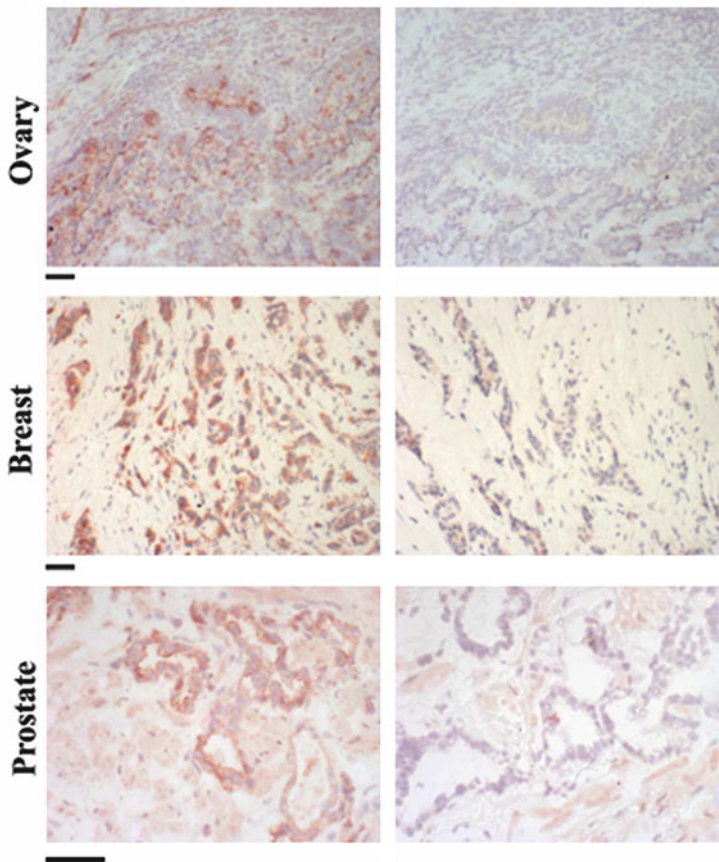


Fig. 8 Examples of results from immunohistochemistry on human malignant tissues, using a purified polyclonal chicken anti-Neu5Gc antibody. As shown, Neu5Gc is expressed on blood vessels and on malignant cells (*left side*) and the binding is blocked when using Neu5Gc rich chimpanzee serum. *Scale bar* = 50 μ m

5 Conclusion and Perspectives

Structural diversity of Sia is not so large in each animal compared with that which we know in nature. In humans, Neu5Ac is a dominant Sia species, with NeuGc and Kdn contributing to only small amounts. 9-*O*-Acetylated Neu5Ac is a major modified Sia, but is still a minor form of total Neu5Ac in human. In rainbow trout, on the other hand, Neu5Ac, Neu5Gc, and Kdn appear to be equally expressed, although in a tissue-specific manner. Therefore, which Sia species each animal expresses for particular biological roles depends on the animal species. Evolutional pressures for

the selection of Sia species may come from critical interactions for survival of life, such as pathogen–host interactions, and species-specific sperm–egg interactions in fertilization. Thus, in some cases, a particular Sia species is specifically functional; in other cases, no specific Sia is necessary. The presence of Kdn-specific sialidase in bacteria, the preferential recognition of Neu5Gc by mouse siglec-2, and an inhibitory effect of polySia, but not di- and oligoSia, on homophilic binding of NCAM may exemplify the significance of the structural diversity of Sia species. Importantly, modifications of Sia are often changed dynamically, including the *O*-acetylation/de-*O*-acetylation cycle and transient degradation of polySia to oligoSia. However, currently there is very little knowledge available about the enzymes involved in these modification/re-modification reactions. Future studies will be directed toward this identification of Sia-modification enzymes and their genes.

References

1. Angata T, Varki A (2002) Chemical diversity in the sialic acid and related α -keto acids: an evolutionary perspective. *Chem Rev* 102:439–469
2. Schauer R, Kamerling JP (2003) In: Montreuil J, Vlieghehart JFG, Schachter H (eds) *Chemistry, biochemistry and biology of sialic acid, glycoproteins II*. Elsevier, Amsterdam, pp 243–402
3. Varki A (1997) Sialic acids as ligands in recognition phenomena. *FASEB J* 11:248–255
4. Varki A, Hooshmand F, Diaz S, Varki NM, Hedrick SM (1991) Developmental abnormalities in transgenic mice expressing a sialic acid-specific 9-*O*-acetyltransferase. *Cell* 65:65–74
5. Cariappa A, Takematsu H, Liu H, Diaz S, Haider K, Boboila C, Kalloo G, Connole M, Shi HN, Varki N, Varki A, Pillai S (2009) B cell antigen receptor signal strength and peripheral B cell development are regulated by a 9-*O*-acetyl sialic acid esterase. *J Exp Med* 206:125–138
6. Parameswaran R, Lim M, Arutyunyan A, Abdel-Azim H, Hurtz C, Lau K, Mischen M, Yu RK, von Itzstein M, Heisterkamp N, Groffen J (2013) *O*-Acetylated *N*-acetylneuraminic acid as a novel target for therapy in human pre-B acute lymphoblastic leukemia. *J Exp Med* 210:805–819
7. Mandal C, Schwartz-Albiez R, Vlasak R (2011) Functions and biosynthesis of *O*-acetylated sialic acids. *Top Curr Chem*. doi:10.1007/128_2011_310
8. Davies LRL, Varki A (2013) Why is *N*-glycolylneuraminic acid rare in the vertebrate brain? *Top Curr Chem*. doi:10.1007/128_2013_419
9. Hildebrandt H, Dityatev A (2013) Polysialic acid in brain development and synaptic plasticity. *Top Curr Chem*. doi:10.1007/128_2013_446
10. Wipfler D, Srinivasan GV, Sadick H, Kniep B, Arming S, Willhauck-Fleckenstein M, Vlasak R, Schauer R, Schwartz-Albiez R (2011) Differentially regulated expression of 9-*O*-acetyl GD3 (CD60b) and 7-*O*-acetyl-GD3 (CD60c) during differentiation and maturation of human T and B lymphocytes. *Glycobiology* 21:1161–1172
11. Lrhorfi LA, Srinivasan GV, Schauer R (2007) Properties and partial purification of sialate *O*-acetyltransferase from bovine submandibular glands. *Biol Chem* 388:297–306
12. Higa HH, Diaz S, Varki A (1987) Biochemical and genetic evidence for distinct membrane-bound and cytosolic sialic acid *O*-acetyl-esterases: serine-active-site enzymes. *Biochem Biophys Res Commun* 144:1099–1108
13. Shen Y, Kohla G, Lrhorfi AL, Sipos B, Kalthoff H, Gerwing GJ, Kamerling JP, Schauer R, Tiralongo J (2004) *O*-Acetylation and de-*O*-acetylation of sialic acids in human colorectal carcinoma. *Eur J Biochem* 271:281–290

14. Schauer R, Shukla AK (2008) Isolation and properties of two sialate-*O*-acetyl esterases from horse liver with 4- and 9-*O*-acetyl specificities. *Glycoconj J* 25:625–632
15. Schwardt O, Kelm S, Ernst B (2013) SIGLEC4 antagonists. *Top Curr Chem*. doi:10.1007/128_2013_498
16. Chou HH, Takematsu H, Diaz S, Iber J, Nickerson E, Wright KL, Muchmore EA, Nelson DL, Warren ST, Varki A (1998) A mutation in human CMP-sialic acid hydroxylase occurred after the Homo-Pan divergence. *Proc Natl Acad Sci USA* 95:11751–11756
17. Irie A, Koyama S, Kozutsumi Y, Kawasaki T, Suzuki A (1998) The molecular basis for the absence of *N*-glycolylneuraminic acid in humans. *J Biol Chem* 273:15866–15871
18. Taylor RE, Gregg CJ, Padler-Karavani V, Ghaderi D, Yu H, Huang S, Sorensen RU, Chen X, Inostroza J, Nizet V, Varki A (2010) Novel mechanism for the generation of human xeno-autoantibodies against the nonhuman sialic acid *N*-glycolylneuraminic acid. *J Exp Med* 207:1637–1646
19. Hedlund M, Padler-Karavani V, Varki NM, Varki A (2008) Evidence for a human-specific mechanism for diet and antibody-mediated inflammation in carcinoma progression. *Proc Natl Acad Sci USA* 105:18936–18941
20. Pham T, Gregg CJ, Karp F, Chow R, Padler-Karavani V, Cao H, Chen X, Witztum JL, Varki NM, Varki A (2009) Evidence for a novel human-specific xeno-auto-antibody response against vascular endothelium. *Blood* 114:5225–5235
21. Shilova N, Huflejt ME, Vuskovic M, Obukhova P, Navakouski M, Khasbiullina N, Pazynina G, Galanina O, Bazhenov A, Bovin N (2013) Natural antibodies against sialoglycans. *Top Curr Chem*. doi:10.1007/128_2013_469
22. Nadano D, Iwasaki M, Endo S, Kitajima K, Inoue S, Inoue Y (1986) A naturally occurring deaminated neuraminic acid, 3-deoxy-*D*-glycero-*D*-galactono-ulosonic acid (KDN). Its unique occurrence at the nonreducing ends of oligosialyl chains in polysialoglycoprotein of rainbow trout eggs. *J Biol Chem* 261:11550–11557
23. Inoue S, Kitajima K (2006) KDN (deaminated neuraminic acid): dreamful past and exciting future of the newest member of the sialic acid family. *Glycoconj J* 23:277–290
24. Inoue S, Kitajima K, Inoue Y (1996) Identification of 2-keto-3-deoxy-*D*-glycero-*D*-galactono-ulosonic acid (KDN, deaminoneuraminic acid) residues in mammalian tissues and human lung carcinoma cells. Chemical evidence of the occurrence of KDN glycoconjugates in mammals. *J Biol Chem* 271:24341–24344
25. Go S, Sato C, Yin J, Kannagi R, Kitajima K (2007) Hypoxia-enhanced expression of free deaminoneuraminic acid in human cancer cells. *Biochem Biophys Res Commun* 357:537–542
26. Prokazova NV, Mikhailov AT, Kocharov SL, Malchenko LA, Zvezdina ND, Buznikov G, Bergelson LD (1981) Unusual gangliosides of eggs and embryos of the sea urchin *Strongylocentrotus intermedius*. Structure and density-dependence of surface localization. *Eur J Biochem* 115:671–677
27. Kochetkov NK, Smimova GP, Chekareva NV (1976) Isolation and structural studies of a sulfated sialosphingolipid from the sea urchin *Echinocardium cordatum*. *Biochim Biophys Acta* 424:274–283
28. Kubo H, Irie A, Inagaki F, Hoshi M (1990) Gangliosides from the eggs of the sea urchin, *Anthocidaris crassispina*. *J Biochem (Tokyo)* 108:185–192
29. Ijuin T, Kitajima K, Song Y, Kitazume S, Inoue S, Haslam SM, Morris HR, Dell A, Inoue Y (1996) Isolation and identification of novel sulfated and nonsulfated oligosialyl glycosphingolipids from sea urchin sperm. *Glycoconj J* 13:401–413
30. Kitazume S, Kitajima K, Inoue S, Haslam SM, Morris HR, Dell A, Lennarz WJ, Inoue Y (1996) The occurrence of novel 9-*O*-sulfated *N*-glycolylneuraminic acid-capped $\alpha 2 \rightarrow 5$ -*O*-glycolyl-linked oligo/polyNeu5Gc chains in sea urchin egg cell surface glycoprotein. Identification of a new chain termination signal for polysialyltransferase. *J Biol Chem* 271:6694–6701
31. Miyata S, Sato C, Kitamura S, Toriyama M, Kitajima K (2004) A major flagellum sialoglycoprotein in sea urchin sperm contains a novel polysialic acid, an $\alpha 2,9$ -linked poly-*N*-

- acetylneuraminic acid chain, capped by an 8-*O*-sulfated sialic acid residue. *Glycobiology* 14:827–840
32. Miyata S, Sato C, Kumita H, Toriyama M, Vacquier VD, Kitajima K (2006) Flagelliasialin: a novel sulfated α 2,9-linked polysialic acid glycoprotein of sea urchin sperm flagella. *Glycobiology* 16:1229–1241
 33. Slomiany BL, Kojima K, Banas-Gruszka Z, Murty VL, Galicki NI, Slomiany A (1981) Characterization of the sulfated monosialosyltriglycosylceramide from bovine gastric mucosa. *Eur J Biochem* 119:647–650
 34. Slomiany A, Kojima K, Banas-Gruszka Z, Slomiany BL (1981) Structure of a novel sulfated sialoglycosphingolipid from bovine gastric mucosa. *Biochem Biophys Res Commun* 100:778–784
 35. Morimoto N, Nakano M, Kinoshita M, Kawabata A, Morita M, Oda Y, Kuroda R, Kakehi K (2001) Specific distribution of sialic acids in animal tissues as examined by LC-ESI-MS after derivatization with 1,2-diamino-4,5-methylenedioxybenzene. *Anal Chem* 73:5422–5428
 36. Bulai T, Bratosin D, Pons A, Montreuil J, Zanetta JP (2003) Diversity of the human erythrocyte membrane sialic acids in relation with blood groups. *FEBS Lett* 534:185–189
 37. Kawai Y, Takemoto M, Oda Y, Kakehi K, Ohta Y, Yamaguchi S, Miyake M (2000) Inhibition of in vitro fertilization of mouse gametes by sulfated sialic acid polymers. *Biol Pharm Bull* 23:936–940
 38. Kimura K, Mori S, Tomita K, Ohno K, Takahashi K, Shigeta S, Terada M (2000) Antiviral activity of NMSO₃ against respiratory syncytial virus infection in vitro and in vivo. *Antiviral Res* 47:41–51
 39. Kisa F, Yamada K, Miyamoto T, Inagaki M, Higuchi R (2007) Determination of the absolute configuration of sialic acids in gangliosides from the sea cucumber *Cucumaria echinata*. *Chem Pharm Bull(Tokyo)* 55:1051–1052
 40. Bhattacharjee AK, Jennings HJ, Kenny CP, Martin A, Smith IC (1976) Structural determination of the polysaccharide antigens of *Neisseria meningitidis* serogroups Y, W-135, and BO1. *Can J Biochem* 54:1–8
 41. Sato C, Kitajima K (2011) New functions of polysialic acid and its relationship to schizophrenia. *Trends Glycosci Glycotechnol* 23:221–238
 42. Mühlenhoff M, Rollenhagen M, Werneburg S, Gerardy-Schahn R, Hildebrandt H (2013) Polysialic acid: versatile modification of NCAM, SynCAM 1 and Neuropilin-2. *Neurochem Res* doi:10.1007/s11064-013-0979-2
 43. Sato C (2013) Polysialic acid. In: Tiralongo J, Martinez-Duncker I (eds) *Sialobiology: structure, biosynthesis, and function*, Bentham Science, pp 33–75. <http://www.eurekaselect.com/107046/volume/1>
 44. Rutishauser U (2008) Polysialic acid in the plasticity of the developing and adult vertebrate nervous system. *Nat Rev Neurosci* 9:26–35
 45. Kanato Y, Kitajima K, Sato C (2008) Complex formation of a brain-derived neurotrophic factor and glycosaminoglycans. *Glycobiology* 18:1044–1053
 46. Kanato Y, Ono S, Kitajima K, Sato C (2009) Direct binding of polysialic acid to a brain-derived neurotrophic factor depends on the degree of polymerization. *Biotech Biosci Biochem* 73:2735–2741
 47. Sato C, Yamakawa N, Kitajima K (2010) Measurement of glycan-based interactions by frontal affinity chromatography and surface plasmon resonance. *Methods Enzymol* 478:219–232
 48. Isomura R, Kitajima K, Sato C (2011) Structural and functional impairments of polysialic acid by a mutated polysialyltransferase found in schizophrenia. *J Biol Chem* 286:21535–21545
 49. Ono S, Hane M, Kitajima K, Sato C (2012) Novel regulation of FGF2-mediated cell growth by polysialic acid. *J Biol Chem* 287:3710–3722
 50. Vaithianathan T, Matthias K, Bahr B, Schachner M, Suppiramaniam V, Dityatev A, Steinhäuser C (2004) Neural cell adhesion molecule-associated polysialic acid potentiates

- α -amino-3-hydroxy-5-methylisoxazole-4-propionic acid receptor currents. *J Biol Chem* 279:47975–47984
51. Hammond MS, Sims C, Parameshwaran K, Suppiramaniam V, Schachner M, Dityatev A (2006) Neural cell adhesion molecule-associated polysialic acid inhibits NR2B-containing *N*-methyl-*D*-aspartate receptors and prevents glutamate-induced cell death. *J Biol Chem* 281:34859–34869
 52. Kambara Y, Shiba K, Yoshida M, Sato C, Kitajima K, Shingyoji C (2011) Mechanism regulating Ca^{2+} -dependent mechanosensory behaviour in sea urchin spermatozoa. *Cell Struct Funct* 36:69–82
 53. Sato C, Kitajima K (1999) Glycobiology of di- and oligosialyl glycotopes. *Trends Glycosci Glycotechnol* 11:371–390
 54. Sato C (2004) Glycobiology of di- and oligosialyl glycotopes. *Trends Glycosci Glycotechnol* 16:331–344
 55. Nadanaka S, Sato C, Kitajima K, Katagiri K, Irie S, Yamagata T (2001) Occurrence of oligosialic acids on integrin α -5 subunit and their involvement in cell adhesion to fibronectin. *J Biol Chem* 276:33657–33664
 56. Sato C, Matsuda T, Kitajima K (2002) Neuronal differentiation-dependent expression of the disialic acid epitope on CD166 and its involvement in neurite formation in Neuro2A cells. *J Biol Chem* 277:45299–45305
 57. Yasukawa Z, Sato C, Sano K, Ogawa H, Kitajima K (2006) Identification of disialic acid-containing glycoproteins in mouse serum: a novel modification of immunoglobulin light chains, vitronectin, and plasminogen. *Glycobiology* 16:651–665
 58. Miyata S, Yamakawa N, Toriyama M, Sato C, Kitajima K (2011) Co-expression of two distinct polysialic acids, α 2,8- and α 2,9-linked polymers of *N*-acetylneuraminic acid, in distinct glycoproteins and glycolipids in sea urchin sperm. *Glycobiology* 21:1596–1605
 59. Aminoff D (1961) Methods for the quantitative estimation of *N*-acetylneuraminic acid and their application to hydrolysates of sialomucoids. *Biochem J* 81:384–392
 60. Uchida Y, Tsukada Y, Sugimori T (1977) Distribution of neuraminidase in *Arthrobacter* and its purification by affinity chromatography. *J Biochem (Tokyo)* 82:1425–1433
 61. Svennerholm L (1957) Quantitative estimation of sialic acids. II. A colorimetric resorcinol-hydrochloric acid method. *Biochim Biophys Acta* 24:604–611
 62. Hara S, Takemori Y, Yamaguchi M, Nakamura M, Ohkura Y (1987) Fluorometric high-performance liquid chromatography of *N*-acetyl- and *N*-glycolylneuraminic acids and its application to their microdetermination in human and animal sera, glycoproteins, and glycolipids. *Anal Biochem* 164:138–145
 63. Hara S, Yamaguchi M, Takemori Y, Furuhashi K, Ogura H, Nakamura M (1989) Determination of mono-*O*-acetylated *N*-acetylneuraminic acids in human and rat sera by fluorometric high-performance liquid chromatography. *Anal Biochem* 179:162–166
 64. Yamakawa N, Sato C, Miyata S, Maehashi E, Toriyama M, Sato N, Furuhashi K, Kitajima K (2007) Development of sensitive chemical and immunochemical methods for detecting sulfated sialic acids and their application to glycoconjugates from sea urchin sperm and eggs. *Biochimie* 89:1396–1408
 65. Kamerling JP, Schauer R, Shukla AK, Stoll S, Van Halbeek H, Vliegthart JF (1987) Migration of *O*-acetyl groups in *N*, *O*-acetylneuraminic acids. *Eur J Biochem* 162:601–607
 66. Finne J, Krusius T, Rauvala H (1977) Occurrence of disialosyl groups in glycoproteins. *Biochem Biophys Res Commun* 74:405–410
 67. Michon F, Brisson JR, Jennings HJ (1987) Conformational differences between linear α 2,8-linked homosialooligosaccharides and the epitope of the group B meningococcal polysaccharide. *Biochemistry* 26:8399–8405
 68. Sato C, Kitajima K, Tazawa I, Inoue Y, Inoue S, Troy FA II (1993) Structural diversity in the α 2-8-linked polysialic acid chains in salmonid fish egg glycoproteins. Occurrence of poly(Neu5Ac), poly(Neu5Gc), poly(Neu5Ac, Neu5Gc), poly(KDN), and their partially acetylated forms. *J Biol Chem* 268:23675–23684

69. Sato C, Inoue S, Matsuda T, Kitajima K (1998) Development of a highly sensitive chemical method for detecting α 2,8-linked oligo/polysialic acid residues in glycoproteins blotted on the membrane. *Anal Biochem* 261:191–197
70. Sato C, Inoue S, Matsuda T, Kitajima K (1999) Fluorescent-assisted detection of oligosialyl units in glycoconjugates. *Anal Biochem* 266:102–109
71. Zhang Y, Lee YC (1999) Acid-catalyzed lactonization of α 2,8-linked oligo/polysialic acids studied by high performance anion-exchange chromatography. *J Biol Chem* 274:6183–6189
72. Nakata D, Troy FA II (2005) Degree of polymerization (DP) of polysialic acid (polySia) on neural cell adhesion molecules (N-CAMS): development and application of a new strategy to accurately determine the DP of polySia chains on N-CAMS. *J Biol Chem* 275:38305–38316
73. Inoue S, Lin SL, Lee YC, Inoue Y (2001) An ultrasensitive chemical method for polysialic acid analysis. *Glycobiology* 11:759–767
74. Rohr TE, Troy FA (1980) Structure and biosynthesis of surface polymers containing polysialic acid in *Escherichia coli*. *J Biol Chem* 255:2332–2342
75. Kitajima K, Kuroyanagi H, Inoue S, Ye J, Troy FA 2nd, Inoue Y (1994) Discovery of a new type of sialidase, “KDNase,” which specifically hydrolyzes deaminoneuraminyl (3-deoxy-*D*-glycero-*D*-galacto-2-nonulosonic acid) but not *N*-acylneuraminyl linkages. *J Biol Chem* 269:21415–21419
76. Kanamori A, Inoue S, Xulei Z, Zuber C, Roth J, Kitajima K, Ye J, Troy FA 2nd, Inoue Y (1994) Monoclonal antibody specific for α 2,8-linked oligo deaminated neuraminic acid (KDN) sequences in glycoproteins. Preparation and characterization of a monoclonal antibody and its application in immunohistochemistry. *Histochemistry* 101:333–340
77. Song Y, Kitajima K, Inoue Y (1993) Monoclonal antibody specific to α 2,3-linked deaminated neuraminyl beta-galactosyl sequence. *Glycobiology* 3:31–36
78. Angata T, Matsuda T, Kitajima K (1998) Synthesis of neoglycoconjugates containing deaminated neuraminic acid (KDN) using rat liver α 2,6-sialyltransferase. *Glycobiology* 8:277–284
79. Ohta K, Sato C, Matsuda T, Toriyama M, Lennarz WJ, Kitajima K (1999) Isolation and characterization of low density detergent-insoluble membrane (LD-DIM) fraction from sea urchin sperm. *Biochem Biophys Res Commun* 258:616–623
80. Sarff LD, McCracken G, Schiffer MS, Glode MP, Robbins JB, Ørskov I, Ørskov F (1975) Epidemiology of *Escherichia coli* K1 in healthy and diseased newborns. *Lancet* 1975:1099–1104
81. Frosch M, Gorge I, Boulnois GJ, Timmis KN, Bitter-Suremann D (1985) NZB mouse system for production of monoclonal antibodies to weak bacterial antigens: isolation of an IgG antibody to the polysaccharide capsules of *Escherichia coli* K1 and group B meningococci. *Proc Natl Acad Sci USA* 82:1194–1198
82. Jennings HJ, Roy R, Michon F (1985) Determinant specificities of the groups B and C polysaccharides of *Neisseria meningitidis*. *J Immunol* 134:2651–2657
83. Häyrynen J, Bitter-Suermann D, Finne J (1989) Interaction of meningococcal group B monoclonal antibody and its Fab fragment with α 2-8-linked sialic acid polymers: requirement of a long oligosaccharide segment for binding. *Mol Immunol* 26:523–529
84. Sato C, Kitajima K, Inoue S, Seki T, Troy FA II, Inoue Y (1995) Characterization of the antigenic specificity of four different anti-(α 2,8-linked polysialic acid) antibodies using lipid-conjugated oligo/polysialic acids. *J Biol Chem* 270:8923–18928
85. Sato C, Fukuoka H, Ohta K, Matsuda T, Koshino R, Kobayashi K, Troy FA II, Kitajima K (2000) Frequent occurrence of pre-existing α 2→8-linked disialic and oligosialic acids with chain lengths up to 7 Sia residues in mammalian brain glycoproteins. Prevalence revealed by highly sensitive chemical methods and anti-di-, oligo-, and poly-Sia antibodies specific for defined chain lengths. *J Biol Chem* 275:15422–15431
86. Inoko E, Nishiura Y, Tanaka H, Takahashi T, Furukawa K, Kitajima K, Sato C (2010) Developmental stage-dependent expression of an α 2,8-trisialic acid unit on glycoproteins in mouse brain. *Glycobiology* 20:916–928

87. Yasukawa Z, Sato C, Kitajima K (2007) Identification of an inflammation-inducible serum protein recognized by anti-disialic acid antibodies as carbonic anhydrase II. *J Biochem* 141:429–441
88. Jakobsson E, Schwarzer D, Jokilampi A, Finne J (2012) Endosialidasas – versatile tools for the study of polysialic acid. *Top Curr Chem*. doi:10.1007/128_2012_349
89. Hallenbeck PC, Vimer ER, Yu F, Basseler B, Troy FA (1987) Purification and properties of a bacteriophage-induced endo-*N*-acetylneuraminidase specific for poly- α -2,8-sialosyl carbohydrate units. *J Biol Chem* 262:3553–3561
90. Pelkonen S, Perkonen J, Finne J (1989) Common cleavage pattern of polysialic acid by bacteriophage endosialidasas of different properties and origins. *J Virol* 63:4409–4416
91. Miyake K, Muraki T, Hattori K, Machida Y, Watanabe M, Kawase M, Yoshida Y, Iijima S (1997) Screening of bacteriophages producing endo-*N*-acetylneuraminidase. *J Ferm Bioeng* 84:90–93
92. Aalto J, Pelkonen S, Kalimo H, Finne J (2001) Mutant bacteriophage with non-catalytic endosialidase binds to both bacterial and eukaryotic polysialic acid and can be used as probe for its detection. *Glycoconj J* 18:751–758
93. Martin LT, Marth JD, Varki A, Varki NM (2002) Genetically altered mice with different sialyltransferase deficiencies show tissue-specific alterations in sialylation and sialic acid 9-*O*-acetylation. *J Biol Chem* 277:32930–32982
94. Yamamoto K, Konami Y, Irimura T (1997) Sialic acid-binding motif of *Maackia amurensis* lectins. *J Biochem* 121:756–761
95. Kim BS, Oh TK, Cho DH, Kim YJ, Koo WM, Kong KH, Kom H (2004) A sialic-acid binding lectin from the legume *Maackia fauriei*: comparison with lectin from *M. amurensis*. *Plant Sci* 167:1315–1321
96. Brinkmann-van der Linden ECM, Sonnenburg JL, Varki A (2002) Effects of sialic acid substitutions on recognition by *Sambucus nigra* agglutinin and *Maackia amurensis* Hemagglutinin. *Anal Biochem* 303:98–104
97. Nicholls JM, Bourne AJ, Chen H, Guan Y, Malik Peiris JS (2007) Sialic acid receptor detection in the human respiratory tract: evidence for widespread distribution of potential binding sites for human and avian influenza viruses. *Respir Res* 8:73–83
98. Imberty A, Gautier C, Lescar J, Perez S, Wyns L, Loris R (2000) An unusual carbohydrate binding site revealed by the structures of two *Maackia amurensis* lectins complexed with sialic acid-containing oligosaccharides. *J Biol Chem* 275:17541–17548
99. Geisler C, Jarvis DL (2011) Effective glycoanalysis with *Maackia amurensis* lectins requires a clear understanding of their binding specificities. *Glycobiology* 21:988–993
100. Herrler G, Rott R, Klenk HD, Muller HP, Shukla AK, Schauer R (1985) The receptor-destroying enzyme of influenza C virus is neuraminidase O-acetyltransferase. *EMBO J* 4:1503–1506
101. Vlasak R, Krystal M, Nacht M, Palese P (1987) The influenza C virus glycoprotein (HE) exhibits receptor-binding (hemagglutinin) and receptor-destroying (esterase) activities. *Virology* 160:419–425
102. Zimmer G, Reuter G, Schauer R (1992) Use of influenza C virus for detection of 9-*O*-acetylated sialic acids on immobilized glycoconjugates by esterase activity. *Eur J Biochem* 204:209–215
103. Klein A, Krishna M, Varki NM, Varki A (1994) 9-*O*-Acetylated sialic acids have widespread but selective expression: analysis using a chimeric dual-function probe derived from influenza C hemagglutinin-esterase. *Proc Natl Acad Sci USA* 91:7782–7786
104. Shi WX, Chammas R, Varki NM, Powell L, Varki A (1996) Sialic acid 9-*O*-acetylation on murine erythroleukemia cells affects complement activation, binding to I-type lectins, and tissue homing. *J Biol Chem* 271:31526–31532
105. Muchmore E, Varki A (1987) Selective inactivation of influenza C esterase: a probe for detecting 9-*O*-acetylated sialic acids. *Science* 236:1293–1295

106. Mehansho H, Carlson DM (1983) Induction of protein and glycoprotein synthesis in rat submandibular glands by isoproterenol. *J Biol Chem* 258:6616–6620
107. Schwegmann-Webels C, Herrier G (2006) Sialic acids as receptor determinants for coronaviruses. *Glycoconj J* 23:51–58
108. Langereis MA, van Vliet ALW, Boot W, deGroot RJ (2010) Attachment of mouse hepatitis virus to *O*-acetylated sialic acid is mediated by hemagglutinin-esterase and not by the spike protein. *J Virol* 2010:8970–8974
109. Schauer R (1982) Sialic acids: chemistry, metabolism and function, cell biology monographs, vol 10. Springer, Berlin
110. Tangvoranuntakul P, Gagneux P, Diaz S, Bardor M, Varki N, Varki A, Muchmore E (2003) Human uptake and incorporation of an immunogenic nonhuman dietary sialic acid. *Proc Natl Acad Sci USA* 100:12045–12050
111. Diaz SL, Padler-Karavani V, Ghaderi D, Hurtado-Ziola N, Yu H, Chen X, Brinkman-Van der Linden EC, Varki A, Varki NM (2009) Sensitive and specific detection of the non-human sialic acid *N*-glycolyneuraminic acid in human tissues and biotherapeutic products. *PLoS One* 4:e4241

Development and Applications of the Lectin Microarray

Jun Hirabayashi, Atsushi Kuno, and Hiroaki Tateno

Abstract The lectin microarray is an emerging technology for glycomics. It has already found maximum use in diverse fields of glycobiology by providing simple procedures for differential glycan profiling in a rapid and high-throughput manner. Since its first appearance in the literature in 2005, many application methods have been developed essentially on the same platform, comprising a series of glycan-binding proteins immobilized on an appropriate substrate such as a glass slide. Because the lectin microarray strategy does not require prior liberation of glycans from the core protein in glycoprotein analysis, it should encourage researchers not familiar with glycotecnology to use glycan analysis in future work. This feasibility should provide a broader range of experimental scientists with good opportunities to investigate novel aspects of glycoscience. Applications of the technology include not only basic sciences but also the growing fields of bio-industry. This chapter describes first the essence of glycan profiling and the basic fabrication of the lectin microarray for this purpose. In the latter part the focus is on diverse applications to both structural and functional glycomics, with emphasis on the wide applicability now available with this new technology. Finally, the importance of developing advanced lectin engineering is discussed.

Keywords Bio-affinity • Biomarker • Functional glycomics • Glycan profiling • Lectin microarray

J. Hirabayashi (✉), A. Kuno, and H. Tateno
Research Center for Medical Glycoscience, National Institute of Advanced Industrial Science and Technology (AIST), Central-2, 1-1-1, Umezono, Tsukuba, Ibaraki 305-8568, Japan
e-mail: jun-hirabayashi@aist.go.jp

Contents

1	Why Glycomics with Lectins?	106
1.1	Inherent Properties of Glycans	106
1.2	Lectin Microarray, an Advanced Technology for Glycan Profiling	107
1.3	Lectins or Carbohydrate-Binding Antibodies?	108
1.4	Preceding Techniques	109
2	Basic Fabrication of the Lectin Microarray System	110
2.1	Equipment	110
2.2	Microarray Plate: LecChip™	110
2.3	Detection Principles	111
3	General Procedures	112
3.1	Sample Preparation: Cell Lysate	112
3.2	Fluorescent Labeling and Application to the Lectin Microarray	113
3.3	Data Analysis	114
4	Targets of the Lectin Microarray	115
4.1	Samples of Homogeneous Glycome	115
4.2	Samples of Nearly Homogeneous Glycome	118
4.3	Samples of Highly Heterogeneous or Complex Glycomes	118
5	Modified Technologies of the Lectin Microarray	118
6	Perspective	119
	References	120

1 Why Glycomics with Lectins?

1.1 *Inherent Properties of Glycans*

Glycosylation endows a protein with various enhanced properties which the naked protein itself does not possess. Examples of these additional properties in protein glycosylation include polypeptide folding, stability, solubility, destination, functional regulation, and efficiency. In other words, glycosylation makes the role of a protein multidimensional [1, 2]. However, investigating protein glycosylation is not simple experimentally. Glycan structures are extremely diverse, for instance, in comparison with protein phosphorylation, another common post-translational modification. In general, glycans found in natural glycoproteins of higher animals consist of aldohexoses (e.g., Glc, Man, and Gal), their *N*-acetyl derivatives (e.g., GlcNAc and GalNAc), and, in many cases, deoxyhexose (e.g., L-fucose) and sialic acids represented by *N*-acetylneuraminic acid (NeuAc). An aldopentose, xylose, is also a common constituent of proteoglycan, and is involved in a linker region between core protein and glycosaminoglycan chains, such as heparin, heparan sulfate, dermatan sulfate, and chondroitin sulfate. However, the complexity of glycans is not strongly attributed to diversity in such component saccharides but more significantly to that in linkage isomers, because each hexose monosaccharide (e.g., Glc) possesses four potential donor hydroxyl groups (2-, 3-, 4-, and 6-OH) and one acceptor 1-hemiacetal (O–C–OH). The presence of anomerism (α/β) further doubles the number of linkage isomers. Indeed, there are eight possible isomers to

link two identical aldohexoses (e.g., Glc): “Glc α 1-2/3/4/6Glc” and “Glc β 1-2/3/4/6Glc.” The presence of multiple linkage isomers also allows branching in glycans, which is completely absent in nucleic acids and proteins. In theory, six monosaccharides can make as many as 1.05×10^{12} molecules of structural complexity, according to Laine [3]. Apparently, this figure far exceeds those calculated for hexanucleotides (i.e., $4^6 = 4,096$) and hexapeptides (i.e., $20^6 = 64,000,000$). Nobody knows the actual size of the glycome. However, it is considered that complete separation technology for glycans is not available. It also appears unrealistic currently to develop an automated method for glycan sequencing or synthesis [4, 5].

Because of the indirect association between genome and glycome, it is also difficult to predict glycan structures merely from the gene expression profiles: other information is necessary about the amounts of sugar nucleotide donors and location of a series of glycosyltransferases in the Golgi apparatus, some of which may compete with one another. However, the most important fact and the largest obstacle, which is never overcome by the classic concept workable for nucleic acids and proteins, is “heterogeneity” of glycans, partly because glycan biosynthesis is achieved only as a result of consecutive steps of glycosyltransferase and glycosidase reactions. These biosynthetic reactions can be incomplete or compete with one another. The extent of heterogeneity is further increased at the level of glycoconjugates (e.g., glycoproteins), because glycan structures and occupancies are not always the same between individual glycosylation sites. Of course, glycan structures should change if cell types and states that produce glycoproteins are different, even if the core protein structure is the same. This basic principle of glycoproteomics has been successfully applied to glycoprotein biomarker development for better diagnostic molecules [6, 7].

1.2 Lectin Microarray, an Advanced Technology for Glycan Profiling

Recent advances in separation and analytical technologies are noteworthy. They are represented by high-performance liquid chromatography (HPLC) and mass spectrometry (MS) as well as their combination with other separation principles (e.g., capillary electrophoresis, affinity chromatography). As described, however, it is still difficult for these methods to distinguish completely diverse structural isomers, and all of these conventional methods assume prior liberation of glycans and in most cases their separation and labeling. For the latter purpose, appropriate fluorescent tags (e.g., 2-aminopyridine and 2-aminobenzamide) are used because of increased sensitivity and resolution in HPLC. However, these processes require substantial labor and significant sacrifice of “throughput.” Although a commonly-used method for glycan liberation is an enzymatic cleavage of glycans from the core proteins, the procedure has substantial limitation to a class of asparagine-linked

N-glycans. On the other hand, no universal glycosidase has been available for serine/threonine-linked *O*-glycans. As a more fundamental issue, some important biological aspects, e.g., density, depth, and orientation, as well as combination to other biochemical components (e.g., core protein, with which glycan functions are closely associated), would be lost if glycans were to be liberated from the core protein. Mucins are a representative case of such glycoconjugates.

Most conventional methods for glycomics are based on physicochemical principles. However, as emphasized previously, they necessitate separation of “liberated” glycans. This procedure depends largely on the type of glycosylation (i.e., *N*-glycomics, *O*-glycomics, glycosaminoglycomics, glycolipidomics, etc.). Obviously, this is a disadvantage for glycomics if the intention is the comprehensive analysis of all types of glycans/glycoconjugates in a comparative manner. In contrast to these conventional technologies, an alternative is to use biochemical “probes” to glycans, which have a substantial affinity to glycans/glycoconjugates, such as lectins and carbohydrate-binding antibodies. Indeed, various techniques using these glycan probes have already been shown to work in biochemical and cytochemical studies, e.g., lectin-probed western blot analysis, histochemistry, and flow cytometry. While these methods are widely applicable to crude biological samples (e.g., cells and sera), they have a great disadvantage in throughput, sensitivity, and speed, all of which are critical requisites for glycomics. An advanced method, lectin microarray, was developed in 2005 to overcome these drawbacks [8–13]. Lectin microarray has a platform similar to that of antibody microarray, consisting of multiple probes (lectins and carbohydrate-binding antibodies) on an appropriate array substrate (typically a glass slide), which enables simultaneous interaction analyses in an ultrasensitive and high-throughput manner. However, a special attribute of the lectin microarray is differential glycan profiling rather than identification of particular target molecules. It is essentially difficult for the lectin microarray to identify individual glycan structures [4, 14, 15]. Differential glycan profiling can also target a mixture of glycoproteins and clinical samples [16, 17]. Readers can find the essential aspects of the lectin microarray in recent comprehensive reviews (e.g., [18–21]). Therefore, the following are descriptions with a focus on original aspects of the lectin microarray developed in the authors’ laboratory.

1.3 Lectins or Carbohydrate-Binding Antibodies?

According to the classic definition, lectins are carbohydrate-binding proteins, with the exception of antibodies and enzymes [22]. However, this criterion should be reconsidered in the light of current understanding of lectins: indeed, siglecs, a group of animal lectins specific for sialic acid, are members of the immunoglobulin superfamily, and an increasing number of enzymes have been found to contain carbohydrate-recognition domains (CRD) other than catalytic domains. On the other hand, it is now evident that even major lectin families, e.g., C-type lectins

and galectins, contain significant numbers of “non-lectin” members. In this context, it seems more reasonable to regard lectins simply as “a group of proteins that have significant ability to bind carbohydrates regardless of the protein families to which they belong.” The number of lectins identified is rapidly increasing as glycotecnologies to prove lectin activity have improved [23–25]. Currently, the number of established lectin families exceeds 40, and this number will probably increase significantly in the future. Moreover, it is possible to create artificial lectins not only from natural lectins that have sugar-binding activity but also from those that have no such activity. In other words, all the existing proteins are able to become “novel” lectins if they have potential binding pockets and if they are modified appropriately to make up a binding network to target carbohydrates, e.g., through hydrogen bonds and van der Waals contacts. Another requisite to evolve lectins might be multivalency in various terms; e.g., tandem repeat of CRDs or binding modules, subunit oligomerization, and acquisition of subsite specificity. Siglecs are an example of novel lectins, which have evolved from the immunoglobulin superfamily [26]. The opposite probably applies to galectins-10, 11, 13, and 15, because they have no evident β -galactose-binding activity despite their original definition [27].

As described, both antibodies and lectins are useful tools for glycan profiling. However, one should keep in mind that protein structures are significantly different between species, e.g., human and mouse, whereas many glycan structures are common between these species. For instance, high-mannose structures of *N*-linked glycans are strongly conserved in eukaryotes (e.g., mammals) and, thus, are not antigenic in any animal. On the other hand, representative xenoantigens, the α -Gal epitope and *N*-glycolylneuramic acid (NeuGc) are absent in humans and, thus, are antigenic. However, these xenoantigens are rather exceptional among a large number of common structures. This is a basic reason why only relatively poor anti-carbohydrate antibodies are produced in animals, most of which are attributed to IgM. However, for the sake of differential glycan profiling, rigorously specific antibodies do not work. Rather, probes such as conventional lectins with broader specificity are desirable to ensure the “coverage” of the glycome. Importantly, many of the conventional lectins show significant cross-affinity to structurally-related glycans, but to different extents. On the other hand, in order to detect specific structures (epitopes), carbohydrate-binding lectins or antibodies with rigorously defined specificity should be effective as epitope detectors [28]. The method is also expected to provide an extremely high-throughput means for glycan profiling.

1.4 Preceding Techniques

A previously-developed technique called “serial lectin-affinity chromatography” is available. Although the technique is almost identical to the lectin microarray in its essence, it lacks throughput and speed, because it utilizes open columns and

radio-labeled oligosaccharides [29, 30]. Frontal affinity chromatography (FAC) is a more quantitative method for the analysis of lectin-glycan interactions [31]. The method was recently greatly improved in speed, throughput, and accuracy [32, 33]. Nevertheless, the method is essentially applicable to only purified oligosaccharides. On the other hand, the lectin microarray has a more suitable platform providing simultaneous interaction analyses with a panel of carbohydrate-binding proteins. Moreover, attaining high sensitivity is assured by utilizing an established procedure for labeling proteins with common fluorescent reagents (e.g., Cy3, Cy5). Although special equipment is required for reproducible analysis, especially of monovalent glycans, in many cases lectin-glycoprotein interactions are strong enough to tolerate even repeated washing [8, 9, 11–13]. However, in the case of glycoproteins having only a few glycan chains, and thus relatively weak affinity to a lectin, i.e., in the order of 10^{-3} to 10^{-5} M in dissociation constant (K_d), high-wash procedures after a probing reaction should remove the weak binding. In this case, it is advisable to take advantage of an evanescent-field activated fluorescence detector, which enables in situ observation of immobilized lectin-glycan/glycoprotein interactions while preserving equilibrium [10, 34]. However, substantial affinity enhancement is often observed on an appropriate array platform, particularly when multiple lectin–glycan interactions occur. This phenomenon is generally understood as a glycoside cluster effect, first demonstrated by Y. C. Lee [35], and is closely associated with biological functions of lectins under physiological conditions [36].

2 Basic Fabrication of the Lectin Microarray System

2.1 Equipment

Figure 1 shows a general scheme for glycan profiling using the lectin microarray. Commercial products for this purpose are available [37]: the system consists of the scanner GlycoStation™ Reader 1200 and LecChip™, both of which are produced by Glycoptechnica, Co. Ltd. (Yokohama). An advantage is that the system utilizes an evanescent-field activated fluorescence detection principle (described later in greater detail).

2.2 Microarray Plate: LecChip™

LecChip™ Ver.1.0 is a commercial product for the lectin microarray: it contains 45 different lectins (for a list see [5] or [37]), which were carefully selected from a pool of 167 candidate lectins, taking into consideration their binding specificity, stability, and economy. Lectins are briefly classified based on monosaccharide

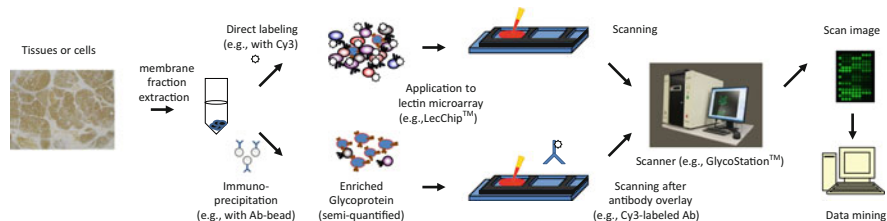


Fig. 1 General procedures of lectin microarray analysis: from sample preparation to data analysis. There are basically two schemes for glycan profiling; i.e., an *upper scheme* for direct profiling and a *lower scheme* for indirect profiling. In the former, all proteins prepared from cells and tissues are fluorescently labeled, e.g., Cy3, and their glycan profiles analyzed directly, while in the latter, target glycoproteins are immuno-precipitated prior to application to the microarray. The target glycoproteins are specifically detected by overlaying fluorescently-labeled detecting antibodies raised against the core protein moiety of the targets (for details, see text or Kuno et al. [28])

specificity; e.g., fucose binders (AAL, etc.), sialic acid binders (SSA, etc.), galactose binders (RCA120, etc.), mannose binders (ConA, etc.), and *O*-glycan binders (Jacalin, etc.), while some of them discriminate branching or more complex features (e.g., PHA-L). LecChip™ (ver.1.0) contains 7 wells on a glass slide, each containing the same set of 45 lectins spotted in triplicate. Basic information about these lectins is available on the website “Lectin frontier DataBase” (LfDB; <http://riodb.ibase.aist.go.jp/rcmg/glycodb/LectinSearch>), in which specificity data obtained by FAC [32, 33] are also available.

2.3 Detection Principles

1. Confocal fluorescence detection method

Many researchers using the lectin microarray take advantage of a confocal fluorescence detection principle, because the method has already been adapted for established systems of DNA microarray [18, 19]. The method is simple, and has therefore been widely used. However, it requires washing of the microarray surface after the binding reaction, because a dried sample is required for the confocal-type detection system. One useful application of the system is multi-color detection using, e.g., Cy3 and Cy5, as described later.

2. Evanescent-field activated fluorescence detection method

Kuno et al. developed a unique lectin microarray system, which utilizes an evanescent-field activated fluorescent detection principle. The evanescent wave is generated from the surface of the glass slide when an appropriate light originally introduced from both edges of the slide makes a total reflection [10]. Because the range of this wave is largely restricted to “near optic field (i.e., substantially <200 nm),” fluorescently labeled materials, such as Cy3-labeled glycoproteins, cannot be activated unless they form a significant complex on the array surface. As a result, the method requires no washing

procedures, which is not the case with a conventional confocal detection system. This specific feature enables in situ aqueous-phase detection without perturbing equilibrium conditions, which should contribute to both sensitive and reproducible analysis.

3. *Bimolecular fluorescence quenching and recovery method*

Hamachi and coworkers manufactured a sophisticated detection principle called “bimolecular fluorescence quenching and recovery” (BFQR) [38]. They used a supramolecular hydrogel matrix, where fluorescent lectins were noncovalently fixed to act as a molecular recognition scaffold. Though the detection mechanism is rather complicated and not straightforward, it enables one to read out fluorescently a series of lectin–saccharide interactions without washing processes, in a manner similar to the evanescent-type detection principle. Moreover, the developed method BFQR does not require prior labeling of the target saccharides. The latter feature resembles that of a surface plasmon resonance detection principle used much more widely [39]. To the present, however, relatively simple, synthetic saccharide quenchers have been utilized, limiting the application of this technology to more complex samples, e.g., body fluids and tissue extracts.

3 General Procedures

A general procedure for glycan profiling experiments using the GlycoStation™ system has been described [37]. Although there might be a variety of sample forms depending on experimental purposes, e.g., sera, urine, cell lysates, and tissue extracts, the core experimental procedure is common to all of the sample forms: i.e., it comprises (1) sample preparation, (2) protein quantification, (3) labeling with an appropriate fluorescent reagent (e.g., Cy-3), (4) application of the labeled sample onto the lectin microarray plate, and (5) data analysis, essentially as described [40, 41]. General procedures are outlined in Fig. 1.

In the following section, a typical procedure using cell lysates and Cy-3 labeling is described, while variously modified procedures are also available, some of which are described later.

3.1 *Sample Preparation: Cell Lysate*

The following protocol assumes that a cell pellet is washed extensively with PBS, and is kept frozen at -80°C until use:

1. Melt the cell pellet gradually on ice
2. Add 1 mL of PBS-Tx (PBS containing 1% (w/v) Triton X-100) to the cell pellet and suspend the cells with a pipette

3. Disrupt the cells by sonication (1–2 min on ice)
4. Centrifuge the above sample at $14,000\times g$ at 4°C for 5 min
5. Recover the supernatant with a pipette
6. Quantify protein by an appropriate method (e.g., Pierce/Micro BCA™ Protein Assay Reagent Kit, #23235)

3.2 Fluorescent Labeling and Application to the Lectin Microarray

1. Cy3 labeling

- (a) Dilute samples to $50\ \mu\text{g}/\text{mL}$ by adding PBS-Tx based on the result of the above protein assay
- (b) Put $20\ \mu\text{L}$ of each sample ($50\ \mu\text{g}/\text{mL}$) into a tube containing $100\ \mu\text{g}$ of Cy3-SE (succinimide ester), mix with a pipette, and spin down by brief centrifugation
- (c) Put the tubes into a shading bag and incubate them for 1 h at room temperature (25°C) in the dark

2. Gel filtration to remove excess free Cy3-SE:

Remove excess Cy3-SE reagent by gel filtration using appropriate equipment (e.g., Pierce/Zeba™ Desalt Spin Columns, $0.5\ \text{mL}$, #89882)

3. Sample application to LecChip™

- (a) Measure each volume of Cy3-labeled samples with a pipette
- (b) Prepare a total volume of $500\ \mu\text{L}$ by adding an appropriate probing buffer (e.g., GP Biosciences Probing Solution); the protein concentration becomes $2\ \mu\text{g}/\text{mL}$ as $1\ \mu\text{g}$ of protein is eluted
- (c) Dilute these samples to $2\ \mu\text{g}/\text{mL}$, $1\ \mu\text{g}/\text{mL}$, $500\ \text{ng}/\text{mL}$, $250\ \text{ng}/\text{mL}$, $125\ \text{ng}/\text{mL}$, $62.5\ \text{ng}/\text{mL}$, and $31.25\ \text{ng}/\text{mL}$ by adding the same probing buffer
- (d) Put LecChip™ in a frozen state at -20°C into an incubation box to melt the keeping solution (provided by the supplier) in wells of the LecChip™
- (e) Remove the keeping solution
- (f) Apply $100\ \mu\text{L}$ of these samples to each well of the LecChips™ with a pipette
- (g) Incubate the LecChips™ in an incubation box at 20°C for 16 h

4. Scanning with GlycoStation™ Reader 1200

Scan the LecChips with GlycoStation™ Reader 1200. The following conditions are recommended for the first trial: gain (70–125), exposure time (133, 199 ms), cumulative count (4). In order to detect relatively weak signals while avoiding saturation of strong signals, take some other scans while adjusting the gain and the exposure time.

3.3 Data Analysis

Data analysis is a critical part of lectin microarray experiments, because the data obtained from each microarray analysis shows systematic variation in microarray quality, scanner detection stability, sample preparation reproducibility, and labeling efficiency. For this purpose, an initial step to normalize the obtained signal intensities is essential. In this section, the two fundamental procedures required for data analysis are described [42–44].

1. Gain merging

Kuno et al. developed a gain-merging procedure to expand the dynamic range of signal intensities obtained by lectin microarray analysis [42]. This procedure is necessary because of the optical properties of the scanning system, but is practically useful when analyzing a series of clinical samples, which often give a wide range of binding signals. When performing an analysis, a lectin microarray slide is scanned under two different gain conditions; one is a higher gain to “rescue” weak signals below 1,000 and the other is a lower gain to “suppress” strong signals over 40,000. Here, appropriate lectins for “merging” are selected, to give moderate signal intensities (i.e., 1,000–40,000) under both higher and lower gain conditions. Then, a merging factor (F) is determined as the average of higher (Int^{H})/lower (Int^{L}) ratios calculated for individual merging lectins by (1):

$$F = \text{Averaged} (\text{Int}^{\text{H}}/\text{Int}^{\text{L}}) \quad (1)$$

The over-range intensities ($>40,000$) obtained under the higher gain condition (e.g., $\text{Int}^{\text{H}}_{(\text{lectin A})}$) are replaced with theoretical intensities ($\text{Int}^{\text{T}}_{\text{d lectin cT}}$) by (2):

$$\text{Int}^{\text{T}}_{(\text{lectin A})} = \text{Int}^{\text{L}}_{(\text{lectin A})} \times F \quad (2)$$

2. Normalization

Four different normalization methods are available to process lectin microarray data: “max,” “mean,” “particular lectin,” and “median.” For these normalization procedures, the signal intensity is multiplied by a normalization factor N for each array, which is calculated by $N = 1/\mu$, where μ is either the highest signal intensity of all of the lectins on the array (max), the mean of all of the lectins on the array (mean), the signal intensity of one selected lectin on the array (particular lectin), or the median of all of the lectins on the array (median). Required procedures depend significantly on experimental procedures and research purposes. For example, the author usually uses a mean normalization method for comparative purposes when dealing with a series of stem cells [45–47], because this gave the best result for glycan analysis of CHO and its mutant LEC cells [44]. On the other hand, a max-normalization procedure is widely used for differential glycan profiling targeting clinical samples for glycan-related bio-marker investigation [16, 17, 28, 48, 49].

4 Targets of the Lectin Microarray

Since its first publication in 2005, most reports of the lectin microarray have concerned its basic aspects; e.g., development of the array substrates and their fabrications. Recently, however, an increasing number of applications of lectin microarray technology have been reported, e.g., in glycan-related biomarker investigations, stem cell profiling toward regeneration medicine, microbial infections, and glycoprotein profiling in the light of functional glycomics (for representative applications, see Table 1). In the following section, relevant technologies required in various orders of hierarchy (i.e., molecular, cellular, tissue and body fluid orders) are briefly described.

4.1 Samples of Homogeneous Glycome

1. *Oligosaccharides*

There are few reports describing the application of the lectin microarray to free oligosaccharides [10, 34], probably because a substantial merit of this microarray is its direct applicability to glycoproteins without liberation of oligosaccharides. Alternatively, most monoamine-coupling fluorescent reagents (e.g., 2-aminopyridine and 2-aminobenzamide) are of the UV-excited type, which are not compatible with the present glass substrate (UV-fluorescence positive). Uchiyama et al. used TAMRA (tetramethylrhodamine)-labeled three representative *N*-linked oligosaccharides (i.e., M6 high-mannose type, α -2-6-disialobiantennary, and asialobiantennary glycans) for analysis by an evanescent-field activated fluorescence-type scanner [34]. Notably, without a washing procedure, relatively weak bindings of these monovalent oligosaccharides could be seen toward a restricted set of lectins, while they were lost immediately after a buffer replacement procedure. Signal patterns observed for these oligosaccharides are relatively clear, as has been evident when sequential glycosidase digestion was performed for a complex-type sialobiantennary *N*-glycan [10].

2. *Purified glycoproteins*

As described, glycoproteins are major targets for direct analysis by the lectin microarray, either in their purified (e.g., glycoprotein drugs) or crude (e.g., cell supernatant and body fluids) forms. The analysis provides glycan profiles regarding both *N*- and *O*-glycans [8–13]. Because the profiles obtained are unique to individual glycoproteins and states of the cells which produce them, the method contributes to the validation of glycoprotein drugs [12]. In combination with a specific antibody against core protein, highly sensitive monitoring of glycan profiles of target glycoproteins in the course of their production is possible (described later).

Table 1 Representative research papers reporting applications of the lectin microarray

Application fields and paper titles (year)	Reference
1. Glycan synthesis and glycoprotein production	
<i>Pichia pastoris</i> -produced mucin-type fusion proteins with multivalent <i>O</i> -glycan substitution as targeting molecules for mannose-specific receptors of the immune system (2011)	Gustafsson et al. [50]
Chemoenzymatic synthesis and lectin array characterization of a class of <i>N</i> -glycan clusters (2009)	Huang et al. [51]
Engineering of mucin-type human glycoproteins in yeast cells (2008)	Amano et al. [52]
A lectin array-based methodology for the analysis of protein glycosylation (2007)	Rosenfeld et al. [12]
2. Glycoprotein profiling relevant to functional glycomics	
Survey of glycoantigens in cells from α -1,3-galactosyltransferase knockout pig using a lectin microarray (2010)	Miyagawa et al. [53]
Transient expression of an IL-23R extracellular domain Fc fusion protein in CHO vs HEK cells results in improved plasma exposure (2010)	Suen et al. [54]
Glycomic analyses of glycoproteins in bile and serum during rat hepatocarcinogenesis (2010)	Nakagawa et al. [55]
Testicular angiotensin-converting enzyme with different glycan modification: characterization on glycosylphosphatidylinositol-anchored protein releasing and dipeptidase activities (2009)	Kondoh et al. [56]
Polylectosamine on glycoproteins influences basal levels of lymphocyte and macrophage activation (2007)	Togayachi et al. [57]
Functional glycosylation of human podoplanin: glycan structure of platelet aggregation-inducing factor (2007)	Kaneko et al. [58]
3. Development of disease-related glycoprotein markers	
A unique <i>N</i> -glycan on human transferrin in CSF: a possible biomarker for iNPH (2011)	Futakawa et al. [59]
Multilectin assay for detecting fibrosis-specific glyco-alteration by means of the lectin microarray (2011)	Kuno et al. [16]
Lectin microarray profiling of metastatic breast cancers (2011)	Fry et al. [60]
Identification of various types of α 2-HS glycoprotein in sera of patients with pancreatic cancer: possible implication in resistance to protease treatment (2010)	Kuwamoto et al. [61]
<i>Wisteria floribunda</i> agglutinin-positive mucin 1 is a sensitive biliary marker for human cholangiocarcinoma (2010)	Matsuda et al. [48]
High levels of E4-PHA-reactive oligosaccharides: potential as marker for cells with characteristics of hepatic progenitor cells (2009)	Sasaki et al. [62]
4. Glycan profiling of stem cells relevant to regenerative medicine	
Possible linkages between the inner and outer cellular states of human induced pluripotent stem cells (2011)	Saito et al. [47]
Glycome diagnosis of human induced pluripotent stem cells using the lectin microarray (2011)	Tateno et al. [46]
Lectin microarray analysis of pluripotent and multipotent stem cells (2011)	Toyoda et al. [45]
5. Glycan profiling relevant to pathogen infection	
HIV-1 and microvesicles from T cells share a common glycome, arguing for a common origin (2009)	Krishnamoorthy et al. [63]
Analyzing the dynamic bacterial glycome with a lectin microarray approach (2006)	Hsu et al. [64]

3. *Eukaryotic cells*

As mentioned previously, glycan structures differ significantly between different types of cells (species and states). According to this concept, i.e., “cellular glycomics,” the lectin microarray should work ideally for differentiation of different types of cells. The first demonstration was made by Ebe et al. [65] using CHO and its LEC mutant strains, and work in this context was further extended in a more systematic manner by Tao et al. [13]. These studies used detergent-solubilized cell membrane fractions as a glycoprotein source. However, significant care is necessary for the preparation of such glycoprotein fractions, which may contain interfering materials, such as non-glycosylated cytoplasmic proteins and immature glycoproteins included in the endoplasmic reticulum and Golgi apparatus.

Tateno et al. reported a convenient method to profile a cell-surface glycome using a commercial CellTracker dye (Invitrogen) [66]. This series of reagents passes freely through the cell membrane, but once inside the cell, they are transformed into cell-impermeant forms. One of the commercial products, Orange™ CMRA, can fluoresce only after it is metabolized by endogenous esterase. The procedure is simple and applicable to an extensive range of eukaryotic cells [44].

4. *Bacterial cells*

Hsu et al. [64] successfully profiled glycosylation patterns of closely related *Escherichia coli* strains, including pathogenic ones. Those authors differentiated commensal and pathogenic strains in a manner of facile fingerprinting. More recently, Yasuda et al. reported an alternative method to analyze bacterial glycomes with emphasis on differential profiling of 16 strains of *Lactobacilli casei* species [67]. They found CYTOX Orange, usually used for intracellular nuclear staining, to be the best dye to be incorporated into the bacterial cells and bound to DNA. Despite the organisms being from the same species, almost all binding patterns obtained for these 16 strains were unique. It should also be noted that the current systems of the lectin microarray, largely composed of plant lectins, can work for bacterial glycome profiling.

5. *Virus and related particles*

Krishnamoorthy et al. reported an interesting analysis of glycan profiles of the HIV virus which focused on comparison with virus-resembling microvesicles, which are secreted from host cells [63]. Lectin microarray analysis using 68 lectin probes revealed that glycome signatures of HIV and host cell microvesicles were almost identical. This observation provides important support for the “exosome” hypothesis that HIV largely relies on the biosynthetic machineries of host cells, and thus they should give totally the same glycan profiles. For this analysis, the authors utilized a sensitive ratiometric two-color detection method (described later).

4.2 Samples of Nearly Homogeneous Glycome

1. Tissue sections

Matsuda et al. developed a skilful technique for differential glycan profiling targeting small areas (i.e. 1.5 mm diameter, 5 μm thickness) of paraffin-embedded and formalin-fixed tissue sections [48]. The method provides clinicians with a very useful approach to differential glycan analysis. Considering that tissue samples to be compared are of the same date and same individual, the data obtained are more reliable than those obtained from sera. In fact, the authors applied the technique to an investigation of glycoprotein markers for diagnosis of cholangiocarcinoma, a representative malignant tumor, for which no useful markers are presently available. As a result, they found that *Wisteria floribunda* agglutinin specific for GalNAc was the best probe to differentiate cholangiocarcinoma lesions from normal bile duct epithelia ($p < 0.0001$) [49]. Though not yet performed, combination with laser micro-dissection followed by high-throughput differential glycan profiling should become a powerful means for future biomarker discovery.

4.3 Samples of Highly Heterogeneous or Complex Glycomes

1. Body fluids (bile, sera) in biomarker investigations

Body fluids, such as patient sera and urine, are primary targets for biomarker diagnosis. Under the concept of glycoproteomics, an increasing number of researchers of both glycomics and proteomics have been involved in the discovery phase of glycoprotein markers. For this realization, however, one should consider the fact that serum is a highly heterogeneous mixture of glycoproteins originating from many different organs. Therefore, their individual cellular glycomes are also heterogeneous. A well-organized strategy is described, which relies on the high-sensitivity technique of the lectin microarray at two different phases [6, 7]. One is the tissue-section targeted analysis described above, and the other an antibody-overlay method described below.

5 Modified Technologies of the Lectin Microarray

1. Antibody-overlay method in biomarker qualification/verification

Kuno et al. reported a highly practical approach to the differential glycan profiling of an antibody-targeted glycoprotein in the course of biomarker development [28]. As the discovery phase of biomarker development proceeds, several biomarker candidates (i.e., glycoproteins) are nominated, whereas it is not yet certain that they really work as robust diagnostic markers, e.g., for hepatocellular carcinoma. For the purpose of pre-validation (i.e., qualification

or verification), it is important to carry out the analysis on several dozen clinical samples (e.g., usually body fluids such as sera). Therefore, establishment of a high-throughput procedure such as the lectin microarray is particularly important in glycoprotein biomarker development. The developed method makes maximum use of an antibody, which is raised against the protein moiety of the target glycoprotein: the antibody is used for enrichment (immunoprecipitation), semi-quantification (Western blotting), and overlay to the lectin microarray (*lower* scheme in Fig. 1). A target glycoprotein included in each clinical sample need not be fluorescence-labeled, and thus high-throughput (>100 samples) analysis is easily performed. Therefore, the lectin microarray is used in two steps in the strategy developed for glycoprotein biomarker investigation. The approach was applied to the investigation of an hepatic fibrosis marker, α -1-acid glycoprotein probed with *Maackia amurensis* lectin (MAL) and *Aspergillus oryzae* lectin (AOL) [16].

2. Dual color measurement

Pilobello et al. developed a method for the rapid evaluation of glycosylation changes of heterogeneous mammalian samples using a ratiometric two-color lectin microarray approach [68]. This approach is reminiscent of a proteomic procedure called two-dimensional differential in-gel electrophoresis (2-D DIGE), which uses two differentially labeled (i.e., Cy3/Cy5) protein samples [69]. When focusing on differential analysis, this approach can be extremely useful to enhance substantial differences in the glycome. However, it should be mentioned that these two methods, i.e., proteomic 2-D DIGE and the glycomic dual color method, are essentially different in that the latter procedure includes “competition” between immobilized lectins toward a set of various glycans. In other words, quantitative comparison of lectin signals needs careful consideration.

6 Perspective

The technology of the lectin microarray described here is expanding its application fields rapidly to broad areas of life sciences, which include both basic and applied sciences. Because the method is still new, relatively few researchers appreciate its innovative features, which previous technologies have lacked, i.e., discrimination based on biological affinity. However, considering that every biological phenomenon comprises cellular communications, it is quite natural to assume that carbohydrate-protein (e.g., lectins) interactions are fundamental for carcinogenesis, embryogenesis, morphogenesis, pathogenesis, etc. Thus it should not be surprising that the lectin microarray platform works very well to analyze, differentiate, and elucidate these complex cellular processes under the basic concept of cellular glycomics/glycoproteomics. For further development of the system, however, production of a series of recombinant lectins is necessary, considering the history of restriction enzymes [18]. In this context, several groups have already shifted to

using recombinant lectins for improved sensitivity [70] and resolution [46]. If a recombinant lectin is established with a good bacterial expression system, its propagation to produce new lectins with novel specificities, a wider glycome coverage and a better cost effectiveness is feasible [71–73]. An alternative approach is to synthesize “artificial” lectins equipped with boronic acid functionalized peptidyl ligands to carbohydrates [74]. In this regard, the lectin microarray enters its really creative phase, and is poised for further development.

References

1. Varki A (1993) Biological roles of oligosaccharides: all of the theories are correct. *Glycobiology* 3:97–130
2. Varki A, Lowe JB (2009) Biological roles of glycans. In: Varki A, Cummings RD, Esko JD, Freeze HH, Stanley P, Bertozzi CR, Hart GW, Etzler ME (eds) *Essentials of glycobiology*, 2nd edn. Cold Spring Harbor, New York
3. Laine RA (1994) A calculation of all possible oligosaccharide isomers both branched and linear yields 1.05×10^{12} structures for a reducing hexasaccharide: the isomer barrier to development of single-method saccharide sequencing or synthesis systems. *Glycobiology* 4:759–767
4. Hirabayashi J (2008) Concept, strategy and realization of lectin-based glycan profiling. *J Biochem* 144:139–147
5. Hirabayashi J, Kuno A, Tateno H (2011) Lectin-based structural glycomics: a practical approach to complex glycans. *Electrophoresis* 32:1118–1128
6. Ito H, Kuno A, Sawaki H, Sogabe M, Ozaki H, Tanaka Y, Mizokami M, Shoda JI, Angata T, Sato T, Hirabayashi J, Ikehara Y, Narimatsu H (2009) Strategy for glycoproteomics: identification of glyco-alteration using multiple glycan profiling tools. *J Proteome Res* 8:1358–1367
7. Narimatsu H, Narimatsu H, Sawaki H, Kuno A, Kaji H, Ito H, Ikehara Y (2010) A strategy for discovery of cancer glyco-biomarkers in serum using newly developed technologies for glycoproteomics. *FEBS J* 277:95–105
8. Angeloni S, Ridet JL, Kusy N, Gao H, Crevoisier F, Guinchard S, Kochhar S, Sigrist H, Sprenger N (2005) Glycoprofiling with micro-arrays of glycoconjugates and lectins. *Glycobiology* 15:31–41
9. Pilobello KT, Krishnamoorthy L, Slawek D, Mahal LK (2005) Development of a lectin microarray for the rapid analysis of protein glycopatterns. *ChemBiochem* 6:985–989
10. Kuno A, Uchiyama N, Koseki-Kuno S, Ebe Y, Takashima S, Yamada M, Hirabayashi J (2005) Evanescent-field fluorescence-assisted lectin microarray: a new strategy for glycan profiling. *Nat Methods* 2:851–856
11. Zheng T, Peelen D, Smith LM (2005) Lectin arrays for profiling cell surface carbohydrate expression. *J Am Chem Soc* 127:9982–9983
12. Rosenfeld R, Bangio H, Gerwig GJ, Rosenberg R, Aloni R, Cohen Y, Amor Y, Plaschkes I, Kamerling JP, Maya RB (2007) A lectin array-based methodology for the analysis of protein glycosylation. *J Biochem Biophys Methods* 70:415–426
13. Tao SC, Li Y, Zhou J, Qian J, Schnaar RL, Zhang Y, Goldstein IJ, Zhu H, Schneck JP (2008) Lectin microarray: a powerful tool for glycan related biomarker discovery. *Glycobiology* 18:761–769
14. Hirabayashi J, Arata Y, Kasai K (2001) Glycome project: concept, strategy and preliminary application to *Caenorhabditis elegans*. *Proteomics* 1:295–303
15. Hirabayashi J (2004) Lectin-based structural glycomics: glycoproteomics and glycan profiling. *Glycoconj J* 21:35–40

16. Kuno A, Ikehara Y, Tanaka Y, Angata T, Unno S, Sogabe M, Ozaki H, Ito K, Hirabayashi J, Mizokami M, Narimatsu H (2011) Multilectin assay for detecting fibrosis-specific glycoalteration by means of lectin microarray. *Clin Chem* 57:48–56
17. Kuno A, Ikehara Y, Tanaka Y, Saito K, Ito K, Tsuruno C, Nagai S, Takahama Y, Mizokami M, Hirabayashi J, Narimatsu H (2011) LecT-Hepa: a triplex lectin-antibody sandwich immunoassay for estimating the progression dynamics of liver fibrosis assisted by a bedside clinical chemistry analyzer and an automated pretreatment machine. *Clin Chim Acta* 412:1767–1772
18. Gupta G, Surolia A, Sampathkumar SG (2010) Lectin microarrays for glycomic analysis. *OMICS* 14:419–436
19. Katrlík J, Svitel J, Gemeiner P, Kozár T, Tkac J (2010) Glycan and lectin microarrays for glycomics and medicinal applications. *Med Res Rev* 30:394–418
20. Vanderschaeghe D, Festjens N, Delanghe J, Callewaert N (2010) Glycome profiling using modern glycomics technology: technical aspects and applications. *Biol Chem* 391:149–161
21. Rakus JF, Mahal LK (2011) New technologies for glycomic analysis: toward a systematic understanding of the glycome. *Annu Rev Anal Chem (Palo Alto Calif)* 4:367–392
22. Goldstein IJ, Hughes RC, Monsigny M, Osawa T, Sharon N (1980) What should be called a lectin? *Nature* 285:66
23. Fukui S, Feizi T, Galustian C, Lawson AM, Chai W (2002) *Nat Biotechnol* 20:1011–1107
24. Blixt O, Head S, Mondala T, Scanlan C, Huflejt ME, Alvarez R, Bryan MC, Fazio F, Calarese D, Stevens J, Razi N, Stevens DJ, Skehel JJ, van Die I, Burton DR, Wilson IA, Cummings R, Bovin N, Wong CH, Paulson JC (2004) Printed covalent glycan array for ligand profiling of diverse glycan binding proteins. *Proc Natl Acad Sci U S A* 101:17033–17038
25. Tateno H, Mori A, Uchiyama N, Yabe R, Iwaki J, Shikanai T, Angata T, Narimatsu H, Hirabayashi J (2008) Glycoconjugate microarray based on an evanescent-field fluorescence-assisted detection principle for investigation of glycan-binding proteins. *Glycobiology* 18:789–798
26. Angata T (2006) Molecular diversity and evolution of the Siglec family of cell-surface lectins. *Mol Divers* 10:555–566
27. Barondes SH, Castronovo V, Cooper DN, Cummings RD, Drickamer K, Feizi T, Gitt MA, Hirabayashi J, Hughes C, Kasai K, Leffler H, Liu F-T, Lotan R, Mercurio AM, Monsigny M, Pillai S, Poirer F, Raz A, Rigby PWJ, Rini JM, Wang JL (1994) Galectins: a family of animal beta-galactoside-binding lectins. *Cell* 76:597–598
28. Kuno A, Kato Y, Matsuda A, Kaneko MK, Ito H, Amano K, Chiba Y, Narimatsu H, Hirabayashi J (2008) Focused differential glycan analysis with the platform antibody-assisted lectin profiling for glycan-related biomarker verification. *Mol Cell Proteomics* 8:99–108
29. Ogata S, Muramatsu T, Kobata A (1975) Fractionation of glycopeptides by affinity column chromatography on concanavalin A-Sepharose. *J Biochem* 78:687–696
30. Cummings RD, Kornfeld S (1982) Characterization of the structural determinants required for the high affinity interaction of asparagine-linked oligosaccharides with immobilized *Phaseolus vulgaris* leucoagglutinating and erythroagglutinating lectins. *J Biol Chem* 257:11235–11240
31. Kasai K, Oda Y, Nishikata M, Ishii S (1986) Frontal affinity chromatography: theory for its application to studies on specific interactions of biomolecules. *J Chromatogr* 376:33–47
32. Nakamura-Tsuruta S, Uchiyama N, Hirabayashi J (2006) High-throughput analysis of lectin-oligosaccharide interactions by automated frontal affinity chromatography. *Methods Enzymol* 415:311–325
33. Tateno H, Nakamura-Tsuruta S, Hirabayashi J (2007) Frontal affinity chromatography: sugar-protein interactions. *Nat Protoc* 2:2529–2537
34. Uchiyama N, Kuno A, Tateno H, Kubo Y, Mizuno M, Noguchi M, Hirabayashi J (2008) Optimization of evanescent-field fluorescence-assisted lectin microarray for high-sensitivity detection of monovalent oligosaccharides and glycoproteins. *Proteomics* 8:3042–3050
35. Lee RT, Lee YC (2000) Affinity enhancement by multivalent lectin-carbohydrate interaction. *Glycoconj J* 17:543–551

36. Brewer CF, Miceli MC, Baum LG (2002) Clusters, bundles, arrays and lattices: novel mechanisms for lectin-saccharide-mediated cellular interactions. *Curr Opin Struct Biol* 12:616–623
37. Yamada M (2009) Lectin microarrays. In: Matson RS (ed) *Microarray methods and protocols*. CRC Press, Taylor & Francis Group, Boca Raton, pp 139–154
38. Koshi Y, Nakata E, Yamane H, Hamachi I (2006) A fluorescent lectin array using supramolecular hydrogel for simple detection and pattern profiling for various glycoconjugates. *J Am Chem Soc* 128:10413–10422
39. Abbas A, Linman MJ, Cheng Q (2011) New trends in instrumental design for surface plasmon resonance-based biosensors. *Biosens Bioelectron* 26:1815–1824
40. Tateno H, Kuno A, Hirabayashi J (2008) How to determine specificity: from lectin profiling to glycan profiling. In: Gabius H-J (ed) *Sugar code: fundamentals of glycoscience*. Wiley, Weinheim, pp 247–259
41. Hirabayashi J (2009) In: Pierce M, Cummings R (eds) *Handbook of glycomics*. Elsevier, Amsterdam, pp 161–176
42. Kuno A, Itakura Y, Toyoda M, Takahashi Y, Yamada M, Umezawa A, Hirabayashi J (2008) Development of a data-mining system for differential profiling of cell glycoproteins based on lectin microarray. *J Proteomics Bioinform* 1:68–72
43. Kuno A, Matsuda A, Ikehara Y, Narimatsu H, Hirabayashi J (2010) Differential glycan profiling by lectin microarray targeting tissue specimens. *Methods Enzymol* 478:165–179
44. Tateno H, Kuno A, Itakura Y, Hirabayashi J (2010) A versatile technology for cellular glycomics using lectin microarray. *Methods Enzymol* 478:181–195
45. Toyoda M, Yamazaki-Inoue M, Itakura Y, Kuno A, Ogawa T, Yamada M, Akutsu H, Takahashi Y, Kanzaki S, Narimatsu H, Hirabayashi J, Umezawa A (2011) Lectin microarray analysis of pluripotent and multipotent stem cells. *Genes Cells* 16:1–11
46. Tateno H, Toyoda M, Saito S, Onuma Y, Ito Y, Hiemori K, Fukumura M, Nakasu A, Nakanishi M, Ohnuma K, Akutsu H, Umezawa A, Horimoto K, Hirabayashi J, Asashima M (2011) Glycome diagnosis of human induced pluripotent stem cells using lectin microarray. *J Biol Chem* 286:20345–20353
47. Saito S, Onuma Y, Ito Y, Tateno H, Toyoda M, Akutsu H, Nishino K, Chikazawa E, Fukawatase Y, Miyagawa Y, Okita H, Kiyokawa N, Shimma Y, Umezawa A, Hirabayashi J, Horimoto K, Asashima M (2011) Possible linkages between the inner and outer cellular states of human induced pluripotent stem cells. *BMC Syst Biol* 5(Suppl):S17
48. Matsuda A, Kuno A, Ishida H, Kawamoto T, Shoda J, Hirabayashi J (2008) Development of an all-in-one technology for glycan profiling targeting formalin-embedded tissue sections. *Biochem Biophys Res Commun* 370:259–263
49. Matsuda A, Kuno A, Kawamoto T, Matsuzaki H, Irimura T, Ikehara Y, Zen Y, Nakanuma Y, Yamamoto M, Ohkohchi N, Shoda J, Hirabayashi J, Narimatsu H (2010) *Wisteria floribunda* agglutinin-positive mucin 1 is a sensitive biliary marker for human cholangiocarcinoma. *Hepatology* 52:174–182
50. Gustafsson A, Sjöblom M, Strindeliuss L, Johansson T, Fleckenstein T, Chatzissavidou N, Lindberg L, Angström J, Rova U, Holgersson J (2011) *Pichia pastoris*-produced mucin-type fusion proteins with multivalent *O*-glycan substitution as targeting molecules for mannose-specific receptors of the immune system. *Glycobiology* 21:1071–1086
51. Huang W, Wang D, Yamada M, Wang LX (2009) Chemoenzymatic synthesis and lectin array characterization of a class of *N*-glycan clusters. *J Am Chem Soc* 131:17963–17971
52. Amano K, Chiba Y, Kasahara Y, Kato Y, Kaneko MK, Kuno A, Ito H, Kobayashi K, Hirabayashi J, Jigami Y, Narimatsu H (2008) Engineering of mucin-type human glycoproteins in yeast cells. *Proc Natl Acad Sci U S A* 105:3232–3237
53. Miyagawa S, Takeishi S, Yamamoto A, Ikeda K, Matsunari H, Yamada M, Okabe M, Miyoshi E, Fukuzawa M, Nagashima H (2010) Survey of glycoantigens in cells from α 1-3galactosyltransferase knockout pig using a lectin microarray. *Xenotransplantation* 17:61–70

54. Suen KF, Turner MS, Gao F, Liu B, Althage A, Slavin A, Ou W, Zuo E, Eckart M, Ogawa T, Yamada M, Tuntland T, Harris JL, Trauger JW (2010) Transient expression of an IL-23R extracellular domain Fc fusion protein in CHO vs. HEK cells results in improved plasma exposure. *Protein Expr Purif* 71:96–102
55. Nakagawa T, Takeishi S, Kameyama A, Yagi H, Yoshioka T, Moriwaki K, Masuda T, Matsumoto H, Kato K, Narimatsu H, Taniguchi N, Miyoshi E (2010) Glycomic analyses of glycoproteins in bile and serum during rat hepatocarcinogenesis. *J Proteome Res* 9:4888–4896
56. Kondoh G, Watanabe H, Tashima Y, Maeda Y, Kinoshita T (2009) Testicular angiotensin-converting enzyme with different glycan modification: characterization on glycosylphosphatidylinositol-anchored protein releasing and dipeptidase activities. *J Biochem* 145:115–121
57. Togayachi A, Kozono Y, Ishida H, Abe S, Suzuki N, Tsunoda Y, Hagiwara K, Kuno A, Ohkura T, Sato N, Sato T, Hirabayashi J, Ikehara Y, Tachibana K, Narimatsu H (2007) Poly-lactosamine on glycoproteins influences basal levels of lymphocyte and macrophage activation. *Proc Natl Acad Sci U S A* 104:15829–15834
58. Kaneko MK, Kato Y, Kameyama A, Ito H, Kuno A, Hirabayashi J, Kubota T, Amano K, Chiba Y, Hasegawa Y, Sasagawa I, Mishima K, Narimatsu H (2007) Functional glycosylation of human podoplanin: glycan structure of platelet aggregation-inducing factor. *FEBS Lett* 581:331–336
59. Futakawa S, Nara K, Miyajima M, Kuno A, Ito H, Kaji H, Shirotani K, Honda T, Tohyama Y, Hoshi K, Hanzawa Y, Kitazume S, Imamaki R, Furukawa K, Tasaki K, Arai H, Yuasa T, Abe M, Arai H, Narimatsu H, Hashimoto Y (2011) A unique N-glycan on human transferrin in CSF: a possible biomarker for iNPH. *Neurobiol Aging* 33(8):1807–1815
60. Fry SA, Afrough B, Lomax-Browne HJ, Timms JF, Velentzis LS, Leatham AJ (2011) Lectin microarray profiling of metastatic breast cancers. *Glycobiology* 21:1060–1070
61. Kuwamoto K, Takeda Y, Shirai A, Nakagawa T, Takeishi S, Ihara S, Miyamoto Y, Shinzaki S, Ko JH, Miyoshi E (2010) Identification of various types of α 2-HS glycoprotein in sera of patients with pancreatic cancer: possible implication in resistance to protease treatment. *Mol Med Report* 3:651–656
62. Sasaki N, Moriwaki K, Uozumi N, Noda K, Taniguchi N, Kameyama A, Narimatsu H, Takeishi S, Yamada M, Koyama N, Miyoshi E (2009) High levels of E4-PHA-reactive oligosaccharides: potential as marker for cells with characteristics of hepatic progenitor cells. *Glycoconj J* 26:1213–1223
63. Krishnamoorthy L, Bess JW Jr, Preston AB, Nagashima K, Mahal LK (2009) HIV-1 and microvesicles from T cells share a common glycome, arguing for a common origin. *Nat Chem Biol* 5:244–250
64. Hsu KL, Pilobello KT, Mahal LK (2006) Analyzing the dynamic bacterial glycome with a lectin microarray approach. *Nat Chem Biol* 2:153–157
65. Ebe Y, Kuno A, Uchiyama N, Koseki-Kuno S, Yamada M, Sato T, Narimatsu H, Hirabayashi J (2006) Application of lectin microarray to crude samples: differential glycan profiling of lectin mutants. *J Biochem* 139:323–327
66. Tateno H, Uchiyama N, Kuno A, Togayachi A, Sato T, Narimatsu H, Hirabayashi J (2007) A novel strategy for mammalian cell surface glycome profiling using lectin microarray. *Glycobiology* 17:1138–1146
67. Yasuda M, Tateno H, Hirabayashi T, Iino T, Sako T (2011) Lectin microarray reveals binding profiles of *Lactobacillus casei* strains in a comprehensive analysis of bacterial cell wall polysaccharides. *Appl Environ Microbiol* 77:4539–4546
68. Pilobello KT, Slawek DE, Mahal LK (2007) A ratiometric lectin microarray approach to analysis of the dynamic mammalian glycome. *Proc Natl Acad Sci U S A* 104:11534–11539
69. Tonge R, Shaw J, Middleton B, Rowlinson R, Rayner S, Young J, Pognan F, Hawkins E, Currie I, Davison M (2001) Validation and development of fluorescence two-dimensional differential gel electrophoresis proteomics technology. *Proteomics* 1:377–396

70. Propheter DC, Hsu KL, Mahal LK (2011) Recombinant lectin microarrays for glycomic analysis. *Methods Mol Biol* 723:67–77
71. Yabe R, Suzuki R, Kuno A, Fujimoto Z, Jigami Y, Hirabayashi J (2007) Tailoring a novel sialic acid-binding lectin from a ricin-B chain-like galactose-binding protein by natural evolution-mimicry. *J Biochem* 141:389–399
72. Hu D, Tateno H, Kuno A, Yabe R, Hirabayashi J (2012) Directed evolution of lectins with sugar-binding specificity for 6-sulfo-galactose. *J Biol Chem* 287:20313–20320
73. Hu D, Tateno H, Sato T, Narimatsu H, Hirabayashi J (2013) Tailoring GalNAc α 1-3Gal β -specific lectins from a multi-specific fungal galectin: dramatic change of carbohydrate specificity by a single amino-acid substitution. *Biochem J* 453:261–270
74. Bicker KL, Sun J, Lavigne JJ, Thompson PR (2011) Boronic acid functionalized peptidyl synthetic lectins: combinatorial library design, peptide sequencing, and selective glycoprotein recognition. *ACS Comb Sci* 13:232–243

Sialoside Arrays: New Synthetic Strategies and Applications

Chi-Hui Liang, Che-Hsiung Hsu, and Chung-Yi Wu

Abstract Sialic acid-containing carbohydrates, or sialosides, play critical roles in many biological events and in diseases, including viral and bacterial infections, the immune response, the progression of tumor cell metastasis, etc. Despite the importance, the limited access to complex sialosides had prevented extensive studies on the function and significance of sialic acid structural diversity. However, recent advances in synthetic sialoside chemistry, such as the novel chemoenzymatic or stereochemical approach, have produced homogeneous size- and structure-defined sialosides to create diverse sialosides for array application. The advantage of sialoside arrays is the multivalent display of arrayed sialosides which can serve to mimic cell surface display; thus, an array-based technique is well suited for investigations of the real sialoside-mediated interactions in nature. In brief, this chapter discusses the novel strategies for synthesizing sialosides with selected examples of applications to illustrate the potential of sialoside arrays and further forecast to the trend of using nanotechnology in sialoside arrays.

Keywords Array · Nanoparticle · Sialic acid · Sialoside · Sialyltransferases

Contents

1	Introduction	127
2	Synthesis of Sialosides	128
2.1	Chemical Synthesis	128
2.2	Chemoenzymatic Synthesis	131
3	Examples of the Clinical and Biological Significance of Sialic Acid Diversity	135
3.1	Gangliosides in Tumor Research	136
3.2	Sialyl Tn in Tumor Research	137

C.-H. Liang (✉), C.-H. Hsu, and C.-Y. Wu
Genomics Research Center, Academia Sinica, 128 Academia Road, Section 2, Nankang,
Taipei 115, Taiwan
e-mail: JamesL@gate.sinica.edu.tw; cyiwu@gate.sinica.edu.tw

3.3	Sialyl Lewis Structures in Tumor Research	138
3.4	Sialosides in Host-Pathogen Infection Research	139
4	Sialoside Array	140
4.1	Applications of Sialoside Arrays in Influenza Virus Study	140
4.2	Applications of Sialoside Arrays in Search of Siglecs Inhibitors	141
5	Synergy of Nanotechnology and Sialoside Arrays	142
6	Conclusions	143
	References	144

Abbreviations

Ac	Acetyl
Ac ₂ O	Acetic anhydride
AcOH	Acetic acid
Bn	Benzyl
Bz	Benzoyl
Cbz	Bezyloxycarbonyl
CMP	Cytidine monophosphate
CSA	Camphorsulfonic acid
DCM	Dichloromethane
DMF	<i>N,N</i> -Dimethylformamide
DMTST	Dimethyl(methylthio)sulfonium trifluoromethanesulfonate
EA	Ethyl acetate
ERBB	Epidermal growth factor receptor
Et	Ethyl
Gal	Galactose
GalNAc	<i>N</i> -Acetylgalactosamine
Glc	Glucose
GlcNAc	<i>N</i> -Acetylglucosamine
KDN	Deaminoneuraminic acid
Man	Mannose
ManNAc	<i>N</i> -Acetylmannosamine
MeOH	Methanol
MeOTf	Methyl trifluoromethane sulfonate
NBz	<i>p</i> -Nitrobenzoyl
Neu5Ac	<i>N</i> -Acetylneuraminic acid
Neu5Gc	<i>N</i> -Glycolylneuraminic acid
NIS	<i>N</i> -Iodosuccinimide
Ph	Phenyl
PhSOTf	Phenyl sulfenyl triflate
RRVs	Relative relativity values
RT	Room temperature
SBox	<i>S</i> -Benzoxazolyl
Siglec	Sialic acid-binding immunoglobulin-type lectins
TfOH	Trifluoromethanesulfonic acid

TMSOTf Trimethylsilyl trifluoromethanesulfonate
Tn GalNAc α -Ser/Thr; Tn is a precursor for many extended O-glycans in animal glycoproteins

1 Introduction

Sialic acids, also called neuraminic acids, are normally found in nature at the terminal positions of *N*-glycans, *O*-glycans, glycosphingolipids, and glycoposphoinositol anchors attached to protein or lipid moieties [1]. Sialic acids are a family of α -keto acids with a nine-carbon backbone. In nature, sialic acids usually do not exist in free sugar forms. They are commonly α 2,3- and α 2,6-linked to terminal galactose, GalNAc or, α 2,8- and α 2,9-linked to another sialic acid residue [2–4]. Sialic acids are usually the outermost residues and direct participants in many molecular recognitions and interactions. Moreover, sialic acid-containing carbohydrates, or sialosides, play significant biological roles in cellular recognition, cell communication, pathogen infections, tumor metastasis, and disease states, which have led to extensive studies in sialobiology [5–7]. Because of their exposed position and diversity, sialosides represent a significant biological target in either masking recognition sites or facilitating cell recognition and adhesion [3]. Additionally, the type of sialyl linkage between the sialic acid and its adjacent carbohydrate moiety can decisively influence the function of sialosides. For example, avian influenza viruses favor the α 2,3-sialic acid receptor, whereas human adapted influenza A viruses use the α 2,6-sialic acid receptor [8, 9]. Obviously, the biological significance of sialosides is enormous; however, because of their structural complexity and diversity, sialosides are extremely difficult to gather from natural sources in amounts abundant enough for biological studies. Recent advances in chemical synthesis and chemoenzymatic approaches successfully produced structurally defined sialosides in homogeneous forms [10, 11]. However, unlike common monosaccharides such as galactose and glucose, the three additional carbons make sialic acids more complicated for structural modification than any other monosaccharides [11]. Because of the various sialyl linkages and different adjacent carbohydrate moieties, obtaining homogeneous sialosides by using chemical or chemoenzymatic synthesis remains a tremendous task.

To assess recent advances involving sialoside array preparation and applications, this chapter first documents new developments in sialoside synthesis and then reviews several examples of sialoside array studies. The discussion includes aspects of sialobiology and sialochemistry and details work towards the new synthetic strategies and development of sialoside arrays, including new sourcing of sialosides, new donors for sialoside synthesis, new sialoside biosynthetic enzymes, new applications for biological study, and perceptions of sialosides for improving the understanding of sialobiology. This chapter also focuses on ways in which synergy of nanotechnology and sialoside arrays can be applied in diseases validation, with the example of influenza virus subtypes determination. Another example in array applications is sialic acid recognizing proteins, such as hemagglutinins,

selectins, and siglecs, an area of increasing biological interest and importance. This chapter further discusses future trends and forecasts possible solutions for current challenges in the development of sialoside arrays.

2 Synthesis of Sialosides

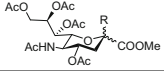


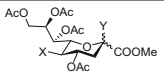


Preparation of structurally diverse sialosides can be achieved through enzymatic synthesis, chemical synthesis, or modification of the natural sialosides [11–13]. Advances in both enzymatic and chemical syntheses have provided reliable procedures to produce complex sialosides. Based on different modifications of the C5 position, sialic acids are classified in three basic forms: Neu5Ac, Neu5Gc, and KDN. Besides the C5 position, single or multiple substitutions can occur at the hydroxyl groups at C4, C7, C8, and C9 positions (substituted by acetyl, sulfate, methyl, lactyl, phosphate, etc.). These modifications generate a diverse family of more than 50 structurally distinct forms of sialosides in nature [2], which are of key importance to cellular recognition, pathogen infections, and disease states [3]. The following pages concisely summarize up-to-date progress on the synthesis of sialosides and offer stimulating information on the rapidly developing field of the chemical and enzymatic syntheses of sialosides. The discussion also includes breakthrough developments in the study of conformational and stereo-selective effects of sialosides synthesis and compares and contrasts the chemical and enzymatic synthesis approaches.

2.1 Chemical Synthesis

Chemical sialylation has been considered to be one of the most challenging glycosylation reactions. The challenges come from the presence of the C1 electron-withdrawing carboxyl group at the tertiary anomeric center which reduces the reactivity of sialic acid as a donor, and the lack of a participating group at C3 to direct the stereochemical outcome of glycosylation. Moreover, each sialic acid carries at least six hydroxyl groups which must be protected and deprotected during synthesis. Glycosylation in each step generates a new stereo-center at the anomeric carbon, and there are no universal methods for the introduction of a desired glycosidic linkage in a stereo-controlled manner. The deciding factors in chemical synthesis of sialosides include protection schemes, functional substitutions, promoter choice, and acceptor architecture. In particular, significant efforts have been directed toward the development of sialic acid donors for efficient α -sialylation [11, 14–20], including the use of anomeric leaving groups, such as halides [21–23], phosphites [24–27], sulfides [28, 29], xanthates [30–32], and phenyltrifluoroacetimidates [33–35], the introduction of auxiliary groups at C-1 [36–40] and C3 [23, 41–49], the modification of the *N*-acetyl functional group at C5 [50, 51], or the optimized combinations of the leaving group with positional modification [52–63]. Among them, glycosylation using various sialic acid donors with different leaving groups and C5 modifications is a most powerful method to increase yield and

improve α -selectivity. Table 1 summarizes the results of glycosylation using various sialic acid donors with different leaving groups and C5-modifications. Acceptors in Table 1 are frequently used for direct comparison. Two glycosylation results, one with primary and the other with secondary acceptors, were included for

Table 1 Review of sialic acid donors with different leaving groups and C5 modifications

Entry	Donor	Acceptor	Yield (%)		Promoter	Solvent ^a	Reference	
			α	β				
1		I (1°)	67	0	Ag ₂ CO ₃	e	[22]	
		X (2°)	6	9	Hg(CN) ₂ / HgBr ₂	f	[64]	
2	R=OP(OBn) ₂	II (1°)	67	13	TMSOTf	a	[24]	
		XI (2°)	67	11	TMSOTf	a	[24]	
3	R=OP(OEt) ₂	II (1°)	56	14	TMSOTf	a	[26]	
		XII (2°)	55		TMSOTf	a	[26]	
4		III (1°)	48	16	DMTST	a	[31]	
		XIII (2°)	70	4	PhSOTf	c	[32]	
5	R=SMe	IV (1°)	70	0	DMTST	a	[65]	
		XIV (2°)	52	0	DMTST	a	[65]	
6	R=SPh	II (1°)	47	8	NIS/TfOH	a	[66]	
		XIV (2°)	70	0	NIS/TfOH	a	[67]	
7		II (1°)	69	10	TMSOTf	c	[33]	
		XV (2°)	61	20	TMSOTf	c	[33]	
8	R=SBox	I (1°)	60	30	MeOTf	a	[28]	
		XIV (2°)	71	4	MeOTf	a	[28]	
9	R=OPO(OBn) ₂	II (1°)	26	9	TMSOTf	a	[24]	
10	R=OPO(OEt) ₂	XII (2°)	11	11	TMSOTf	a	[25]	
11		X=NAc ₂ , Y=SPh	II (1°)	40	25	NIS/TfOH	a	[66]
		X=NAc ₂ , Y=SMe	XIV (2°)	72		NIS/TfOH	a	[68]
12	X=NHTroc, Y=SPh	II (1°)	81	10	NIS/TfOH	a	[66]	
		XIV (2°)	35	8	NIS/TfOH	a	[69]	
13	X=NHTFA, Y=SPh	II (1°)	85	7	NIS/TfOH	a	[66]	
		X=NHTFA, Y=SMe	XVI (2°)	84	0	NIS/TfOH	a	[70]
14	X=N ₃ , Y=STol	V (1°)	65	0	NIS/TfOH	a	[63]	
15	X=NPhth, Y= 	VII (1°)	92	0	TMSOTf	d	[61]	
		XVII (2°)	74	2	TMSOTf	d	[61]	
16	X=TCA, Y= 	VIII (1°)	77	0	TMSOTf	c	[59]	
		XVIII (2°)	68	0	TMSOTf	c	[59]	

(continued)

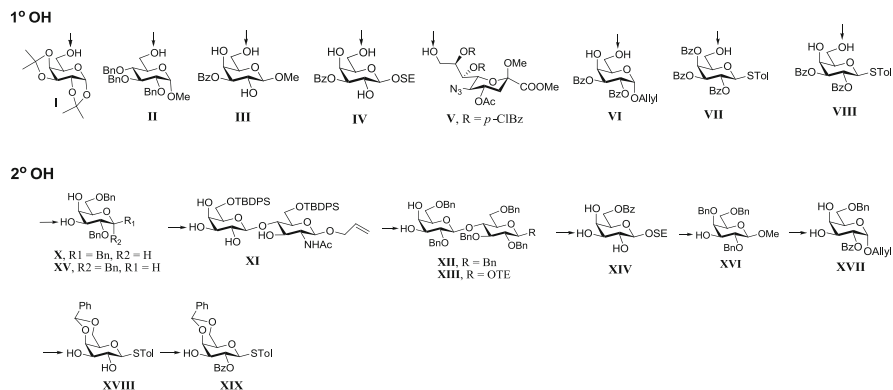
Table 1 (continued)

Entry	Donor	Acceptor	Yield (%)		Promoter	Solvent ^a	Reference
			α	β			
17	X=H, Y=SPh	II (1°)	100	0	NIS/TfOH	b	[71]
		XIV (2°)	50	0	NIS/TfOH	a	[52]
18	X=Ac, Y=SPh	II (1°)	92	0	NIS/TfOH	b	[53]
		XVI (2°)	10	75	NIS/TfOH	b	[53]
19	X=Ac, Y=STol	II (1°)	72	18	NIS/TfOH	b	[57]
		XVI (2°)	44	37	NIS/TfOH	c	[57]
20	X=Ac, Y =	II (1°)	91	0	NIS/TfOH	b	[52]
		XVI (2°)	64	21	NIS/TfOH	c	[52]
21	X=Ac, Y=OPO (OBu) ₂	IX	85	0	TMSOTf	b	[56]
		XIX	83	0	TMSOTf	b	

^aSolvents: a=MeCN; b=CH₂Cl₂; c=MeCN + CH₂Cl₂; d=EtCN; e=CHCl₃; f=Cl(CH₂)₂Cl

every sialyl donor. Generally, the yields of sialylation by using donors with different leaving groups (entries 1–10) did not exceed 70% yield, and a higher α -selectivity was obtained when a less hindered glycosyl acceptor was used (entries 2, 5, 6, and 8). Notably, most of these reactions need to be conducted in acetonitrile. A second type of sialic acid donor is with C5-modification, which often greatly enhanced not only the α -selectivity but also the yield toward primary or secondary acceptors (entries 11–21). After comparing the glycosylation results of N-modified donors by coupling with primary acceptor **II** (entries 11–13 and 17–20), 5-*N*, 4-*O*-oxazolidinone-protected donors showed the best result (entries 17–20). On the other hand, for less hindered secondary acceptor **XIV** (entries 11, 12, and 17), the *N*, *N*-diacetyl sialyl donor gave the best outcome. In addition, for more hindered secondary acceptor **XV** (entries 13, 18, 19, and 20), *N*-TFA sialyl donor seems to be a good choice, and conversion of the leaving group to the adamantanylthio group of *N*-acetyl-5-*N*,4-*O*-oxazolidinone-protected donors showed further improvement (entry 20). Next, more improvement was achieved by using a combination of C5-modification and efficient leaving group (entries 15 and 16). The products gave excellent α -selectivity and yield toward both primary and secondary alcohols. The *N*-acetyl-5-*N*,4-*O*-oxazolidinone-protected donor indeed showed an excellent α -selectivity as well as high yield toward both primary and secondary acceptors, and the phosphate leaving group appears to be an excellent choice (entry 21).

Acceptors



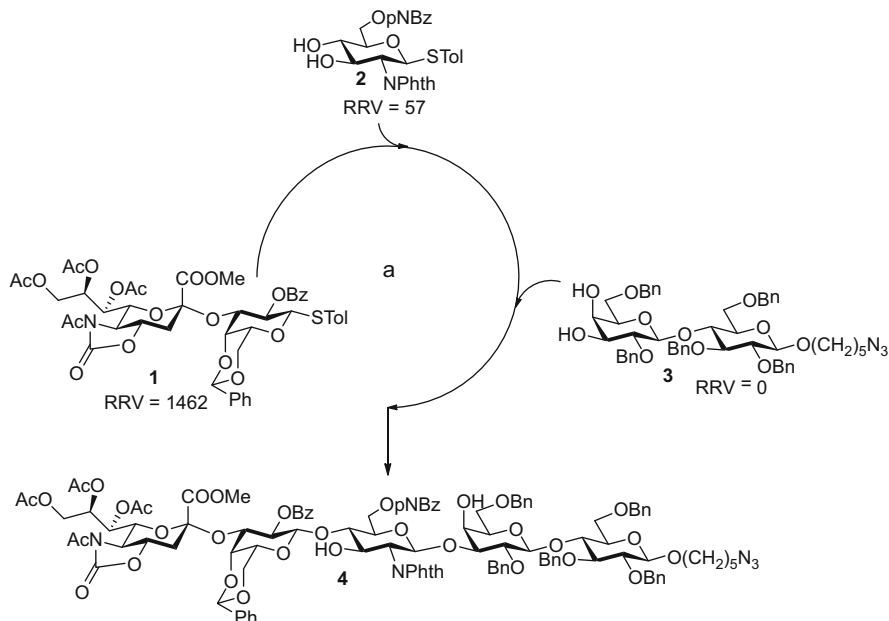
Another advantage of the phosphate-based methodology includes the use of tolylthio glycosides as acceptors with defined relative reactivity values (RRVs) for sialylation to give sialyl disaccharides as building blocks for the subsequent reactivity-based programmable one-pot synthesis. In this way, the limitation of relatively low reactivity of thiosialoside donor and the difficult control of stereoselectivity of sialylation could be resolved. Moreover, the RRVs of sialylated disaccharides can be programmed by manipulating the protecting groups of the second sugar residue at the reducing end and can be applied to the synthesis of α 2,3-linked sialylated pentasaccharide **4** (Scheme 1) [56].

More importantly, the phosphate donor can also be applied to the convergent block synthetic strategy to synthesize the various lengths of α 2,9-linked oligosialic acid up to dodecamer in good yield and alpha selectivity (Scheme 2) [72].

2.2 Chemoenzymatic Synthesis

In sialoside synthesis, the use of enzymes can dramatically reduce the multiple protection-deprotection steps and provide great stereo- and regio-selectivity. Many enzymes accept unnatural substrates. Genetic engineering can further alter their stability, broaden their substrate specificity, and increase their specific activity [73]. Therefore, chemoenzymatic approaches using enzymes in combination with classical carbohydrate chemistry opened up a new venue, avoiding many of the problems encountered in traditional chemical synthesis.

Sialyltransferases introducing the sialic acid moiety onto saccharides for constructing the sialosidic linkage are typically employed in chemoenzymatic synthesis. The chiral nature of sialyltransferases results in the formation of stereo- and regio-chemically defined sialosides. Depending on the type of sialyltransferase used, α 2,3-, α 2,6-, or α 2,8-linked sialosides can be synthesized with remarkable rate acceleration. Generally speaking, each sialyltransferase is specific for a

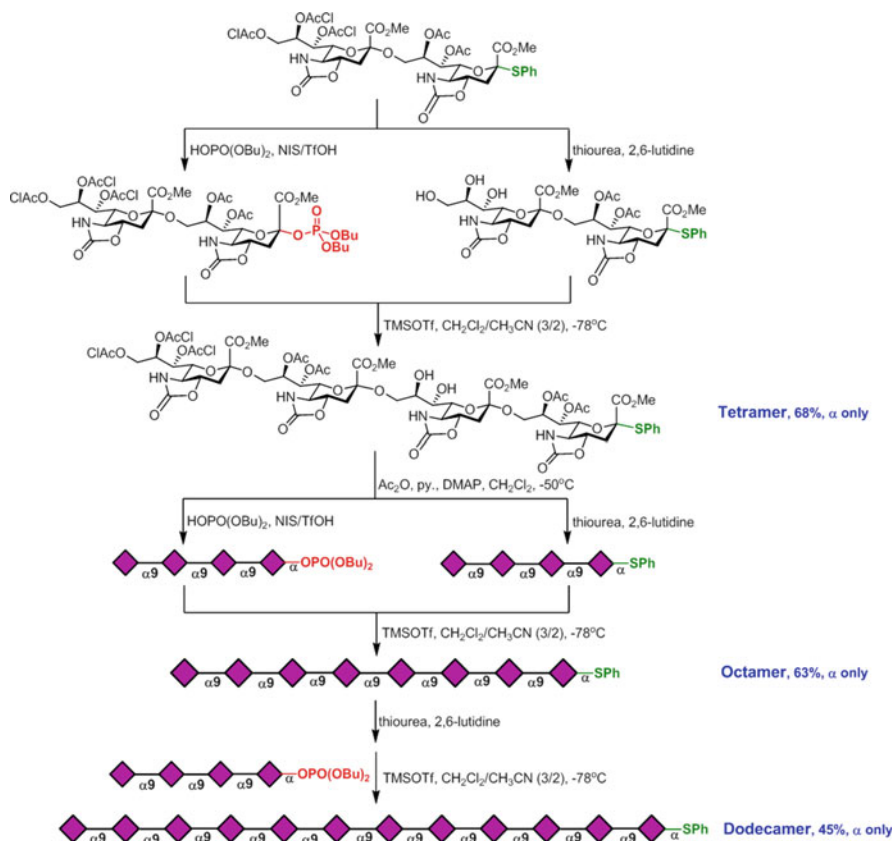


Scheme 1 Programmable one-pot synthesis of sialylated pentasaccharide **4**

1. NIS, TfOH, 4Å molecular sieves, CH₂Cl₂, -78 °C
2. NIS, TfOH, -20 °C to RT

particular substrate and can be classified into four families according to the carbohydrate linkage they synthesize: ST3Gal (α 2,3-ST), ST6Gal (α 2,6-ST), ST6GalNAc, and ST8Sia (α 2,8-ST) families. In detail, the ST3Gal family transfers sialic acid residues in α 2,3-linkage to terminal galactose residues. The ST6Gal family comprises two subfamilies, ST6Gal-I and -II, which both use the Gal β -1,4GlcNAc-R as the acceptor substrate and add sialic acid with an α 2,6-linkage to galactose or GalNAc. The ST8Sia family mediates the transfer of sialic acid residues in α 2,8-linkage to other sialic acid residues forming polysialic acid structures.

The total synthesis of sialosides by using the chemoenzymatic approach is as follows [74]. Sialic acid itself can be synthesized from ManNAc, mannose, or their derivatives by sialic acid aldolase enzyme through aldol condensation reaction. If ManNAc is chemically or enzymatically modified at C2, C4–C6 positions, sialic acid has structural modifications at C5, C7–C9 positions, respectively. The sialic acids are subsequently activated by a CMP-sialic acid synthetase to form a CMP-sialic acid, which is the donor used by sialyltransferases. Because CMP-sialic acid is unstable, the CMP-Neu5Ac synthetase is valuable for the preparative enzymatic synthesis of sialosides. In the last steps, the CMP-sialic acid is transferred to galactose or GalNAc terminated glycosides by sialyltransferases to form structurally defined sialosides. Examples are that Chen and co-workers have recently developed a one-pot multienzyme system for the efficient synthesis of α -sialosides (Table 2) [12, 76, 79]. In this system, recombinant *E. coli* K-12 sialic acid aldolase catalyzed the synthesis of sialic acid precursors for



Scheme 2 Preparation of α 2,9-oligosialic acids up to dodecamer

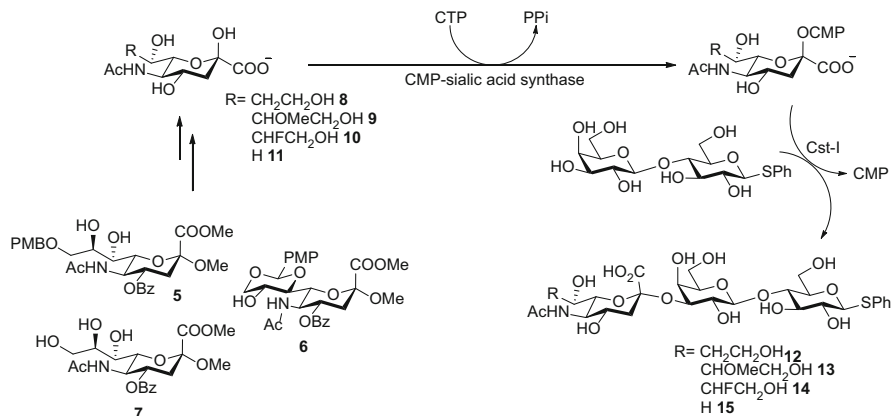
CMP-sialic acid from their corresponding six-carbon derivatives (mannose, *N*-acetylmannosamine, and other derivatives). CMP-sialic acid derivatives were then generated in situ from sialic acid precursors catalyzed by a recombinant CMP-sialic acid synthetase NmCSS and subsequently used by sialyltransferase to transfer the sialic acid residue to a sugar acceptor for the formation of natural sialosides. These three steps can be combined into a one-pot sequence without purification of intermediates. This approach has been demonstrated to be highly efficient for systematically synthesizing sialoside libraries with great diversity. A number of nature-occurring or non-natural sialoside derivatives containing α 2,3- [76], α 2,6- [10], and α 2,8- [77] linkages were synthesized in large quantities and high yields with versatile structural modifications, mainly at C5 and C9 positions of sialic acid residues. It is also of great importance that 3-fluoro Neu5Ac- and Neu5Gc-containing sialyl lactose trisaccharides, which are particularly difficult to obtain by organic synthesis, could be synthesized by the use of NmCSS and multifunctional sialyltransferase PmST1 via one-pot two-enzyme synthesis [80]. The 3-fluorinated sialic acid derivatives have been demonstrated to be competitive inhibitors for virus sialidases [81, 82].

Table 2 Example of one-pot multienzyme approach for the synthesis of sialosides with versatile structural modification

Enzyme	From	Function
Aldolase [75]	<i>E. coli</i> K-12	Sialic acid aldolase
NmCSS [75]	<i>Neisseria meningitidis</i>	CMP-sialic acid synthetase
SiaT	<i>Pasteurella multocida</i>	α 2,3-Sialyltransferase
	<i>Photobacterium damsela</i>	α 2,6-Sialyltransferase
	<i>Campylobacter jejuni</i>	α 2,8-Sialyltransferase

$\text{pyruvate} \xrightarrow{\text{Aldolase}} \left[\text{HO} \begin{array}{c} \text{R}_1 \\ \text{R}_2\text{O} \end{array} \begin{array}{c} \text{OH} \\ \text{OH} \end{array} \text{COO}^- \right] \xrightarrow{\text{NmCSS}} \left[\text{HO} \begin{array}{c} \text{R}_1 \\ \text{R}_2\text{O} \end{array} \begin{array}{c} \text{OH} \\ \text{OH} \end{array} \text{CO}_2^- \right] \xrightarrow{\text{SiaT}} \left[\text{HO} \begin{array}{c} \text{R}_1 \\ \text{R}_2\text{O} \end{array} \begin{array}{c} \text{OH} \\ \text{OH} \end{array} \text{OR} \right] + \text{CMP}$

Man(NAc) or derivatives
 $\text{R}_1 = \text{NHAc, NHGc, OH, OAc, NHAcOMe, HHAcOAc, OMe}$
 $\text{R}_2 = \text{H, Ac, lactyl}$



Scheme 3 Chemoenzymatic synthesis of sialyl lactose trisaccharide with C8 modifications

A strategy of chemoenzymatic synthesis of C8-modified sialic acid analogues was recently reported by Withers and co-workers (Scheme 3) [83]. The C8-modified sialic acid precursors **8–11** were synthesized from compounds **5–7**, and each was converted to its CMP donor using a bacterial CMP-sialic acid synthetase. The Cst-1, an α 2,3-sialyltransferase from *Campylobacter jejuni*, was then selected and successfully used for the synthesis of sialyl thiolactoside **12–15**. Notably, Cst-1 was shown to be more desirable for the synthesis of C8-modified sialyl lactose in comparison with another sialyltransferase PM0188h, which exhibited more hydrolysis activity toward natural substrates.

In another piece of work reported by Chen and co-workers, two bacterial β -1,4-galactosyltransferases, NmLgtB and Hp1-4GalT, showed promiscuous and complementary specificity for GlcNAc monosaccharide substrates with 6-O-sulfation or N-sulfation [84]. Using these enzymes in the one-pot multienzyme synthesis efficiently generated an array of LacNAc and lactose derivatives [84]. Moreover, chemoenzymatic synthesis of sulfur-linked saccharides was also investigated [85]. For example, Withers' group recently reported an efficient synthesis of sulfur-linked disaccharides Gal- β -S-1,4-GlcNAc_pNP and uncommon Gal- β -1,4-Man-*p*NP by a β -1,4-galactosyltransferase HP0826 from *Helicobacter pylori* [86].

3 Examples of the Clinical and Biological Significance of Sialic Acid Diversity

Recent developments in sialosides put new possibilities on the horizon to achieve a profiling of sialoside diversity to discover novel molecular targets. Studies have found that sialyl Tn, sialyl Lewis structures, and the G_M and G_D gangliosides are sialo-tumor associated carbohydrate antigens which are abundantly expressed by a

number of tumors [87]. Their associated biological phenomena are important for tumor proliferation, invasion, metastasis, and angiogenesis [87, 88]. Therefore, sialo-tumor associated carbohydrate antigens become important targets in tumor therapy [89]. In addition to tumor research, most recent interests by the pharmaceutical industry in sialosides rely on their involvement in flu virus infections [90]. In the following sections, the significant clinical and biological roles played by sialosides, both in relation to disease states and microbial infections, are discussed.

3.1 *Gangliosides in Tumor Research*

Gangliosides are a family of glycolipids containing a variable number of sialic acid residues. In a shorthand nomenclature system, the letter G refers to ganglioside, and the subscripts M, D, and T refer to the number of sialic acid residues on the backbone of molecule (mono-, di-, or tri-sialogangliosides, respectively). Findings on the effect of gangliosides on the formation and progression of many diseases, such as Guillain-Barré syndrome [91], Sandhoff disease [92], diabetes [93], Alzheimer's disease [94], and Huntington's disease [95] etc., have been reported

In tumor research, gangliosides play crucial roles at various pathophysiological steps of tumor progression. Gangliosides overexpression and shedding are believed to contribute to tumor formation and progression. For instance, shedding of gangliosides from the surface of the tumor cell has been shown to stimulate endothelial cell proliferation and migration [96–98], inhibit tumor-specific cytotoxic T lymphocyte generation [99], suppress lymphocyte proliferation and antigen-presenting cell function [100, 101], and enhance tumor development by increasing the levels of vascular endothelial growth factor [102, 103]. Gangliosides can also trigger tumor transformation from a dormant to a malignant state. Evidence indicates that the gangliosides G_{M3} or G_{D3} , which are commonly overexpressed in tumors, are able to regulate growth signaling through interactions with receptor tyrosine kinases or protein kinase C [88, 104]. The ganglioside G_{M1} is particularly necessary for ERBB2 and ERBB3 receptors to form heterodimers [105], thereby facilitating ERBB signaling, which serves as important growth-promoting functions for tumor proliferation (Fig. 1a, d). This signaling through the ERBB receptors has already become a target for several tumor drugs [106]. In addition, tumorigenic properties of certain gangliosides affect patient survival, causing faster progression and metastasis, and worse prognosis [107–109]. Clinical studies have found certain ganglioside levels in serum to be upregulated in tumor patients compared to healthy subjects. Examples are that the gangliosides G_{M3} and G_{D3} are overexpressed in about 50% of breast carcinomas [110], and the *N*-glycolyl- G_{M3} is overexpressed in stage II breast cancer with 100% efficiency [111]. To date, mounting evidence suggest that gangliosides are important for prevention and treatment of tumors.

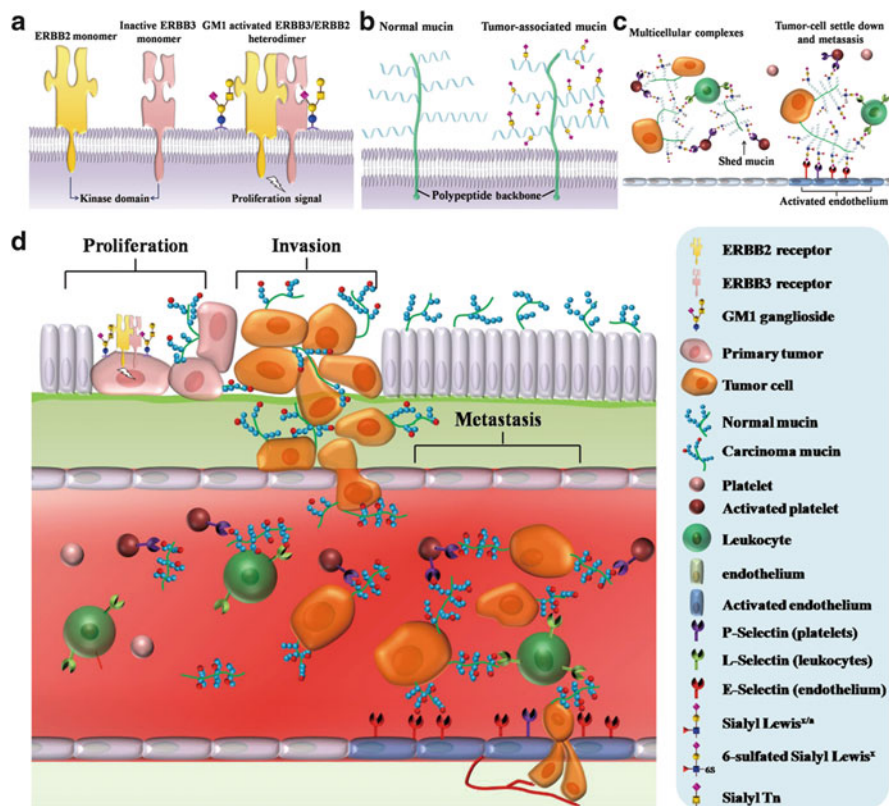


Fig. 1 Schematic representation of the multi-steps of tumor progression. (a) It starts with the tumor transformation from a dormant to a malignant state. (b) Highly expressed sialyl Tn induces the initiation of metastasis. (c) The tumor cells then invade in the bloodstream, where they interact with blood cells, finally adhering to endothelial cells in the vessel walls. (d) After extravasation, they establish new metastatic colonies

3.2 Sialyl Tn in Tumor Research

Tumor invasion is usually associated with the expression of sialosides which promote dissociation of tumor cells. The sialyl Tn antigen, a mucin-type O-linked disaccharide, is one such class of sialosides overexpressed in tumor cells [112]. Membrane-bound mucin known as MUC1 is commonly carrying sialyl Tn in carcinomas. Because the carboxyl group of sialic acids is deprotonated under physiological pH, sialyl Tn-contained mucin has a net negative charge. The rod-like structures of the sialyl Tn-contained mucins are thought to repel intercellular interactions and sterically prevent other adhesion molecules [113, 114]. Thus, sialyl Tn-contained mucins act as “anti-adhesion” factors which can promote displacement of a cell from the primary tumor during the initiation of

metastasis [115] (Fig. 1b, d). Researchers have found that sialyl Tn expression is highly associated with a decreased overall survival of tumor patients [116–118]. Because sialyl Tn is absent in normal healthy tissues but can be detected in almost all kinds of epithelial tumors (i.e., gastric, pancreatic, colorectal, ovarian, and breast tumors), both sialyl Tn and MUC1 have been considered as targets for the development of antitumor vaccine [119–121].

3.3 Sialyl Lewis Structures in Tumor Research

Tumor migration, invasion, extravasation, circulation with blood, interaction with endothelium, and establishment of a microenvironment are all dependent on specific adhesive interactions of tumor cells with other cells and components of the extracellular matrix. Some of these important interactions are mediated by selectins and their ligands—sialyl Lewis^x and sialyl Lewis^a [122–124]. Selectins are adhesion receptors expressed on activated platelets (P-selectin), leukocytes (L-selectin), and endothelial cells (E-selectin) [125]. Unlike normal healthy endothelial cells which do not express E-selectin, primary tumors in patients secrete inflammatory cytokines such as TNF α , which induce E-selectin on endothelial cells [126]. Sialyl Lewis^a and sialyl Lewis^x are ligands for E-selectin, and therefore responsible for the adhesion of human tumor cells to endothelium [127]. With platelets and leukocytes, tumor cells can form multicellular complexes (via P-selectin) and then settle down in the microvasculature (via L-selectin) of distant organs and eventually extravasate and establish metastatic colonies [128] (Fig. 1c, d).

In chemistry, sialyl Lewis^x is a positional isomer of sialyl Lewis^a. Notably, sialyl Lewis^a and sialyl Lewis^x are minimal structures required for selectins recognition. The mucin-type glycoproteins and gangliosides are found to be carriers of sialyl Lewis^a and sialyl Lewis^x in tumors. Several lines of evidence suggest that sialyl Lewis^a antigen is mainly responsible for adhesion of human colon, pancreas, and gastric tumor cells to the endothelium, whereas binding of lung, breast, liver, and ovarian tumor cells is mediated by sialyl Lewis^x [129–131]. Clinico-pathological studies revealed that patients having both increased levels of sialyl Lewis^x and sialyl Lewis^a and enhanced E selectin expression are at a greater risk of developing hematogenous metastasis [132]. Anti-adhesion therapy directed to block the binding between sialyl Lewis^a and E-selectin might confine tumor metastasis. A promising example is to use P-selectin glycoprotein ligand-1, a glycoprotein receptor containing sialyl Lewis^x structures, which inhibit the interaction of selectins with their sialyl Lewis ligands [133, 134].

3.4 Sialosides in Host-Pathogen Infection Research

3.4.1 *Neisseria meningitidis*

Polysialic acids are linear polymers consists of contiguous sialic acid residues linked by α 2,8-, α 2,9-, and alternating α 2,8- and α 2,9-linkages. The degree of polymerization can extend beyond 200 sialic acid residues [135]. Expression of the polysialic acids capsule on the surface of bacteria is important in pathogenesis, because it appears to facilitate bacterial invasion and colonization [136]. The main function of polysialic acid is thought to be to serve as a repulsive element contributing to the net negative charge of the cell surface which regulates intermolecular and intercellular adhesion. Polysialic acids are found on the surfaces of bacteria in different compositions. For example, the capsular polysaccharide of group C *Neisseria meningitidis* consists of α 2,9-linkaged polysialic acid and that of *Escherichia coli* K92 consists of alternating α 2,8- and α 2,9-linkaged sialic acid residues [137–139]. The current vaccines against meningococcal group C diseases are glycoconjugates of isolated α 2,9-polysialic acids with carrier protein [140]. In particular, the capsular polysaccharide of group B *Neisseria meningitidis* consists of α 2,8-linkaged polysialic acid, which not only occurs in bacteria but also is a universal mammalian developmental antigen, and expressed by many human tumors [136]. Because the bacterial α 2,8-polysialic acid chains are structurally identical to the polysialic acid epitopes on human cells, bacteria with α 2,8-polysialic acid capsule may escape immune surveillance. Furthermore, antigenic similarities between brain components and *N. meningitidis* B allow these pathogens to migrate into the brain and cause meningitis [141]. Conversely, surface expression of the polysialic acid capsule on pathogens may provide good targets for the development of bactericidal agents and antibacterial vaccines. Because it is a self-antigen, α 2,8-polysialic acid was structurally modified and then applied as a vaccine against group B meningococci [142].

3.4.2 Influenza Virus

Sialosides on the host cell can be recognized by viral hemagglutinins and neuraminidase which are viral surface proteins facilitating the attachment and release of viruses, respectively. Hemagglutinin binding to sialoside receptor sites on the cell membrane is a crucial part of the infection process. Hemagglutinins of influenza viruses (A, B, and C), mouse polyoma virus, mouse hepatitis virus, and some others have shown binding preference to specific sialic acid linkage [143–147]. Therefore, understanding the linkage type and distribution of sialic acid on cell surfaces is important for determining host tropism. Different sialic acid isomers on the tissues of various host animals are listed as following: α 2,3- and α 2,6-linkages on the lung and intestinal epithelial cells of chickens, α 2,3-linkages in the gut of waterfowl, α 2,3- and α 2,6-linkages on the respiratory epithelial cells of pig, and α 2,3-linkages

in the lower upper respiratory epithelium and α 2,6-linkages in the upper respiratory epithelium of humans [148–152].

Influenza virus is one of the most dangerous infectious pathogens, affecting wide range of hosts. As mentioned above, avian-adapted influenza virus preferentially binds α 2,3-sialic acid receptor, while human-adapted influenza virus has a binding preference for α 2,6-sialic acid receptor. Respiratory epithelium of pigs contain both α 2,3- and α 2,6-sialic acid receptor and are therefore regarded as a hypothetical “mixing” vessel where re-assortment of avian viruses and human viruses can take place, potentially leading to the emergence of pandemic influenza [45, 153]. In addition to α 2,6-structural requirement, long α 2,6-sialosides with an umbrella-like topology promote influenza transmission in humans, whereas short α 2,6-sialosides which adopt a cone-like topology hinder transmission [154]. Therefore, it is not a simple task to determine the authentic sialoside for the cognate receptor. The more information we gain from the hemagglutinin–sialoside interaction, the better our chances to develop novel anti-influenza drug or a vaccine to prevent this binding from occurring, thus preventing host infection.

4 Sialoside Array

Sialoside arrays are a new technology developed for high-throughput evaluation of interactions between sialosides and proteins, enzymes, antibodies, vertebrate cells, or viruses [155–162]. Although sialyl linkages are quite specific, the terminal sialic acid-linked carbohydrates are expanded with more internal sugar units in different backbone types, chain lengths, and branching patterns, generating a big library of sialosides. Advances in both chemical and enzymatic syntheses have provided reliable routes to the production of many complex sialosides as mentioned in Sect. 2. Prepared by immobilization of hundreds of sialosides to a surface in a discrete pattern, sialoside arrays minimize the amount of sialoside compounds as well as the quantity of binding samples, lowering cost for experiments and provide a high-throughput methodology for screening sialoside interactions. One of the other main advantages of sialoside array analysis is the multivalent display of arrayed sialosides, which can serve to mimic cell surface display; thus, array-based technique is particularly well suited for investigations of the real sialoside-mediated interactions in nature. Many aspects of sialoside binding can be addressed. For example, binding preferences can be determined, competition studies enable drug screening, and whole cell binding is of high diagnostic value. Therefore, sialoside arrays are expected to be of great value in contributing to pioneering new fields of sialobiology.

4.1 Applications of Sialoside Arrays in Influenza Virus Study

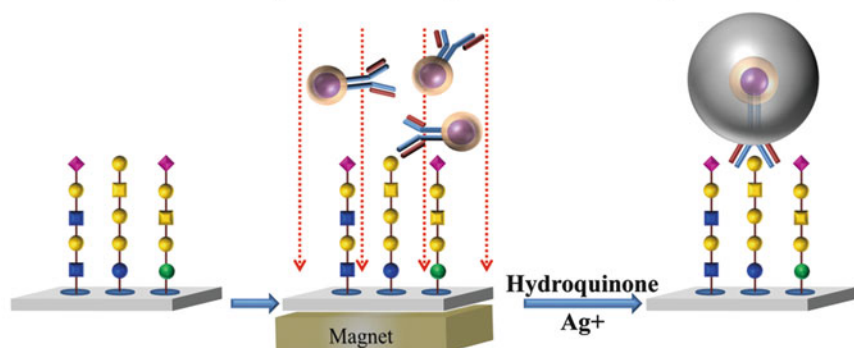
During the last decade, several important discoveries have significantly enriched our knowledge of the behavior of influenza viruses by applying sialoside array

technologies. Because each influenza virus has a characteristic receptor binding fingerprint, sialoside arrays have been used to identify influenza virus strains and evaluate possible hosts by monitoring changes in the receptor binding profile [161, 162]. In addition to host tropism of influenza virus, mutation and transmission of influenza viruses can also be identified by sialoside arrays [163]. For example, the sialoside binding specificity of wild type and mutant H1, H3, H5, H7, and H9 hemagglutinins are detected by sialoside arrays [161–167]. In the virus transmission study, the binding specificity between hemagglutinins and sialosides is related to transmission efficiency [168]. Switching the receptor binding specificity from α 2,6- to α 2,3- results in losing transmission efficiency in ferrets [169]. In addition, not only hemagglutinins but also neuraminidases play an essential role in the transmission of influenza viruses among mammals [170]. Furthermore, human influenza viruses have higher affinity for long than for short sialosides containing α 2,6-linkages, suggesting that the size and shape of sialic acid-containing receptors can modulate the receptor-binding properties of influenza A viruses [154]. Therefore, expanding the contents of sialoside arrays containing long α 2,6-sialosides is valuable for surveillance of the evolution of emerging viruses which could cause new pandemics or epidemics. As the accessibility of sialosides increases, sialoside array-based constructs are readily applicable to the large-scale identification and characterization of sialoside-related agents for pathological and biological study.

4.2 Applications of Sialoside Arrays in Search of Siglecs Inhibitors

Siglecs (Sialic acid-binding immunoglobulin-type lectins) are a family of cell surface receptor proteins which recognize sialosides. They are expressed on cells of the hemopoietic, immune, and nervous systems. Understanding sialoside ligands and siglec receptors precisely is important for mediating their biological functions. By applying sialoside arrays, the siglec family was demonstrated to exhibit differential specificity for the various sialoside sequences [155]. It was further revealed that sulfation is an important modulator of siglec–sialoside interactions [155]. For example, binding of the Siglec-2 to 6-sulfo-6'-sialyl-LacNAc, Siglec-8 to 6'-sulfo-sialyl Lewis^x, and mouse Siglec-F to 6'-sulfo-sialyl Lewis^x was all detected on a sialoside array [171]. In addition, the sialoside array also holds great potential for inhibitor screening of siglecs [155]. To date, several of the siglecs are emerging as potential targets for the treatment of diseases, such as Siglec-7 for chronic myeloid leukemia, Siglec-9 for hyper inflammation, and Siglec-8 and Siglec-10 for allergic disorders [172–176]. Siglec-8 cross-linking with antibodies can rapidly induce eosinophil apoptosis and then reduce numbers of eosinophils (and perhaps basophils and mast cells) in allergic and other diseases where these cells are important [177]. Therefore, the interest in the development of sialoside mimetics as inhibitors of sialic acid-recognizing proteins undoubtedly contributes to pharmaceutical

a Schematic Description of Nanoparticle Based Glycan Array



b

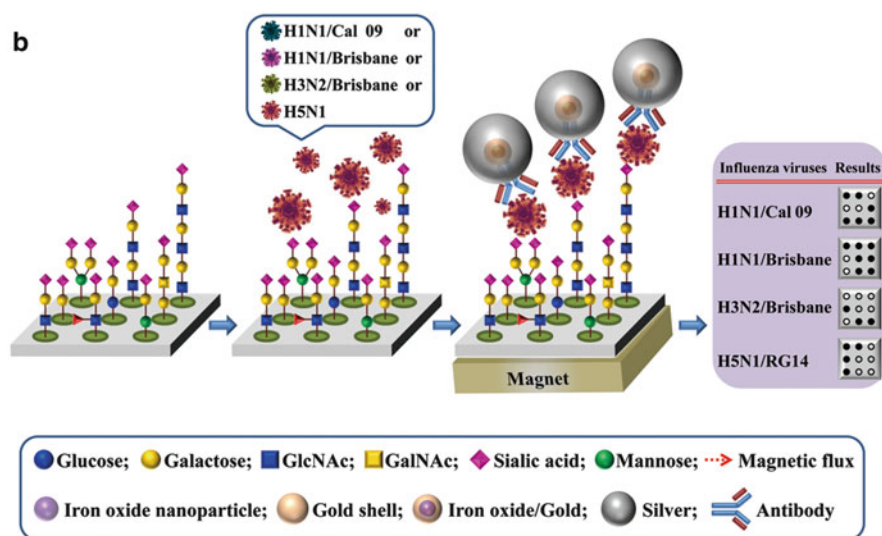


Fig. 2 Schematic description of functions of nanoparticle in sialoside array applications. (a) Core/shell nanoparticles contribute to improve the array sensitivity. (b) Influenza virus subtypes are observed by the naked eye

research. Sialoside arrays would also represent an invaluable tool to address therapeutic opportunities of sialic acid-mediated biological recognition effectively.

5 Synergy of Nanotechnology and Sialoside Arrays

Many biological systems are natural nanostructures, such as polysialic acids, gangliosides, glycoproteins, and viruses. Not surprisingly, nanoscale detections are emerging for sialoside array analysis. Nanotechnology provides opportunities

for developing new methods which contribute to improving the sensitivity and extending the present limits of molecular diagnostics. Synergy of nanotechnology and sialoside arrays is a new trend, dealing with the application of nanomaterials for performing array analyses. Recently, iron oxide/gold core/shell nanoparticles were applied in the detection of sugar binding proteins on carbohydrate arrays [178]. The core/shell nanoparticles consist of discrete domains of different materials and, thus, can exhibit the properties of different components in the same structure. This powerful combination enables researchers to concentrate proteins/viruses quickly by an external magnetic field, easily conjugate biomolecules on gold surfaces, and amplify signals by depositing silver on gold surfaces of core/shell nanoparticles (Fig. 2a). This nanoparticle-based assay can reach sub-attomole detection levels and has clear advantages in improving the sensitivity of carbohydrate array analysis. By the same concept, the utility of nanoparticle probes in sialoside array analysis for virus analysis is achievable. In Fig. 2b, a set of sialosides is selected from a library of sialosides to capture influenza viruses and form a unique fingerprint for each subtype of influenza virus. Captured viruses can be detected by modified gold nanoparticles. According to this result, the subtypes of influenza A viruses can be simply identified by using this approach; conversely, it is impossible for the commercially available quick test kits to achieve the same success. These quick test kits only allow for the detection of influenza type A and type B antigens, but not subtypes of influenza A. One of the most notable aspects about this nanoparticle-based detection method is that it requires nothing more than the naked eye to read out results which currently require chemical labeling and confocal laser scanners. The advantages of this method are less expense and no specialized training required. In summary, synergy of nanotechnology and sialoside arrays in the establishment of a comprehensive methodology to improve detection sensitivity of viruses with possible applications in the diagnosis of influenza subtype in a rapid, inexpensive, and efficient manner is introduced.

6 Conclusions

The advanced progress made thus far in this exciting field of sialoside chemistry and biology is only the beginning. As described in this chapter, it is quite apparent that sialosides play diverse and crucial roles in a wide variety of biological systems. Modern methods in chemical or enzymatic synthesis of sialosides have clearly made a major contribution toward improved production of many complex sialosides, especially with advancements in formation of α selective glycoside bonds. In addition, sialoside arrays have provided an efficient platform for quick identification of preferred ligands for sialic acid-binding proteins or pathogens. Combined with nanotechnology, a sialoside array which has provided for more reliable and sensitive methods in diagnostic applications, can reveal a great deal about the bind specifics of pathogen–sialic acid interactions with possible applications in the rapid diagnosis of influenza subtype. Research efforts can not only lead

to a better understanding of the pathological importance of sialic acids and their diversity but also could lead to the development of therapeutics. The increasing knowledge of sialosides in tumors suggests that further studies may assist both in determining their role in every step of tumor progression and in the design of new therapeutic and diagnostic approaches. Given what sialosides might do at different stages of tumor progression, targeting particular sialosides could be a way to achieve specificity in tumor detection. For example, blocking selectin interactions would affect platelet, lymphocyte, and endothelial adhesion. Moreover, multi-therapeutic strategies can be achieved by blocking selectin–sialyl Lewis^x interactions and inhibiting interactions between G_{M1} and growth-factor. Sialyl Lewis^x analogues therefore have potential in anti-inflammatory and antitumor therapy. To date, tumor research has led to the development of sialyl Tn-based vaccine for metastatic breast tumor. Pathogen research has investigated the use of polysialic acid in an anti-meningococcal type C vaccine. Moreover, sialyl glycosides are antiviral agent candidates against the influenza virus. Indeed, an understanding of the roles these sialosides play in the pathogenicity of infectious virus is essential to the development of these new-wave drugs. Predictably, the potential of the nascent sialoside array technologies in exploring the rich biological information content in sialosides and its use in diagnostic applications is enormous.

References

1. Traving C, Schauer R (1998) *Cell Mol Life Sci* 54:1330
2. Angata T, Varki A (2002) *Chem Rev* 102:439
3. Varki A (1993) *Glycobiology* 3:97
4. Schauer R (1982) *Cell biology monographs*, vol 10. Springer, Wien
5. Schauer R (2009) *Curr Opin Struct Biol* 19:507
6. Varki A (2008) *Trends Mol Med* 14:351
7. Varki A (2007) *Nature* 446:1023
8. Connor RJ, Kawaoka Y, Webster RG, Paulson JC (1994) *Virology* 205:17
9. Rogers GN, Dsouza BL (1989) *Virology* 173:317
10. Yu H, Huang SS, Chokhawala H, Sun MC, Zheng HJ, Chen X (2006) *Angew Chem Int Ed* 45:3938
11. Boons GJ, Demchenko AV (2000) *Chem Rev* 100:4539
12. Chen X, Varki A (2010) *ACS Chem Biol* 5:163
13. Klefel MJ, von Itzstein M (2002) *Chem Rev* 102:471
14. Ando H, Imanura A (2004) *Trends Glycosci Glycotechnol* 16:293
15. Boons GJ, Demchenko AV (2003) In: Wong C-H (ed) *Carbohydrate-based drug discovery*, vol 1. Wiley-VCH, Weinheim, 55 p
16. Halcomb RL, Chappel MD (2001) In: Wang PG, Bertozzi CR (eds) *Glycochemistry: principles, synthesis and applications*. Marcel Dekker, New York, 177 p
17. Hasegawa A, Kiso M (1997) In: Hanessian S (ed) *Preparative carbohydrate chemistry*. Marcel Dekker, New York, 357 p
18. Kiso M, Ishida H, Ito H (2000) In: Ernst B, Hart GW, Sinay P (eds) *Carbohydrates in chemistry and biology*, vol 1. Wiley-VCH, Weinheim, 345 p
19. Ress DK, Linhardt RJ (2004) *Curr Org Chem* 1:31

20. Roy R (1997) Glycoscience synthesis of substrate analogs and mimetics, vol 187. Springer, Berlin, 241 p
21. Deninno MP (1991) *Synthesis* 583
22. Ogawa T, Sugimoto M (1985) *Carbohydr Res* 135:C5
23. Okamoto K, Goto T (1990) *Tetrahedron* 46:5835
24. Kondo H, Ichikawa Y, Wong CH (1992) *J Am Chem Soc* 114:8748
25. Martin TJ, Brescello R, Toepfer A, Schmidt RR (1993) *Glycoconjugate J* 10:16
26. Martin TJ, Schmidt RR (1992) *Tetrahedron Lett* 33:6123
27. Sim MM, Kondo H, Wong CH (1993) *J Am Chem Soc* 115:2260
28. De Meo C, Parker O (2005) *Tetrahedron-Asymmetry* 16:303
29. Zhang ZY, Ollmann IR, Ye XS, Wischnat R, Baasov T, Wong CH (1999) *J Am Chem Soc* 121:734
30. Dziadek S, Brocke C, Kunz H (2004) *Chem Eur J* 10:4150
31. Marra A, Sinay P (1990) *Carbohydr Res* 195:303
32. Martichonok V, Whitesides GM (1996) *J Org Chem* 61:1702
33. Cai ST, Yu B (2003) *Org Lett* 5:3827
34. Liu YP, Ruan XH, Li XP, Li YX (2008) *J Org Chem* 73:4287
35. Tanaka SI, Goi T, Tanaka K, Fukase K (2007) *J Carbohydr Chem* 26, 369 p
36. Danishefsky SJ, Deninno MP, Chen S (1988) *J Am Chem Soc* 110:3929
37. Haberman JM, Gin DY (2001) *Org Lett* 3:1665
38. Haberman JM, Gin DY (2003) *Org Lett* 5:2539
39. Hanashima S, Akai S, Sato K (2008) *Tetrahedron Lett* 49:5111
40. Takahashi T, Tsukamoto H, Yamada H (1997) *Tetrahedron Lett* 38:8223
41. Castro-Palomino JC, Tsvetkov YE, Schmidt RR (1998) *J Am Chem Soc* 120:5434
42. Ercegovic T, Magnusson G (1995) *J Org Chem* 60:3378
43. Ercegovic T, Magnusson G (1996) *J Org Chem* 61:179
44. Hossain N, Magnusson G (1999) *Tetrahedron Lett* 40:2217
45. Ito T, Couceiro JNSS, Kelm S, Baum LG, Krauss S, Castrucci MR, Donatelli I, Kida H, Paulson JC, Webster RG, Kawaoka Y (1998) *J Virol* 72:7367
46. Ito Y, Ogawa T (1990) *Tetrahedron* 46:89
47. Martichonok V, Whitesides GM (1996) *J Am Chem Soc* 118:8187
48. Okamoto K, Kondo T, Goto T (1987) *Tetrahedron* 43:5909
49. Ito Y, Ogawa T (1988) *Tetrahedron Lett* 29:3987
50. De Meo C, Priyadarshani U (2008) *Carbohydr Res* 343:1540
51. Uchinashi Y, Nagasaki M, Zhou JZ, Tanaka K, Fukase K (2011) *Org Biomol Chem* 9:7243
52. Crich D, Li WJ (2007) *J Org Chem* 72:7794
53. Crich D, Li WJ (2007) *J Org Chem* 72:2387
54. Farris MD, De Meo C (2007) *Tetrahedron Lett* 48:1225
55. Harris BN, Patel PP, Gobble CP, Stark MJ, De Meo C (2011) *Eur J Org Chem* 4023
56. Hsu CH, Chu KC, Lin YS, Han JL, Peng YS, Ren CT, Wu CY, Wong CH (2010) *Chem Eur J* 16:1754
57. Liang FF, Chen L, Xing GW (2009) *Synlett* 425
58. Lin CC, Huang KT (2005) *Org Lett* 7:4169
59. Sun B, Srinivasan B, Huang XF (2008) *Chem Eur J* 14:7072
60. Tanaka H, Tatenno Y, Nishiura Y, Takahashi T (2008) *Org Lett* 10:5597
61. Tanaka K, Goi T, Fukase K (2005) *Synlett* 2958
62. Wang YJ, Jia J, Gu ZY, Liang FF, Li RC, Huang MH, Xu CS, Zhang JX, Men Y, Xing GW (2011) *Carbohydr Res* 346:1271
63. Yu CS, Niikura K, Lin CC, Wong CH (2001) *Angew Chem Int Ed* 40:2900
64. Paulsen H, Tietz H (1984) *Carbohydr Res* 125:47
65. Hasegawa A, Ohki H, Nagahama T, Ishida H, Kiso M (1991) *Carbohydr Res* 212:277
66. Tanaka H, Adachi M, Takahashi T (2005) *Chem Eur J* 11:849

67. Hasegawa A, Nagahama T, Ohki H, Hotta K, Ishida H, Kiso M (1991) *J Carbohydr Chem* 10:493
68. Demchenko AV, Boons GJ (1998) *Tetrahedron Lett* 39:3065
69. Ando H, Koike Y, Ishida H, Kiso M (2003) *Tetrahedron Lett* 44:6883
70. De Meo C, Demchenko AV, Boons GJ (2002) *Aust J Chem* 55:131
71. Tanaka H, Nishiura Y, Takahashi T (2006) *J Am Chem Soc* 128:7124
72. Chu KC, Ren CT, Lu CP, Hsu CH, Sun TH, Han JL, Pal B, Chao TA, Wu SH, Wong CH, Wu CY (2011) *Angew Chem Int Ed* 50:9391
73. Chokhawala HA, Huang SS, Lau K, Yu H, Cheng JS, Thon V, Hurtado-Ziola N, Guerrero JA, Varki A, Chen X (2008) *ACS Chem Biol* 3:567
74. Yu H, Chokhawala HA, Huang SS, Chen X (2006) *Nat Protocol* 1:2485
75. Yu H, Yu H, Karpel R, Chen X (2004) *Bioorg Med Chem* 12:6427
76. Yu H, Chokhawala H, Karpel R, Wu B, Zhang J, Zhang Y, Jia Q, Chen X (2005) *J Am Chem Soc* 127:17618
77. Yu H, Cheng J, Ding L, Khedri Z, Chen Y, Chin S, Lau K, Tiwari VK, Chen X (2009) *J Am Chem Soc* 131:18467
78. Cheng JS, Yu H, Lau K, Huang SS, Chokhawala HA, Li YH, Tiwari VK, Chen X (2008) *Glycobiology* 18:686
79. Muthana S, Cao H, Chen X (2009) *Curr Opin Chem Biol* 13:573
80. Chokhawala HA, Cao HZ, Yu H, Chen X (2007) *J Am Chem Soc* 129:10630
81. Sun X-L, Kanie Y, Guo C-T, Kanie O, Suzuki Y, Wong C-H (2000) *Eur J Org Chem* 2643
82. Guo C-T, Sun X-L, Kanie O, Shortridge KF, Suzuki T, Miyamoto D, Hidari KI, Wong C-H, Suzuki Y (2002) *Glycobiology* 12:183
83. Morley TJ, Withers SG (2010) *J Am Chem Soc* 132:9430
84. Lau K, Thon V, Yu H, Ding L, Chen Y, Muthana MM, Wong D, Huang R, Chen X (2010) *Chem Commun* 46:6066
85. Aharoni A, Thieme K, Chiu CPC, Buchini S, Lairson LL, Chen HM, Strynadka NCJ, Wakarchuk WW, Withers SG (2006) *Nat Methods* 3:609
86. Namdjou DJ, Chen HM, Vinogradov E, Brochu D, Withers SG, Wakarchuk WW (2008) *ChemBiochem* 9:1632
87. Fuster MM, Esko JD (2005) *Nat Rev Cancer* 5:526
88. Hakomori S (1996) *Cancer Res* 56:5309
89. Dube DH, Bertozzi CR (2005) *Nat Rev Drug Discov* 4:477
90. Sun XL (2007) *Curr Med Chem* 14:2304
91. Kaida K, Ariga T, Yu RK (2009) *Glycobiology* 19:676
92. Sandhoff K (1969) *FEBS Lett* 4:351
93. Gillard BK, Thomas JW, Nell LJ, Marcus DM (1989) *J Immunol* 142:3826
94. Matsuzaki K, Kato K, Yanagisawa K (2010) *BBA Mol Cell Biol Lipids* 1801:868
95. Desplats PA, Denny CA, Kass KE, Gilmartin T, Head SR, Sutcliffe JG, Seyffied TN, Thomas EA (2007) *Neurobiol Dis* 27:265
96. Liu YH, Li RX, Ladisch S (2004) *J Biol Chem* 279:36481
97. Li RX, Manela J, Kong Y, Ladisch S (2000) *J Biol Chem* 275:34213
98. Li RX, Liu YH, Ladisch S (2001) *J Biol Chem* 276:42782
99. McKallip R, Li RX, Ladisch S (1999) *J Immunol* 163:3718
100. Ladisch S, Gillard B, Wong C, Ulsh L (1983) *Cancer Res* 43:3808
101. Caldwell S, Heitger A, Shen WP, Liu YH, Taylor B, Ladisch S (2003) *J Immunol* 171:1676
102. Ladisch S, Kitada S, Hays EF (1987) *J Clin Invest* 79:1879
103. Koochekpour S, Pilkington GJ (1996) *Cancer Lett* 104:97
104. Hakomori SI (1990) *J Biol Chem* 265:18713
105. Nagy P, Vereb G, Sebestyen Z, Horvath G, Lockett SJ, Damjanovich S, Park JW, Jovin TM, Szollosi J (2002) *J Cell Sci* 115:4251
106. Zhang HT, Berezov A, Wang Q, Zhang G, Drebin J, Murali R, Greene MI (2007) *J Clin Invest* 117:2051

107. Valentino L, Moss T, Olson E, Wang HJ, Elashoff R, Ladisch S (1990) *Blood* 75:1564
108. Hakomori S (2001) In: Wu AM (ed) *Molecular immunology of complex carbohydrates-2*, vol 491. Plenum, New York, 369 p
109. Ladisch S, Wu ZL, Feig S, Ulsh L, Schwartz E, Floutsis G, Wiley F, Lenarsky C, Seeger R (1987) *Int J Cancer* 39:73
110. Marquina G, Waki H, Fernandez LE, Kon K, Carr A, Valiente O, Perez R, Ando S (1996) *Cancer Res* 56:5165
111. Mirkin BL, Clark SH, Zhang C (2002) *Cell Prolif* 35:105
112. Ogawa H, Inoue M, Tanizawa O, Miyamoto M, Sakurai M (1992) *Histochemistry* 97:311
113. Butenhof KJ, Gerken TA (1993) *Biochemistry* 32:2650
114. Shogren R, Gerken TA, Jentoft N (1989) *Biochemistry* 28:5525
115. Julien S, Lagadec C, Krzewinski-Recchi MA, Courtand G, Le Bourhis X, Delannoy P (2005) *Breast Cancer Res Treat* 90:77
116. Imada T, Rino Y, Hatori S, Takahashi M, Amano T, Kondo J, Suda T (1999) *Hepatogastroenterology* 46:208
117. Leivonen M, Nordling S, Lundin J, von Boguslawski K, Haglund C (2001) *Oncology* 61:299
118. Nakagoe T, Sawai T, Tsuji T, Jibiki MA, Nanashima A, Yamaguchi H, Yasutake T, Ayabe H, Arisawa K, Ishikawa H (2002) *Anticancer Res* 22:451
119. MacLean GD, Reddish MA, Koganty RR, Longenecker BM (1996) *J Immunother* 19:59
120. Ragupathi G, Koganty RR, Qiu DX, Lloyd KO, Livingston PO (1998) *Glycoconjugate J* 15:217
121. Sandmaier BM, Oparin DV, Holmberg LA, Reddish MA, MacLean GD, Longenecker BM (1999) *J Immunother* 22:54
122. Biancone L, Araki M, Araki K, Vassalli P, Stamenkovic I (1996) *J Exp Med* 183:581
123. Dennis JW, Laferte S (1987) *Cancer Metastasis Rev* 5:185
124. Nakamori S, Kameyama M, Imaoka S, Furukawa H, Ishikawa O, Sasaki Y, Kabuto T, Iwanaga T, Matsushita Y, Irimura T (1993) *Cancer Res* 53:3632
125. Lasky LA (1995) *Annu Rev Biochem* 64:113
126. Barthel SR, Gavino JD, Descheny L, Dimitroff CJ (2007) *Expert Opin Ther Targets* 11:1473
127. Varki A (1994) *Proc Natl Acad Sci USA* 91:7390
128. Thomas SN, Schnaar RL, Konstantopoulos K (2009) *Am J Physiol Cell Physiol* 296:C505
129. Takada A, Ohmori K, Yoneda T, Tsuyuoka K, Hasegawa A, Kiso M, Kannagi R (1993) *Cancer Res* 53:354
130. Sozzani P, Arisio R, Porpiglia M, Benedetto C (2008) *Int J Surg Pathol* 16:365
131. Ugorski M, Laskowska A (2002) *Acta Biochim Pol* 49:303
132. Ito K, Ye CL, Hibi K, Mitsuoka C, Kannagi R, Hidemura K, Ando H, Kasai Y, Akiyama S, Nakao A (2001) *J Gastroenterol* 36:823
133. Ley K (2003) *Trends Mol Med* 9:263
134. Somers WS, Tang J, Shaw GD, Camphausen RT (2000) *Cell* 103:467
135. Rohr TE, Troy FA (1980) *J Biol Chem* 255:2332
136. Troy FA (1992) *Glycobiology* 2:5
137. Glode MP, Robbins JB, Liu TY, Gotschlich EC, Orskov I, Orskov F (1977) *J Infect Dis* 135:94
138. McGowen MM, Vionnet J, Vann WF (2001) *Glycobiology* 11:613
139. Liu TY, Gotschli E, Dunne FT, Jonssen EK (1971) *J Biol Chem* 246:4703
140. Costantino P, Viti S, Podda A, Velmonte MA, Nencioni L, Rappuoli R (1992) *Vaccine* 10:691
141. Finne J, Leinonen M, Makela PH (1983) *Lancet* 2:355
142. Jennings HJ, Roy R, Gamian A (1986) *J Immunol* 137:1708
143. Wilson IA, Skehel JJ, Wiley DC (1981) *Nature* 289:366
144. Crennell S, Takimoto T, Portner A, Taylor G (2000) *Nat Struct Biol* 7:1068
145. Rosenthal PB, Zhang XD, Formanowski F, Fitz W, Wong CH, Meier-Ewert H, Skehel JJ, Wiley DC (1998) *Nature* 396:92

146. Sauter NK, Glick GD, Crowther RL, Park SJ, Eisen MB, Skehel JJ, Knowles JR, Wiley DC (1992) *Proc Natl Acad Sci USA* 89:324
147. Stehle T, Yan YW, Benjamin TL, Harrison SC (1994) *Nature* 369:160
148. Baum LG, Paulson JC (1990) *Acta Histochem* 40:35
149. Wan HQ, Perez DR (2006) *Virology* 346:278
150. Shinya K, Ebina M, Yamada S, Ono M, Kasai N, Kawaoka Y (2006) *Nature* 440:435
151. Guo CT, Takahashi N, Yagi H, Kato K, Takahashi T, Yi SQ, Chen Y, Ito T, Otsuki K, Kida H, Kawaoka Y, Hidari KIPJ, Miyamoto D, Suzuki T, Suzuki Y (2007) *Glycobiology* 17:713
152. Gagneux P, Cheriyan M, Hurtado-Ziola N, van der Linden ECMB, Anderson D, McClure H, Varki A, Varki NM (2003) *J Biol Chem* 278:48245
153. Oxford JS, Lambkin R, Sefton A, Daniels R, Elliot A, Brown R, Gill D (2005) *Vaccine* 23:940
154. Chandrasekaran A, Srinivasan A, Raman R, Viswanathan K, Raguram S, Tumpey TM, Sasisekharan V, Sasisekharan R (2008) *Nat Biotechnol* 26:107
155. Blixt O, Han SF, Liao L, Zeng Y, Hoffmann J, Futakawa S, Paulson JC (2008) *J Am Chem Soc* 130:6680
156. Blixt O, Allin K, Bohorov O, Liu XF, Andersson-Sand H, Hoffmann J, Razi N (2008) *Glycoconjugate J* 25:59
157. Blixt O, Head S, Mondala T, Scanlan C, Huflejt ME, Alvarez R, Bryan MC, Fazio F, Calarese D, Stevens J, Razi N, Stevens DJ, Skehel JJ, van Die I, Burton DR, Wilson IA, Cummings R, Bovin N, Wong CH, Paulson JC (2004) *Proc Natl Acad Sci USA* 101:17033
158. Gout E, Garlatti V, Smith DF, Lacroix M, Dumestre-Perard C, Lunardi T, Martin L, Cesbron JY, Arlaud GJ, Gaboriaud C, Thielens NM (2010) *J Biol Chem* 285:6612
159. Liang CH, Wu CY (2009) *Exp Rev Proteomics* 6:631
160. Rillahan CD, Paulson JC (2011) *Annu Rev Biochem* 80:797
161. Stevens J, Blixt O, Glaser L, Taubenberger JK, Palese P, Paulson JC, Wilson IA (2006) *J Mol Biol* 355:1143
162. Stevens J, Blixt O, Paulson JC, Wilson IA (2006) *Nat Rev Microbiol* 4:857
163. Liao HY, Hsu CH, Wang SC, Liang CH, Yen HY, Su CY, Chen CH, Jan JT, Ren CT, Chen CH, Cheng TJR, Wu CY, Wong CH (2010) *J Am Chem Soc* 132:14849
164. Suzuki Y (2005) *Biol Pharm Bull* 28:399
165. Stevens J, Blixt O, Chen LM, Donis RO, Paulson JC, Wilson IA (2008) *J Mol Biol* 381:1382
166. Childs RA, Palma AS, Wharton S, Matrosovich T, Liu Y, Chai WG, Campanero-Rhodes MA, Zhang YB, Eickmann M, Kiso M, Hay A, Matrosovich M, Feizi T (2009) *Nat Biotechnol* 27:797
167. Wang CC, Chen JR, Tseng YC, Hsu CH, Hung YF, Chen SW, Chen CM, Khoo KH, Cheng TJ, Cheng YSE, Jan JT, Wu CY, Ma C, Wong CH (2009) *Proc Natl Acad Sci USA* 106:18137
168. Maines TR, Jayaraman A, Belser JA, Wadford DA, Pappas C, Zeng H, Gustin KM, Pearce MB, Viswanathan K, Shriver ZH, Raman R, Cox NJ, Sasisekharan R, Katz JM, Tumpey TM (2009) *Science* 325:484
169. Tumpey TM, Maines TR, Van Hoeven N, Glaser L, Solorzano A, Pappas C, Cox NJ, Swayne DE, Palese P, Katz JM, Garcia-Sastre A (2007) *Science* 315:655
170. Yen HL, Liang CH, Wu CY, Forrest HL, Ferguson A, Choy KT, Jones J, Wong DDY, Cheung PPH, Hsu CH, Li OT, Yuen KM, Chan RWY, Poon LLM, Chan MCW, Nicholls JM, Krauss S, Wong CH, Guan Y, Webster RG, Webby RJ, Peiris M (2011) *Proc Natl Acad Sci USA* 108:14264
171. Bochner BS, Alvarez RA, Mehta P, Bovin NV, Blixt O, White JR, Schnaar RL (2005) *J Biol Chem* 280:4307
172. Crocker PR, Paulson JC, Varki A (2007) *Nat Rev Immunol* 7:255
173. von Gunten S, Yousefi S, Seitz M, Jakob SM, Schaffner T, Seger R, Takala J, Villiger PM, Simon HU (2005) *Blood* 106:1423
174. Yokoi H, Myers A, Matsumoto K, Crocker PR, Saito H, Bochner BS (2006) *Allergy* 61:769
175. O'Reilly MK, Paulson JC (2009) *Trends Pharmacol Sci* 30:240

176. Vitale C, Romagnani C, Puccetti A, Olive D, Costello R, Chiossone L, Pitto A, Bacigalupo A, Moretta L, Mingari MC (2001) *Proc Natl Acad Sci USA* 98:5764
177. Nutku E, Aizawa H, Hudson SA, Bochner BS (2003) *Blood* 101:5014
178. Liang CH, Wang CC, Lin YC, Chen CH, Wong CH, Wu CY (2009) *Anal Chem* 81:7750

SIGLEC-4 (MAG) Antagonists: From the Natural Carbohydrate Epitope to Glycomimetics

Oliver Schwardt, Soerge Kelm, and Beat Ernst

Abstract Siglec-4, also known as myelin-associated glycoprotein (MAG), is a member of the siglec (sialic acid-binding immunoglobulin-like lectins) family. MAG binds with high preference to sialic acids $\alpha(2-3)$ -linked to D-galactose. Although the involvement and relevance of its sialic acid binding activity is still controversial, it could be demonstrated that interactions of MAG with sialylated gangliosides play an important role in axon stability and regeneration. In this article we describe in detail our current understanding of the biological role and the carbohydrate specificity of siglec-4. Furthermore, this review compiles the intensive research efforts leading from the identification of the minimal oligosaccharide binding epitope in gangliosides via micromolar oligosaccharide mimics to the development of small molecular weight and more drug-like sialic acid derivatives binding with low nanomolar affinities. Such compounds will be useful to elucidate MAG's biological functions, which are currently not fully understood.

Keywords Binding kinetics · Drug discovery · Lead optimization · Myelin-associated glycoprotein (MAG) · Oligosaccharide mimics · Sialosides · Siglecs

Contents

1	Introduction	152
1.1	History of MAG Research	152
1.2	MAG as Siglec	153
2	Biological Functions of MAG	154
3	Oligosaccharides and Mimics as MAG Antagonists	156

O. Schwardt and B. Ernst (✉)
Institute of Molecular Pharmacy, Pharmacenter, University of Basel,
Klingelbergstrasse 50, CH-4056, Basel, Switzerland
e-mail: beat.ernst@unibas.ch

S. Kelm
Institute for Physiological Biochemistry, University Bremen, D-28334, Bremen, Germany

4	Neuraminic Acid Derivatives as MAG Antagonists – First Generation	172
4.1	Modification in the 2-Position	173
4.2	Modification in the 5-Position	173
4.3	Modification in the 9-Position	173
5	Neuraminic Acid Derivatives as MAG Antagonists – Second Generation	181
5.1	Modification in the 2-Position	182
5.2	Modification in the 5-Position	182
5.3	Modification in the 4-Position	185
6	Design of a Non-Carbohydrate Antagonist of MAG	185
7	Kinetic Studies	187
8	High-Affinity MAG Antagonists by a <i>Fragment-Based In Situ Combinatorial Approach</i>	190
9	Summary of the Structure Affinity Relationship	192
9.1	The Terminal Neuraminic Acid	193
9.2	The Subterminal Galactose	195
9.3	The Third Terminal Monosaccharide	195
9.4	The Internal $\alpha(2-6)$ -Linked Neuraminic Acid	196
10	Conclusions	196
	References	197

1 Introduction

Myelin-associated glycoprotein (abbreviated as MAG) is a member of the siglec (sialic acid-binding immunoglobulin-like lectins) family with at least 16 members [1, 2]. Therefore, MAG is also known under the name Siglec-4 [3]. In this chapter we will focus on the specificity of MAG for sialylated carbohydrate structures and the development of sialic acid derivatives binding with low nanomolar affinity. Such compounds will be useful to investigate its biological function, which has remained largely unclear despite intensive research over 3 decades as documented by about 2,300 publications listed in PubMed with almost 100,000 citations.

1.1 History of MAG Research

Thirty years ago, Quarles et al. described a heavily glycosylated protein from myelin preparations [4], which they called myelin-associated glycoprotein [5], abbreviated MAG. Over the following 10 years, Quarles and his coworkers characterized this protein before it drew further attention from other laboratories, induced by the seminal observation that MAG is developmentally regulated in myelinating cells [6] and is possibly involved in multiple sclerosis [7], two publications, which together have been cited over 500 times.

The next decade in MAG research was fostered largely by the availability of polyclonal and monoclonal antibodies and upcoming molecular biology tools, including the cloning of MAG mRNAs with a publication frequency of 20–30 papers each year. Clearly the most prominent highlights of that time were the characterization of MAG as a cell adhesion molecule (CAM) carrying the HNK-1 epitope [8–11], altogether over 1,700 citations, the molecular cloning of cDNAs

demonstrating the existence of alternatively spliced variants having two different cytoplasmic tails [12, 13], and the structural similarity to members of the immunoglobulin gene superfamily (IgSF) [14].

In 1994, four different major observations marked a turning point regarding the biological functions of MAG. First, to the great surprise of the “MAG-community,” MAG-deficient mice created by gene inactivation in two independent laboratories showed almost normal myelination [15, 16], despite the general expectation that MAG would play an essential role in myelination of axons, based on its distribution during development and its properties as CAM. Second, tyrosine residues in the cytoplasmic tail of L-MAG are phosphorylated by the tyrosine kinase *fyn*, suggesting that this larger splice variant of MAG mainly found in the central nervous system is involved in signal transduction into oligodendrocytes [17]. Third, sequence similarities of MAG with CD22, a B cell specific cell adhesion protein, and sialoadhesin, a sialic acid-recognizing cell adhesion molecules found on tissue macrophages, led to the discovery that this subgroup of IgSF bind sialylated glycoconjugates, providing a clue for the specificity of MAG as a CAM [18] and initiating the discovery the siglec family. Fourth, two independent laboratories investigating the effect of myelin preparations [19] or MAG-expressing fibroblasts [20] on neuronal outgrowth made the unexpected discovery that MAG strongly inhibits the outgrowth of neuronal axons, suggesting that MAG is a key player in the inhibitory effect of CNS myelin on neurite outgrowth, an important factor preventing the regeneration of injured nerves. These four observations had a significant impact on the field as shown by the almost 2,500 citations of these six publications and the increase of publications dealing with MAG to about 100 every year since then. Only the aspect of signal transduction via tyrosine phosphorylation into the oligodendrocytes has not been investigated further to a significant extent, although approximately 250 citations of the relevant publication [17] are an indicator of the interest it has raised. Clearly, since 1994 most attention has been directed towards the inhibitory activity of MAG on neuronal outgrowth. The hope to develop treatments leading to the regeneration of injured nerves had led to an increase in the research on MAG and other inhibitory molecules, their interaction, and the potential signaling pathways in neurons. Along that line, the carbohydrate specificity of MAG has also been studied and sialic acid-based high affinity inhibitors of MAG have been developed as described in more detail below. More recently, the interest in other functions of MAG as in stabilizing and protecting myelinated axons has also gained some momentum [21].

1.2 *MAG as Siglec*

Native MAG or Siglec-4 has been characterized as a heavily glycosylated protein of about 100 kDa with 30% of its mass being made up by carbohydrates [22]. As shown in Fig. 1, MAG has five Ig-like domains, one N-terminal V-set domain, four C2-set domains, and eight *N*-glycosylation sites, all of which are used [24]. Differences in the glycan structures lead to microheterogeneity, causing native MAG to migrate in

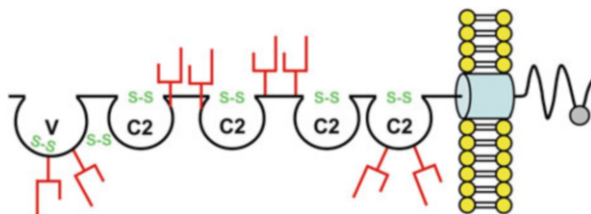


Fig. 1 Schematic drawing of the domain structure of MAG; shown are its five extracellular domains, the disulfide bridges (*green*) and eight *N*-glycosylation sites (*red*). At the cytoplasmic tail the location of the tyrosine phosphorylation site of L-MAG is indicated by a *gray circle* (adapted from [23])

SDS-PAGE as a broad smear. As mentioned above, MAG is expressed in two splice variant forms which have different cytoplasmic tails. Based on the amino acid sequence, S-MAG is 67 kDa and L-MAG is 72 kDa and contains a tyrosine phosphorylation site [12, 13].

MAG is a typical siglec with the sialic acid binding site located in the N-terminal V-set domain. Characteristic features of the binding site are the amino acids Arg118, Trp22, and Tyr124 interacting with the carboxylate, the *N*-acetyl substituent at C5, and the glycerol side chain of sialic acids (see below for details).

Siglec-4 is the best conserved protein among all siglecs, supporting the notion that the function of MAG is essential in all vertebrates, including the specificity for sialylated glycans [25]. Between fish Siglecs-4 and the mammalian orthologs, 38% of all amino acids are identical and in the first two N-terminal domains the identity is even higher (Fig. 2).

Siglec-4 binds with high preference to sialic acids $\alpha(2-3)$ -linked to D -galactose. The underlying structures can further enhance binding as discussed in detail below [27]. Similar to Siglec-1 (sialoadhesin), and in contrast to Siglec-2 (CD22), MAG does not bind well to *N*-glycolylneuraminic acid (Neu5Gc) [28, 29]. At least in vitro Siglec-4 binds well to both, glycoproteins and glycolipids [18, 30–34].

2 Biological Functions of MAG

Despite 30 years of research, the biological functions of MAG are still unclear and have been a matter of intense controversy, in particular the relevance for neurite outgrowth inhibition. However, several of the results appearing to be contradictory at first sight can be combined in a model introducing a network of interaction partners of MAG involving both, sialic acid-dependent and sialic acid-independent binding events. These partners include gangliosides like GT1b and GD1a, Nogo receptors (NgR1 and NgR2), neurotrophin receptor p75, Lingo-1, TROY, and pirB (paired-immunoglobulin-like-receptor-B). For some of these (e.g., NgR1, p75, and Lingo-1) the formation of signaling receptor complexes triggering RhoA activity

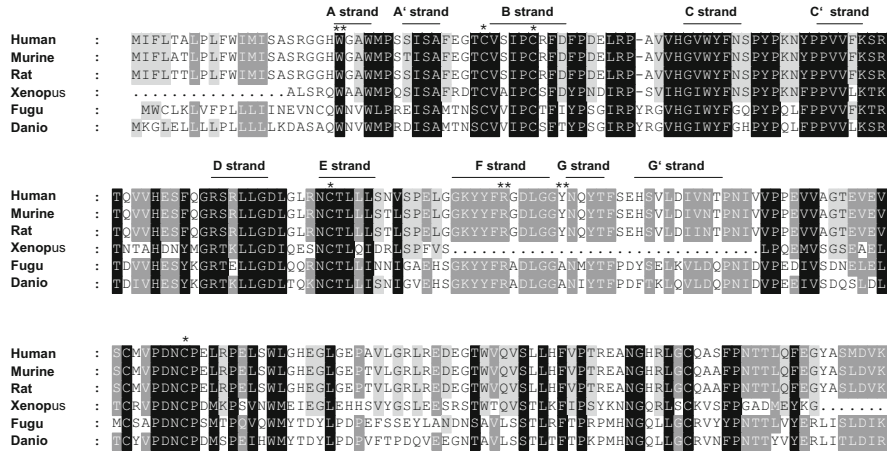


Fig. 2 Alignment of the two N-terminal domains of vertebrate MAG. Amino acid alignment of Siglec-4 domains 1 and 2 from three mammalian species with bird (Quail), frog (Xenopus) and two fish (Fugu and Danio [25]) species. Sequences have been aligned using Geneious 5.5.6 based on the ClustalW algorithm using the Blossum matrix with a gap opening penalty of 10 and gap extension penalty of 0.1. The strand assignment in domain 1 is based on a homology model of the crystal structure of Siglec-1 [26]. Amino acid residues known to be essential for the siglec-characteristic disulfide bridges and Sia binding are indicated with one or two asterisks, respectively

dependent inhibition of axon regeneration has been proposed. However, even after knocking out three potential contributors (MAG, NgR1, and OMgp) the results and their interpretation are still controversial [35, 36]. Such scenarios have been reviewed intensively [21, 35–39] and, therefore, will not be discussed here.

Although the initial studies on MAG^{-/-} mice revealed almost normal myelination [15, 16], later studies demonstrated that with increasing age these mice develop quite severe pathological abnormalities associated with myelin instability [40, 41]. More recently, the roles of MAG in protecting and stabilizing the myelinated axon have been addressed in several studies. For example, good evidence has been provided that sialylated gangliosides play important roles in the process by which MAG protects neurons from acute toxicity [42], but not in the protection from excitotoxicity, a process in which the interaction of MAG with NgR is involved [43].

The involvement and relevance of MAG’s sialic acid binding activity is still controversial, one reason being the difficulty in addressing this interaction directly without disturbing other parts of the system. Until now, the main tools have been genetically modified animals in which one or more of the binding partners had been knocked out. In a different approach, sialidase has been used to demonstrate the relevance of MAG interactions with gangliosides in axon stability and regeneration [44]. A more direct and specific tool to address selectively the sialic acid binding site of MAG would be to use small molecular weight inhibitors, which bind with

high affinity and selectivity to MAG blocking its binding site. Our current understanding of the carbohydrate specificity and development of such inhibitors is described in detail in the next sections.

3 Oligosaccharides and Mimics as MAG Antagonists

Based on the observation that gangliosides carry 75–80% of the sialic acid moieties in the brain, and the major brain gangliosides GD1a (**1**) and GT1b (**3**) (Fig. 1) contain the preferred Neu5Ac- α (2–3)-Gal- β (1–3)-GalNAc epitope for MAG [45], it was hypothesized that brain gangliosides located on the surface of nerve cells are functional ligands responsible for MAG-mediated inhibition of neural regeneration [30, 34, 46].

This hypothesis is supported by a considerable amount of experimental evidence: First, MAG binds with high affinity and specificity to GD1a, GT1b, and similar gangliosides [30, 31, 34, 46–48]. Second, the binding of MAG with gangliosides is blocked by the mAb 513, a conformation-dependent specific anti-MAG antibody that also blocks MAG-neuron binding [31]. Finally, mice genetically lacking the Neu5Ac- α (2–3)-Gal- β (1–3)-GalNAc epitope on gangliosides (but not on glycoproteins) suffer from axon degeneration and dysmyelination similar to that in MAG knockout mice [49].

In order to clarify the binding specificities of MAG and other siglecs [Schwann cell myelin protein (SMP, Siglec-4b) and sialoadhesin (Sn, Siglec-1)], Schnaar et al. [30, 31, 47, 48] measured the siglec-mediated half-maximal cell adhesion to several immobilized natural and synthetically modified [50–55] gangliosides (Table 1).

Apart from the two major brain gangliosides GD1a (**1**) and GT1b (**3**) (entries 9 and 10), which adhere to MAG with moderate potency, other gangliosides like GM4 (**5**) and GM3 (**6**) (entries 1 and 2) also support adhesion of MAG, albeit at more than tenfold higher concentrations. In contrast, gangliosides missing the terminal α (2–3)-linked sialic acid such as GD3 (**7**), GM1 (**8**), and GM1 α (**9**) (entries 3–5) do not bind to MAG. In an effort to identify the relevant functional groups of the terminal α (2–3)-linked sialic acid required for MAG binding, chemically modified derivatives of GM3 and GD1a were tested [31, 48]. However, all the modifications, including 4-, 7-, 8-, or 9-deoxy derivatives, 1-alcohols as well as several esters or amides, resulted in a decrease or even loss of affinity for MAG.

GM1 α (**9**), GD1 α (**13**), GT1 α (**2**), and GQ1 β (**4**) (Fig. 3) are members of a less abundant family of brain gangliosides. Members of this family are called Chol-1 based on their reactivity towards a polyclonal antiserum raised against cholinergic neurons [56]. They bear a Neu5Ac- α (2–6)-GalNAc residue on the gangliotetraose core, making them part of the “ α -series” ganglioside family [57]. In contrast to GM1 α (**9**), the other members of the α -series displayed enhanced avidity for MAG. A comparison of MAG-binding to the Chol-1 gangliosides GD1 α (**13**), GT1 α (**2**), and GQ1 β (**4**) (entries 11–13), and the closely related gangliosides, GM1b

Table 1 MAG-mediated cell adhesion to natural and modified gangliosides [30, 31, 47, 48]

Entry	Ganglioside	Conc. of ganglioside supporting half-maximal cell adhesion		Ref.
		[pmol/well]	[μ mol/L]	
1	Neu5Ac- α (2-3)-Gal- β -Cer GM4 (5)	>100 ^a	>0.5	[48]
2	Neu5Ac- α (2-3)-Gal- β (1-4)-Glc- β -Cer GM3 (6)	>100 ^a	>0.5	[47, 48]
3	Neu5Ac- α (2-8)-Neu5Ac- α (2-3)-Gal- β (1-4)-Glc- β -Cer GD3 (7)	n.d. ^b	–	[47, 48]
4	Gal- β (1-3)-GalNAc- β (1-4)-Gal- β (1-4)-Glc- β -Cer Neu5Ac- α (2-3) GM1 (8)	n.d. ^b	–	[30, 47]
5	Neu5Ac- α (2-6) Gal- β (1-3)-GalNAc- β (1-4)-Gal- β (1-4)-Glc- β -Cer GM1 α (9)	n.d. ^b	–	[30, 47]
6	Neu5Ac- α (2-3)-Gal- β (1-3)-GalNAc- β (1-4)-Gal- β (1-4)-Glc- β -Cer GM1b (10)	80	0.4	[30]
7	Neu5Ac- α (2-3)-Gal- β (1-3)-GlcNAc- β (1-4)-Gal- β (1-4)-Glc- β -Cer (11)	240	1.2	[30]
8	Neu5Ac- α (2-3)-Gal- β (1-6)-GalNAc- β (1-4)-Gal- β (1-4)-Glc- β -Cer (12)	87	0.435	[30]
9	Neu5Ac- α (2-3)-Gal- β (1-3)-GalNAc- β (1-4)-Gal- β (1-4)-Glc- β -Cer Neu5Ac- α (2-3) GD1a (1)	50	0.25	[30, 47]
10	Neu5Ac- α (2-3)-Gal- β (1-3)-GalNAc- β (1-4)-Gal- β (1-4)-Glc- β -Cer Neu5Ac- α (2-8)-Neu5Ac- α (2-3) GT1b (3)	50	0.25	[30, 47]
11	Neu5Ac- α (2-3)-Gal- β (1-3)-GalNAc- β (1-4)-Gal- β (1-4)-Glc- β -Cer Neu5Ac- α (2-6) GD1 α (13)	19	0.095	[30]
12	Neu5Ac- α (2-3)-Gal- β (1-3)-GalNAc- β (1-4)-Gal- β (1-4)-Glc- β -Cer Neu5Ac- α (2-3) GT1a α (2)	17	0.085	[30]
13	Neu5Ac- α (2-3)-Gal- β (1-3)-GalNAc- β (1-4)-Gal- β (1-4)-Glc- β -Cer Neu5Ac- α (2-6) GQ1b α (4)	6.0	0.03	[30, 47]
14	Neu5Ac- α (2-3)-Gal- β (1-3)-GalNAc- β (1-4)-Gal- β (1-4)-Glc- β -Cer HO ₃ S-(6) (14)	22	0.11	[30]
15	Neu5Ac- α (2-3)-Gal- β (1-3)-GalNAc- β (1-4)-Gal- β (1-4)-Glc- β -Cer HO ₃ S-(6) HO ₃ S-(3) (15)	1.5	0.0075	[30]

^aLow but statistically significant adhesion over background^bNo detectable adhesion at >100 pmol/well

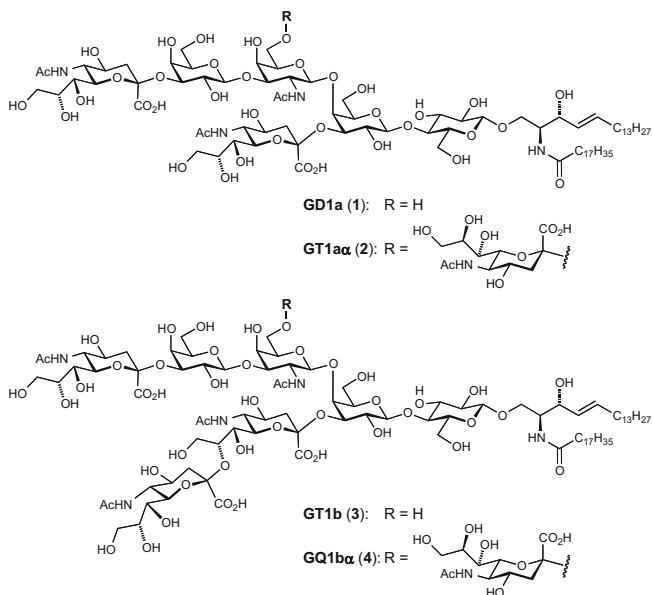
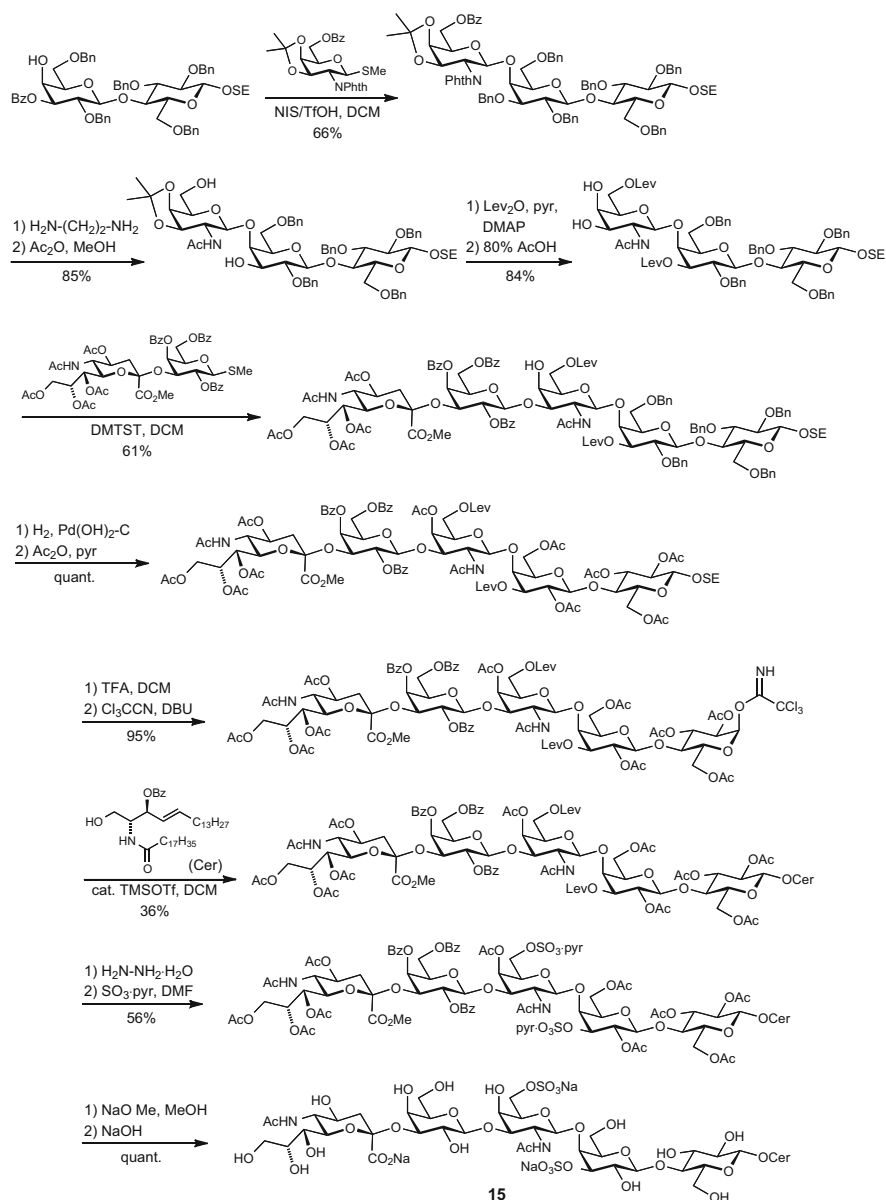


Fig. 3 The brain gangliosides GD1a (**1**), GT1b (**2**), GT1a α (**3**), and GQ1b α (**4**)

(**10**), GD1a (**1**), and GT1b (**3**) (entries 6, 9, and 10), lacking the $\alpha(2-6)$ -Neu5Ac residue, reveals that an $\alpha(2-6)$ -linked sialic acid leads to a three- to tenfold enhanced adhesion. GQ1b α (**4**, Fig. 3, Table 1, entry 13) represents the most potent natural ligand of MAG discovered to date. Further GM1b-based structure-activity studies, which addressed alterations in the gangliotetraose core, e.g., GlcNAc replacing GalNAc (entry 7) or $\beta(1-6)$ - instead of $\beta(1-3)$ -linked Gal (entry 8), did not improve binding affinities compared to GM1b.

The low abundance of GQ1b α (**4**) and the other “ α -series” gangliosides (0.5 mg/kg of brain), compared to the major brain ganglioside GD1a (**1**) (1,200 mg/kg), prompted several synthetic efforts. The first total synthesis of GQ1b α (**4**) on a milligram scale was achieved in 1995 by Kiso et al. [53]

The increased affinity of Chol-1 gangliosides suggests that the $\alpha(2-6)$ -Neu5Ac residue is in direct contact with Siglec-4. A detailed structure-activity analysis with derivatives of GD1 α (**13**), including a 7-, 8-, or 9-deoxy-Neu5Ac moiety linked $\alpha(2-6)$ to GalNAc [30, 55], revealed that the hydroxy groups of the glycerol side-chain are not essential for binding. In fact, the entire $\alpha(2-6)$ -Neu5Ac moiety of GD1 α can be replaced by a sulfate group (entry 14) without deleterious effect on MAG binding, suggesting that the negatively charged carboxylate at the anomeric position of the $\alpha(2-6)$ -Neu5Ac moiety is essential for binding [50]. Furthermore, the internal $\alpha(2-3)$ -linked sialic acid in GT1a α (**2**) can also be replaced by a sulfate, giving rise to the very potent MAG antagonist **15** [30, 50] (Table 1, entry 15; for its synthesis see Scheme 1), which displays tenfold higher binding affinity for MAG than the parent ganglioside GT1a α (**2**).



Scheme 1 Synthesis of the sulfated GT1a α mimic **15** [50]

In an effort to compare the glycan specificity of MAG (Siglec-4) and sialoadhesin (Sn, Siglec-1), Kelm et al. studied their interactions with a series of synthetic monovalent oligosaccharides using a hapten inhibition assay (Table 2) [27]. Native human erythrocytes were incubated with antibody-complexed MAG_{d1-3}-Fc in the

Table 2 Relative inhibitory potencies (rIPs) of oligosaccharide inhibitors of sialoadhesin (Sn, Siglec-4) and MAG (Siglec-4) [27]. The rIP of each substance was calculated by dividing the IC₅₀ of the reference compound by the IC₅₀ of the compound of interest. This results in rIPs above 1.00 for derivatives binding better than the reference and rIPs below 1.00 for compounds with a lower affinity. OSE: 2-(trimethylsilyl)ethoxy; Kdn: 2-keto-3-deoxy-D-glycero-D-galactononic acid; *n.a.*: not applicable since less than 50% inhibition at the highest concentration tested; *n.d.*: not determined

Entry	Compound	Ref. compd	rIP (Sn)	rIP (MAG)
<i>Influence of Neu5Ac linkage</i>				
1	Neu5Ac- α (2-3)-Gal- β (1-4)-Glc (16)	16	1	1
2	Neu5Ac- α (2-6)-Gal- β (1-4)-Glc (17)	16	0.21	0.12
3	Neu5Ac- α (2-3)-Gal- β (1-4)-AlNAc- β (1-3)-Gal- β (1-4)-Glc- β -OSE (18)	16	1.3	6.3
4	Neu5Ac- α (2-6)-Gal- β (1-3)-GlcNAc- β (1-3)-Gal- β (1-4)-Glc- β -OSE (19)	16	0.52	0.82
<i>Modifications of the terminal Neu5Ac</i>				
5	4-O-Methyl -Neu5Ac- α (2-3)-Gal- β (1-4)-Glc- β -EtN ₃ (20)	16	<i>n.d.</i>	0.43
6	4-Deoxy -Neu5Ac- α (2-3)-Gal- β (1-4)-Glc- β -EtN ₃ (21)	16	<i>n.d.</i>	0.56
7	7-Deoxy -Neu5Ac- α (2-3)-Gal- β (1-4)-Glc- β -EtN ₃ (22)	16	<i>n.d.</i>	1.56
8	8-Deoxy -Neu5Ac- α (2-3)-Gal- β (1-4)-Glc- β -EtN ₃ (23)	16	<i>n.d.</i>	<i>n.a.</i>
9	9-O-Methyl -Neu5Ac- α (2-3)-Gal- β (1-4)-Glc- β -EtN ₃ (24)	16	<i>n.d.</i>	0.18
10	Kdn - α (2-3)-Gal- β (1-4)-GlcNAc- β (1-3)-Gal- β (1-4)-Glc- β -OSE (25)	33	<i>n.a.</i>	6.47
<i>Modifications of the subterminal galactose</i>				
11	Neu5Ac- α (2-3)- (4-deoxy) Gal- β (1-4)-GlcNAc- β (1-3)-Gal- β (1-4)-Glc- β -OSE (26)	18	1.15	2.25
12	Neu5Ac- α (2-3)- (6-deoxy) Gal- β (1-4)-GlcNAc- β (1-3)-Gal- β (1-4)-Glc- β -OSE (27)	18	0.92	1.5
13	Neu5Ac- α (2-3)- (6-NH₂) Gal- β (1-4)-GlcNAc- β (1-3)-Gal- β (1-4)-Glc- β -OSE (28)	18	1.05	1.35
14	Neu5Ac- α (2-3)- (6-O-methyl) Gal- β (1-4)-GlcNAc- β (1-3)-Gal- β (1-4)-Glc- β -OSE <div style="margin-left: 100px;"> $\begin{array}{c} \text{Fuc-}\alpha(1-3) \\ \\ \text{Gal-}\beta(1-4) \end{array}$ </div>	18	0.77	0.5
<i>Linkage to the third terminal monosaccharide</i>				
15	Neu5Ac- α (2-3)-Gal- β (1-3)-GalNAc- β -OSE (30)	16	0.97	1.73
16	Neu5Ac- α (2-3)-Gal- β (1-3)-GlcNAc- β -OSE (31)	16	0.86	1.9

17	Neu5Ac- α (2-3)-Gal- β (1-4)-GlcNAc- β -OSE (32)	16	1.5	2.71
<i>Modifications of the third terminal hexosamine</i>				
18	Neu5Ac- α (2-3)-Gal- β (1-4)-GlcNAc- β (1-3)-Gal- β (1-4)-Glc- β -OSE (33)	33	n.d.	1
19	Neu5Ac- α (2-3)-Gal- β (1-4)-(3-deoxy)GlcNAc- β (1-3)-Gal- β (1-4)-Glc- β -OSE (34)	18	1.06	0.63
20	Neu5Ac- α (2-3)-Gal- β (1-4)-(3-O-methyl)GlcNAc- β (1-3)-Gal- β (1-4)-Glc- β -OSE (35)	33	n.d.	1
21	Neu5Ac- α (2-3)-Gal- β (1-4)-GlcNAc- β (1-3)-Gal- β (1-4)-Glc- β -OSE (36)	33	n.d.	0.7
Fuc-α(1-3)				
22	Neu5Ac- α (2-3)-Gal- β (1-4)-GlcNbutanoyl- β (1-3)-Gal- β (1-4)-Glc- β -OSE (37)	18	1.77	0.39
23	Neu5Ac- α (2-3)-Gal- β (1-4)-GlcNoctanoyl- β (1-3)-Gal- β (1-4)-Glc- β -OSE (38)	18	0.66	1.69
24	Neu5Ac- α (2-3)-Gal- β (1-4)-GlcNdecanoyl- β (1-3)-Gal- β (1-4)-Glc- β -OSE (39)	18	n.a.	n.a.
25	Neu5Ac- α (2-3)-Gal- β (1-4)-GlcNbenzoyl- β (1-3)-Gal- β (1-4)-Glc- β -OSE (40)	18	1.05	1.35
26	Neu5Ac- α (2-3)-Gal- β (1-4)-GlcNphtaloyl- β (1-3)-Gal- β (1-4)-Glc- β -OSE (41)	18	1.77	1.59
<i>Influence of additional Neu5Ac</i>				
27	Neu5Ac- α (2-6)	33	1.08	3.06
Neu5Ac-α(2-6)				
Neu5Ac- α (2-3)-Gal- β (1-3)-GalNAc- β (1-4)-Gal- β (1-4)-Glc- β -OSE (42)				
Neu5Ac-α(2-6)				
28	Neu5Ac- α (2-3)-Gal- β (1-3)-GalNAc- β (1-4)-Gal- β (1-4)-Glc- β -OSE (43)	33	1.21	0.95
Neu5Ac-α(2-3)				

presence of the monovalent oligosaccharides at different concentrations. The results are quoted as relative inhibitory potencies (rIPs), calculated by dividing the IC_{50} of the reference compound Neu5Ac- $\alpha(2-3)$ -Gal- $\beta(1-4)$ -Glc (**16**) by the IC_{50} of the compound of interest. Consequently, antagonists inhibiting better than the reference compound have rIP values above 1.0.

Whereas Schnaar and coworkers showed that only gangliosides bearing a terminal $\alpha(2-3)$ -linked sialic acid bind to MAG (Table 1), Kelm et al. also found modest inhibition by oligosaccharides with terminal $\alpha(2-6)$ -bound Neu5Ac (entries 2 and 4 in Table 2). However, a direct comparison of these two studies is not possible, since different assay formats were used. Whereas Schnaar et al. measured binding of full-length cell surface MAG to saccharides oriented on an apposing membrane monolayer, the group of Kelm used a fluorescent hapten inhibition assay to measure the affinity of soluble saccharides for a soluble MAG-Fc chimera (MAG_{d1-3}-Fc) engineered to have only the *N*-terminal three (of five) Ig-like domains [58].

Influence of Neu5Ac Linkage. In general, oligosaccharides with $\alpha(2-3)$ -linked sialic acid show an approximately eightfold higher activity than their isomers with $\alpha(2-6)$ -linked Neu5Ac (Table 2, entries 1–4). This ratio is independent of the underlying glycan structure, demonstrating the importance of the $\alpha(2-3)$ -linked sialic acid for MAG binding. However, since the rIP value of the pentasaccharide **18** (entry 3) is sixfold higher than that of $\alpha(2-3)$ -sialyl lactose (**16**, entry 1) and about twice as large as that of trisaccharide **32** (entry 17) with a similar terminal sequence, a binding epitope extending beyond the three terminal saccharide units can be assumed.

Modifications of the Terminal Neu5Ac. Methylation of the 4-OH (entry 5) or removal of the hydroxyl groups at C-4 (entry 6), C-8 (entry 8), or C-9 (entry 9) of the terminal sialic acid moiety are detrimental for activity, whereas deoxygenation at C-7 (entry 7) slightly improves the activity compared to reference compound **16**. Interestingly, the replacement of the entire terminal Neu5Ac moiety by Kdn (2-keto-3-deoxy-D-glycero-D-galacto-nononic acid) (\rightarrow **25**, entry 10) resulted in a 6.5-fold improved inhibition compared to the Neu5Ac-containing analogue **33** (entry 18).

Modifications of the Subterminal Galactose. Deoxygenation at C-4 or C-6 on the subterminal galactose (entries 11 and 12) or replacement of the 6-hydroxy by an amino group (entry 13) slightly improved affinity, whereas methylation of the 6-OH (entry 14) or the substitution of the ring oxygen by an *N*-alkyl group [27] had deleterious effects.

Linkage to the Third Terminal Monosaccharide. In contrast to the observations by Schnaar et al. (Table 1), the study of Kelm suggested that the linkage [$\beta(1-3)$ or $\beta(1-4)$] to and the nature (GalNAc, GlcNAc or AllNAc) of the second subterminal monosaccharide are of minor importance (entries 3 and 6–8). Nevertheless, the *N*-acetyl moiety appears to be beneficial for the inhibitory potency, since these trisaccharides are better inhibitors than $\alpha(2-3)$ -sialyl lactose (**16**, entry 1).

Modifications of the Third Terminal Hexosamine. The 3-position of GlcNAc is obviously not of high importance, since the inhibitory potencies of the 3-deoxy (entry 19), 3-*O*-methyl (entry 20), or 3-Fuc (entry 21) derivatives, or the 3-epimer (AllNAc, entry 3) are quite similar, suggesting that the binding pocket in this area is

relatively spacious, accepting even a fucose residue. Compounds bearing hydrophobic acyl groups such as octanoyl (entry 23), benzoyl (entry 25), or phthaloyl (entry 26) on the nitrogen of glucosamine showed moderately improved inhibition of MAG, indicating additional hydrophobic interactions with the lectin. In contrast, the butanoyl group in derivative **37** (entry 22) appears to be too small to reach this potential lipophilic site, whereas the decanoyl moiety in **39** (entry 24) is apparently too space demanding.

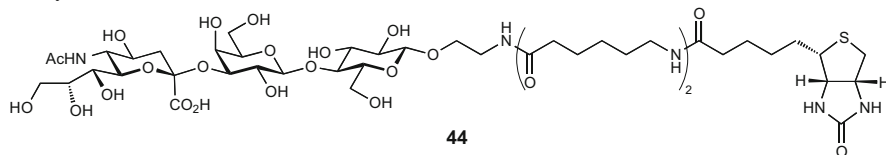
Influence of Additional Neu5Ac. Inhibition experiments with synthetically generated oligosaccharide fragments of gangliosides GD1a and GT1 α (entries 27 and 28) revealed that an additional $\alpha(2-6)$ -linked sialic acid enhances the affinity for MAG significantly (entry 27), whereas the introduction of a third sialic acid residue at position 3 of the underlying Gal moiety (entry 28) reverses this effect, giving an rIP value similar to that of compound **33** (entry 18). This result conflicts with the data obtained by Schnaar et al. [30], where GT1 α (**2**) was reported to be a fourfold better inhibitor than GM1b (**10**) lacking the two additional Neu5Ac moieties (Table 1). However, a direct comparison of data from the two studies [27, 30] should be carried out with caution, because different assay formats were applied.

In 2003, Paulson et al. investigated the relative specificity of ten human and murine siglecs towards neuraminic acid-containing saccharides representing terminal sequences on *N*- and *O*-linked oligosaccharides of natural glycoproteins and glycolipids [59]. The oligosaccharides were chemo-enzymatically synthesized [59–62] and tested directly or linked to *N*-hydroxy-succinimido-activated biotin in a competitive inhibition assay [59]. In this assay, a conjugate of a biotinylated oligosaccharide (for MAG compound **44**) and streptavidin-alkaline phosphatase (SAAP) was used to detect the siglec in the presence of increasing concentrations of each monovalent sialoside. The relative inhibitory potencies (rIPs) and IC₅₀ values obtained for MAG are summarized in Table 3.

As expected, MAG preferentially bound the Neu5Ac- $\alpha(2-3)$ -Gal epitope with a 5- to 13-fold increased affinity (Table 3, entries 2–4) compared to those with $\alpha(2-6)$ -Neu5Ac (entries 9–11). In addition, MAG exhibited a clear preference for Neu5Ac over Neu5Gc (entries 4 and 5 compared to entries 7 and 8). Modifications in the terminal Neu5Ac (entries 18 and 19) led to a strong decrease in affinity. Interesting results were obtained with compounds corresponding to sialylated derivatives of Tn (GalNAc- α -OThr) and T (Gal- $\beta(1-3)$ -GalNAc- α -OThr) antigens. MAG bound to the sialyl-Tn sequence **56** (entry 12) with two- to threefold higher affinity compared to the Neu5Ac- $\alpha(2-6)$ -Gal sequence (entries 10 and 11). Whereas the non-sialylated T antigen Gal- $\beta(1-3)$ -GalNAc- α -OThr showed no inhibition at 10 mM [59], the addition of sialic acid to either the 6-position of GalNAc (**50**, entry 13) or the 3-position of Gal (**57**, entry 6) dramatically increased the affinity, again with a preference for the $\alpha(2-3)$ -linked terminal sialic acid, indicated by the sixfold higher rIP of **57** compared to **50**. The additional sialic acid in the disialyl-T antigen **58** (entry 14) led to a further fivefold enhancement of affinity toward MAG, making **58** the first nanomolar MAG-antagonist reported.

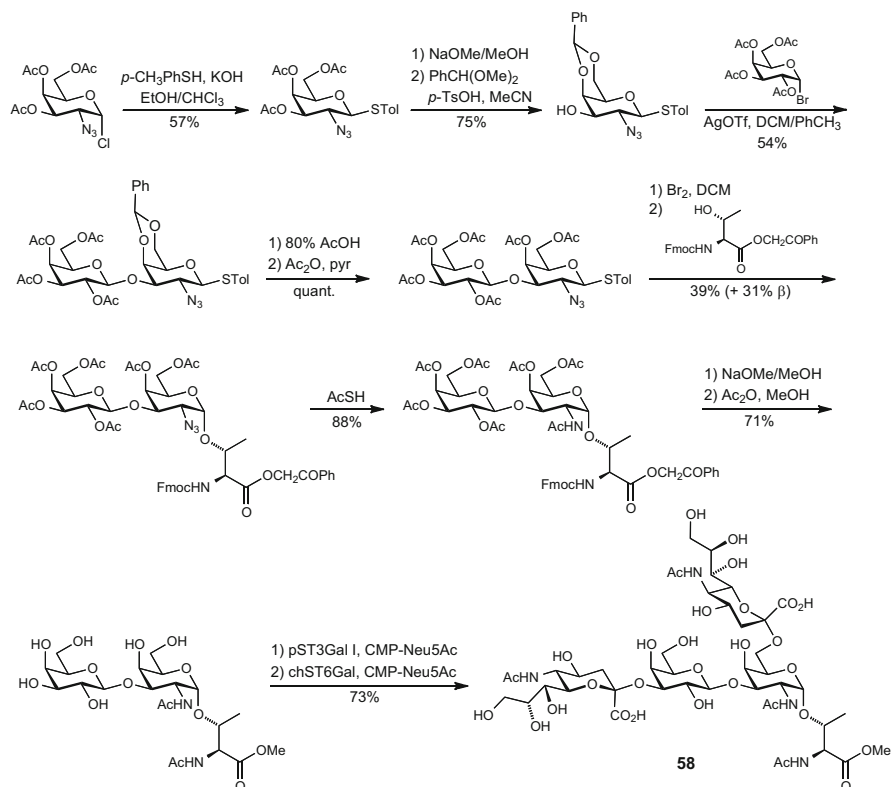
The disialyl-T antigen **58** (entry 14; for the synthesis see Scheme 2) contains the same terminal sequence as the gangliosides GD1 α (**13**), GT1 α (**2**), and GQ1b α

Table 3 Competitive inhibition of monovalent sialosides to murine MAG using a conjugate of biotinylated trisaccharide **44** and streptavidin-alkaline phosphatase (SAAP) [59]. Trisaccharide **48** (entry 4) was selected as reference compound. The relative inhibitory potency (rIP) for each sialoside is calculated as a percentage of the reference compound **48** (100%) according to the formula: $rIP = (IC_{50} \text{ reference compound} / IC_{50} \text{ sialoside}) \times 100$. ThrOCH₃: *N*-acetyl-threonine methyl ester



Entry	Compound	rIP [%]	IC ₅₀ [μM]
1	Neu5Ac-α-OCH ₃ (45)	12	4,083
2	Neu5Ac-α(2-3)-Gal-β-SCH ₃ (46)	192	255
3	Neu5Ac-α(2-3)-Gal-β(1-4)-Glc-β-OCH ₂ CH ₂ N ₃ (47)	94	521
4	Neu5Ac-α(2-3)-Gal-β(1-4)-GlcNAc-β-OCH ₂ CH ₂ N ₃ (48)	100	490
5	Neu5Ac-α(2-3)-Gal-β(1-3)-GlcNAc-β-OCH ₂ CH ₂ N ₃ (49)	94	521
6	Neu5Ac-α(2-3)-Gal-β(1-3)-GalNAc-α-OThrOCH ₃ (50)	29,600	1.6
7	Neu5Gc-α(2-3)-Gal-β(1-4)-GlcNAc-β-OCH ₂ CH ₂ N ₃ (51)	46	1,065
8	Neu5Gc-α(2-3)-Gal-β(1-3)-GlcNAc-β-OCH ₂ CH ₂ N ₃ (52)	48	1,021
9	Neu5Ac-α(2-6)-Gal-β-SCH ₃ (53)	14	3,500
10	Neu5Ac-α(2-6)-Gal-β(1-4)-Glc-β-OCH ₂ CH ₂ N ₃ (54)	21	2,333
11	Neu5Ac-α(2-6)-Gal-β(1-4)-GlcNAc-β-OCH ₂ CH ₂ N ₃ (55)	14	3,500
12	Neu5Ac-α(2-6)-GalNAc-α-OThrOCH ₃ (56)	40	1,225
13	Neu5Ac-α(2-6) Gal-β(1-3)-GalNAc-α-OThrOCH ₃ (57)	4,640	10.6
14	Neu5Ac-α(2-6) Neu5Ac-α(2-3)-Gal-β(1-3)-GalNAc-α-OThrOCH ₃ (58)	154,000	0.3
15	Neu5Gc-α(2-6)-Gal-β(1-4)-Glc-β-OCH ₂ CH ₂ N ₃ (59)	6	8,167
16	Neu5Gc-α(2-6)-Gal-β(1-4)-GlcNAc-β-OCH ₂ CH ₂ N ₃ (60)	4	12,250
17	Neu5Gc-α(2-6)-GalNAc-α-OThrOCH ₃ (61)	60	817
18	Neu5Ac(7-aldehyde)-α(2-3)-Gal-β(1-4)-GlcNAc-β-OCH ₂ CH ₂ N ₃ (62)	4	12,250
19	Neu5Ac(1-ethanolamide)-α(2-3)-Gal-β(1-4)-GlcNAc-β-OCH ₂ CH ₂ N ₃ (63)	4	12,250
20	Neu5Ac-α(2-3)-Gal-β(1-3)-GalNAc-β-SC ₆ H ₄ CH ₃ (64)	178	275
21	Neu5Ac-α(2-3)-Gal-β(1-3)-GalNAc-α-OCH ₂ Ph (65)	943	52.0
22	Neu5Ac-α(2-6)-Gal(2 N ₃)-β-SC ₆ H ₄ CH ₃ (66)	36	1,361

(**4**), which were identified as the most potent natural MAG antagonists (Table 1) [30, 34, 46]. However, whereas in the gangliosides the terminal tetrasaccharide epitope Neu5Ac-α(2-3)-Gal-β(1-3)-[Neu5Ac-α(2-6)]-GalNAc is linked β(1-4) to lactosylceramide, it is α-linked to the side-chain hydroxyl group of threonine in the case of disialyl-T antigen **58**. The α-configuration is essential for affinity as indicated by a comparison of the linear sequence Neu5Ac-α(2-3)-Gal-β(1-3)-GalNAc linked either β- (entry 20) or α- (entry 21) to an aromatic aglycone.



Scheme 2 Synthesis of disialyl-T antigen **58** [59, 63, 64]

Furthermore, a comparison of the benzyl aglycone (entry 21) to the threonine aglycone (entry 6) shows a further 30-fold improvement.

The sialylated T antigen derivatives **50**, **57**, and **58**, and the sialyllactoside **47**, were then tested by Schnaar, Paulson, and coworkers [65] for their ability to reverse MAG-mediated inhibition of axon outgrowth from rat cerebellar granule neurons in vitro (Table 4).

Whereas sialyllactoside **47** (entry 1) did not cause any significant reversal of inhibition, 6-sialyl-T **57** (entry 2) induced modest reversal at the highest concentration tested (100 μM). The MAG-binders with the highest affinities, 3-sialyl-T **50** (entry 3) and disialyl-T **58** (entry 4) caused statistically significant reversal of inhibition at all concentrations tested, with **58** having an approximately tenfold higher potency. These results suggest that MAG indeed uses sialidated glycans as its major axonal ligands.

In an effort to study the sialic acid binding pocket of MAG, the binding properties of the tri- and tetrasaccharides **67–71** [66, 67] (Table 5) towards a soluble recombinant protein consisting of the three N-terminal domains of MAG and the Fc part of human IgG (MAG_{d1-3}-Fc) were investigated using a fluorescent hapten inhibition assay [58]. The affinities were measured relative to the reference

Table 4 Comparison of MAG binding affinities and reversal of MAG-mediated axon outgrowth inhibition in vitro [65]

Entry	Compound	MAG binding IC ₅₀ [μM]	Axon outgrowth IC ₅₀ [μM]
1	Neu5Ac-α(2-3)-Gal-β(1-4)-Glc-β-OCH ₂ CH ₂ N ₃ (47)	521	–
2	Neu5Ac-α(2-6) Gal-β(1-3)-GalNAc-α-OThrOCH ₃ (57)	10.6	>100 ^a
3	Neu5Ac-α(2-3)-Gal-β(1-3)-GalNAc-α-OThrOCH ₃ (50)	1.6	80
4	Neu5Ac-α(2-6) Neu5Ac-α(2-3)-Gal-β(1-3)-GalNAc-α-OThrOCH ₃ (58)	0.3	6.2

^a22% reversal at 100 μM

compound **67** with a relative inhibitory potency (rIP) of 1 (entry 1; for the synthesis see Scheme 3). Tetrasaccharide **67** represents the relevant binding epitope of GQ1bα (**4**) responsible for MAG binding. In addition, for tetrasaccharide **67** as well as for the trisaccharides **69** and **70** the dissociation constants (K_D) of the ligand–MAG complex were determined by titrating a protein sample with a ligand solution and repeated acquisition of Saturation Transfer Difference (STD) NMR experiments [68]. For the parent compound **67** a K_D of 180 μM was obtained.

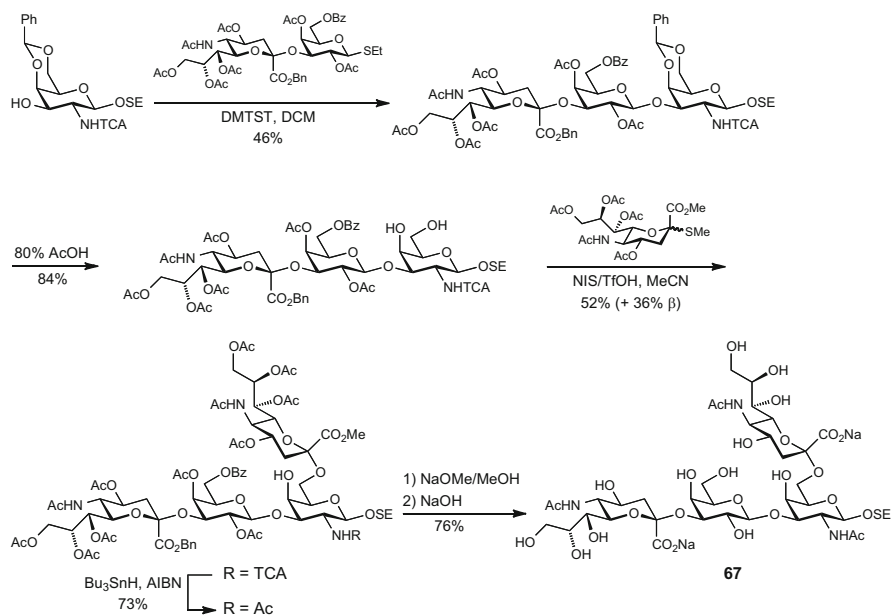
As in previous investigations [30, 31, 34, 46–48, 59], the tetrasaccharides **67** and **68** (entries 1 and 2) showed a fourfold higher affinity than the corresponding trisaccharides **69** and **70** (entries 3 and 4) missing the additional α(2-6)-linked Neu5Ac. Similar to the results obtained by Kelm et al. [27], the replacement of the GalNAc moiety present in the natural epitopes **67** and **69** (entries 1 and 3) by a Gal residue (→ **68** and **70**, entries 2 and 4) had no significant influence on the rIPs in the hapten assay. However, in the NMR titration experiments, compound **70** surprisingly showed a twofold higher affinity than **69** (K_D 390 μM vs 820 μM). The fivefold lower rIP for trisaccharide **71** (entry 5) compared to **70** (entry 4) again confirmed the preference of MAG for the α(2-3)-linked Neu5Ac epitope over α(2-6)-linked Neu5Ac.

The binding epitope, i.e., the proximity of individual parts of tetrasaccharide **67** to MAG, was elucidated by STD NMR [68] (Fig. 4). The largest STD response was observed for the *N*-acetyl function of the α(2-3)-linked sialic acid. Further significant contacts originate from H-6 and H-7 of the α(2-3)-linked sialic acid as well as from H-3, H-4, and H-5 of the central galactose residue indicating a hydrophobic interaction of the α-face of D-Gal with MAG. For the α(2-6)-linked sialic acid, the only sizeable STD effect stems from the *N*-acetyl group. In analogy to the knowledge gained from the crystal structure of the N-terminal domain of sialoadhesin (Siglec-1) in complex with 3'-sialyllactose [26], it was assumed that the carboxylic acid of the α(2-3)-linked sialic acid forms an important salt bridge with Arg118 [69].¹ In addition, docking of tetrasaccharide **67** to a homology model of

¹The numbering of the amino acids was carried out according to [69]; see also Swiss-Prot entry P20916.

Table 5 Relative inhibitory potencies (rIPs) and dissociation constants (K_D) of tri- and tetrasaccharides as MAG antagonists [66, 68]. The rIP of each substance was calculated by dividing the IC_{50} of the reference compound **67** by the IC_{50} of the compound of interest. This results in rIPs above 1.00 for derivatives binding better than **67** and rIPs below 1.00 for compounds with a lower affinity. *n.d.*: not determined; OSE: 2-(trimethylsilyl)ethoxy

Entry	Compound	rIP	K_D [μ M]
1	Neu5Ac- α (2-3)-Gal- β (1-3)-[Neu5Ac- α (2-6)]-GalNAc- β -OSE (67)	1.00	180
2	Neu5Ac- α (2-3)-Gal- β (1-3)-[Neu5Ac- α (2-6)]-Gal- β -OSE (68)	1.27	n.d.
3	Neu5Ac- α (2-3)-Gal- β (1-3)-GalNAc- β -OSE (69)	0.26	820
4	Neu5Ac- α (2-3)-Gal- β (1-3)-Gal- β -OSE (70)	0.30	390
5	Gal- β (1-3)-[Neu5Ac- α (2-6)]-Gal- β -OSE (71)	0.06	n.d.



Scheme 3 Synthesis of tetrasaccharide **67** [66] OSE: 2-(trimethylsilyl)ethoxy, TCA: trichloroacetyl

MAG [66] based on the X-ray structure of sialoadhesin (Siglec-1) [26] indicated that the carboxylic acid of the α (2-6)-linked sialic acid forms a second salt bridge with Lys67 (Fig. 5). Finally, the comparably low STD values found for the protons of the GalNAc residue (Fig. 4) indicated that no significant contribution to binding originates from this moiety of tetrasaccharide **67**.

Based on these results, and with the aim of obtaining structurally simplified and pharmacokinetically improved MAG antagonists, a number of mimics of tetrasaccharide **67** were prepared and their affinity towards MAG tested in the fluorescent hapten inhibition assay [58] relative to **67** [70, 71]. In these studies, the Gal- β (1-3)-GalNAc core was simplified, the α (2-6)-linked sialic acid was replaced

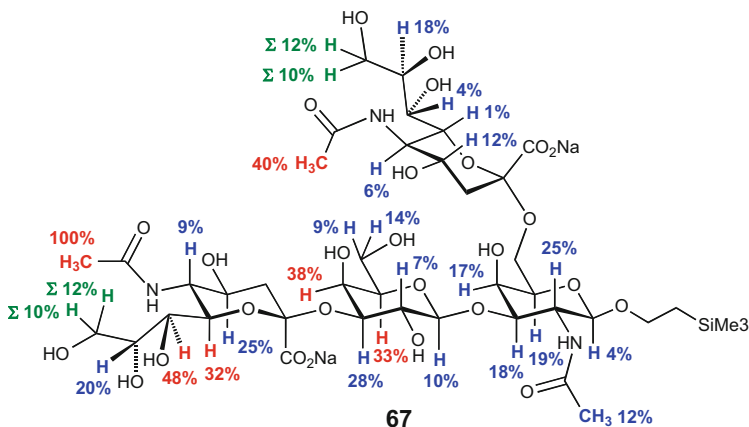


Fig. 4 Binding epitopes of tetrasaccharide **67** obtained from the integration of signals in the STD HSQC experiment [68]. The strongest signal (*N*-acetyl group at the $\alpha(2\text{--}3)$ -linked Neu5Ac, left) was scaled to 100%, and all other signals were related to it. The signal intensity was divided into three categories: red for 31–100% and blue for 1–30%, while overlapping signals were coded in green and designated by Σ , meaning that analysis was only possible for the sum of the signals. The absolute STD value of the *N*-acetyl group was 16%

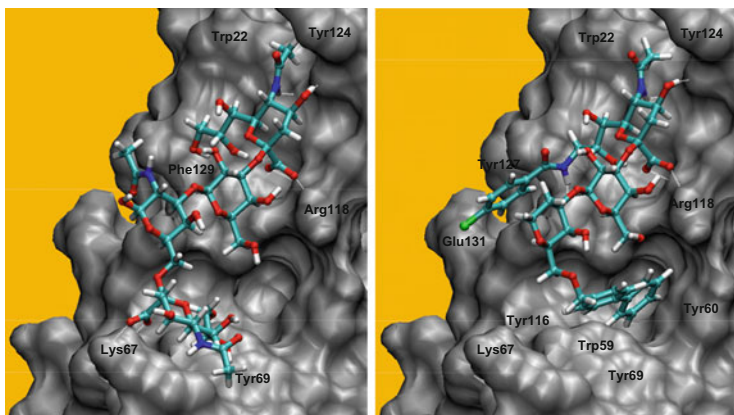


Fig. 5 Homology model of MAG – based on a 3D-structure of sialoadhesin (Siglec-1) [26] – complexed with tetrasaccharide **67** (left) and glycomimetic **85** (right) [66, 70]

by hydrophilic lactate residues or lipophilic benzyl ethers, and the terminal $\alpha(2\text{--}3)$ -linked Neu5Ac modified at the non-reducing end by benzamides (Table 6).

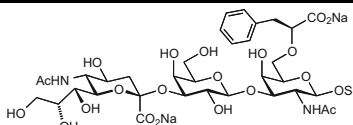
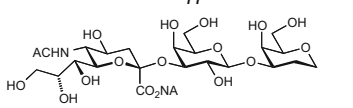
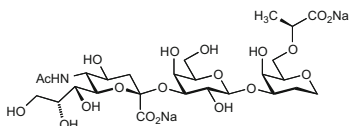
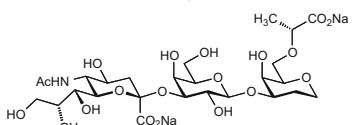
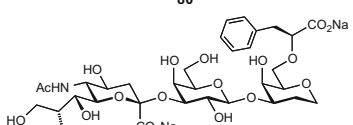
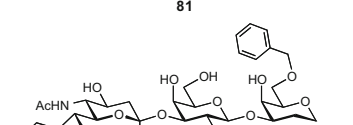
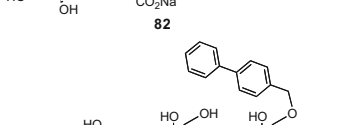
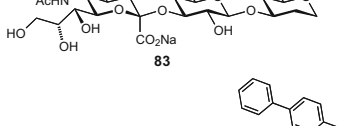
Replacement of the Gal- $\beta(1\text{--}3)$ -GalNAc core with a lipophilic biphenyl moiety (compound **72**, entry 3) resulted in a twofold reduced affinity compared to parent compound **67** (entry 1). The additional substitution of the $\alpha(2\text{--}6)$ -linked sialic acid by (*S*)-phenyllactate (\rightarrow **73**, entry 4) further reduced the affinity, indicating that the (*S*)-phenyllactate moiety is only incompletely mimicking the $\alpha(2\text{--}6)$ -linked sialic acid.

Table 6 Relative inhibitory potencies (rIPs) of oligosaccharide mimics. The rIP of each substance was calculated by dividing the IC_{50} of the reference compound **67** by the IC_{50} of the compound of interest. This results in rIPs above 1.00 for derivatives binding better than **67** and rIPs below 1.00 for compounds with a lower affinity. OSE: 2-(trimethylsilyl)ethoxy

Entry	Compound	rIP	Ref.
1		1.00	[66]
2		0.30	[66]
3		0.42	[71]
4		0.28	[71]
5		61.5	[70]
6		0.31	[70]
7		0.17	[70]

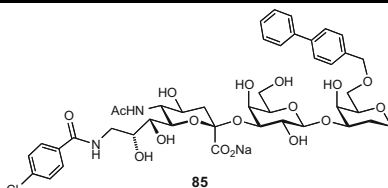
(continued)

Table 6 (continued)

Entry	Compound	rIP	Ref.
8	 <p style="text-align: center;">77</p>	0.40	[70]
9	 <p style="text-align: center;">78</p>	0.76	[70]
10	 <p style="text-align: center;">79</p>	1.24	[70]
11	 <p style="text-align: center;">80</p>	0.98	[70]
12	 <p style="text-align: center;">81</p>	0.40	[70]
13	 <p style="text-align: center;">82</p>	0.59	[70]
14	 <p style="text-align: center;">83</p>	1.25	[70]
15	 <p style="text-align: center;">84</p>	314	[70]

(continued)

Table 6 (continued)

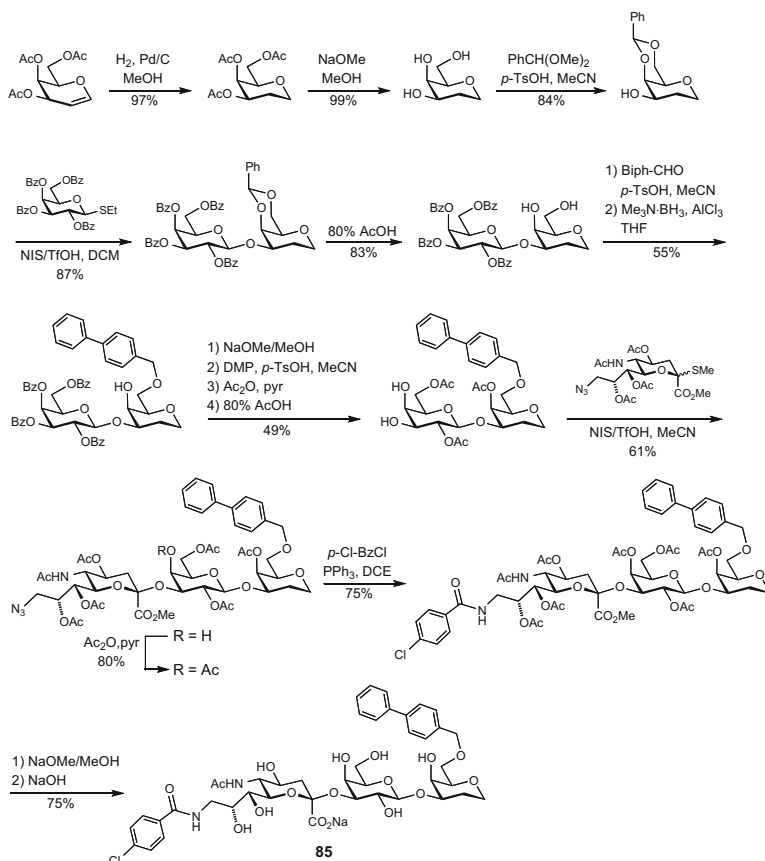
Entry	Compound	rIP	Ref.
16		377	[70]

The application of the known beneficial effect of aryl modifications in the 9-position of methyl and benzyl sialosides [72, 73] (see below) to trisaccharide **70** (\rightarrow **74**, entry 5) gave an approximately 60-fold improvement in affinity. The $\alpha(2-6)$ -linked sialic acid in the natural epitope **67** (entry 1) most probably is involved in a salt bridge [66], which obviously is not mimicked well by lactate residues, leading to a relevant decrease of affinity for both stereoisomers **75** and **76** (entries 6 and 7). The additional introduction of a lipophilic phenyl group at C-3 of the (*S*)-lactic acid moiety in **77** (entry 8) did not show a measurable effect.

The replacement of the second galactose moiety in trisaccharide **70** by 1,2-dideoxy-Gal in **78** (entry 9) raised the affinity close to that of tetrasaccharide **67** (rIP 0.76 vs 1.0). When the similar modification was applied to the pseudo-tetrasaccharides **75–77**, improved affinities were obtained for the (*R*)- and (*S*)-lactic acid derivatives **79** and **80** (entries 10 and 11), but not for the phenyllactic acid derivative **81** (entry 12). For further investigating the lipophilic contact emanating from the $\alpha(2-6)$ -linked Neu5Ac residue [68], the 6-position of the *lyxo*-hexitol moiety was derivatized with lipophilic substituents. Whereas benzyl ether **82** (entry 13) turned out to be less active than tetrasaccharide **67**, the biphenylmethyl derivative **83** (entry 14) proved to be superior.

Finally, concomitant modifications at the site of the $\alpha(2-3)$ - as well as at the $\alpha(2-6)$ -linked sialic acid proved the additivity of the beneficial interactions, leading to mimic **84** (entry 15) exhibiting a more than 300-fold enhanced affinity compared to the parent tetrasaccharide **67**. Similar to previous findings [73], the *p*-chloro substituent on the benzamide in **85** (entry 16) led to a further significant enhancement of the affinity compared to **84** (rIP: 377 vs 314). The synthesis of compound **85** is depicted in Scheme 4.

The beneficial contribution of the two lipophilic substituents on the opposing ends in the most active ligand **85** could be rationalized by comparing the docking modes of tetrasaccharide **67** and **85** to a homology model of MAG [66] based on the X-ray structure of sialoadhesin (Siglec-1) [26] (Fig. 5). Both compounds form an important salt bridge between the carboxylate of the $\alpha(2-3)$ -linked Neu5Ac with Arg118, a relevant hydrogen bond between the backbone carbonyl of Phe129 with 9-OH (in **67**) and 9-NH (in **85**), and a hydrophobic contact between the acetamido group of the $\alpha(2-3)$ -linked Neu5Ac and the aromatic side chains of Trp22 and Tyr124. However, whereas **67** establishes a second salt bridge between the



Scheme 4 Synthesis of tetrasaccharide mimic **85** [70]

carboxylate of the $\alpha(2\text{--}6)$ -linked Neu5Ac and Lys67, prominent interactions with two hydrophobic pockets of MAG are formed by **85**: the *p*-chlorobenzamide substituent is homed by Glu131 and Tyr127, and the biphenylmethyl group is accommodated in the main hydrophobic pocket lined by the side chains of Trp59, Tyr60, Tyr69, and Tyr116.

4 Neuraminic Acid Derivatives as MAG Antagonists – First Generation

In order to explore further the substructural determinants of sialic acid required for MAG binding, Kelm and Brossmer [28, 72], and later our group [73, 74], evaluated the inhibitory potential of simple neuraminic acid derivatives towards MAG and

sialoadhesin (Table 7). All affinities were determined with the hapten binding assay [58] and are given relative to the trisaccharide **70** for better comparison.

4.1 Modification in the 2-Position

Neu5Ac- α -Bn (**87**, entry 3 in Table 7) was found to have a tenfold higher affinity for MAG than the corresponding methyl sialoside **86** (entry 2) [28, 72]. Interestingly, compared to the trisaccharide reference compound **70** (entry 1), the potency of **87** was reduced only by less than half (rIP 0.62 vs 1.00). This effect is not unexpected and can be rationalized by the STD-NMR investigations of tetrasaccharide **67** (see above, Fig. 2), indicating a hydrophobic interaction of the α -face of D-galactose with MAG. Replacement of the OMe-aglycone in **86** with SMe or SBn (entries 4 and 5) did not markedly influence the affinity [72]. When the size of the aglycone was increased from benzyl in **87** to biphenylmethyl (entry 6) or phenoxyphenylmethyl (entry 7) the potency was reduced by half [74]. However, the replacement of the benzyl aglycone by phenoxyphenyl (entry 8), or by biphenyl- (entry 9) or phenyltriazole-based substituents (entries 10–12) resulted in an up to threefold increase in affinity (\rightarrow **96**, entry 12) compared to **87**, indicating an improved hydrophobic contact of the aglycone with MAG [74].



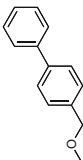
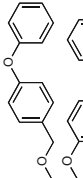
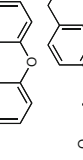
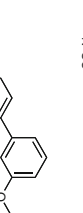
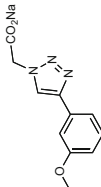
4.2 Modification in the 5-Position

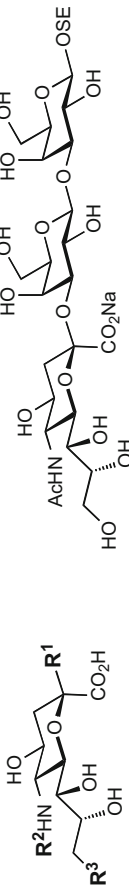
Since the substituents at positions 5 and 9 of neuraminic acid are critical for binding to siglecs [48, 75, 76], the inhibitory potencies of a series of sialosides with modifications at the 5-amino group were studied [28, 72]. Replacement of *N*-acetyl by propionyl (entries 13 and 14) or butyryl (entry 15) markedly reduced the rIP of sialosides **97–99**, whereas an *N*-thioacetyl residue (entry 16) leads to a fourfold increase in affinity for methyl sialoside **100**. Halogen substitution led to improved potencies, as shown by the chloroacetyl (**103**, entry 19) and trifluoroacetyl (**102**, entry 18) derivatives, which exhibited seven- and fourfold higher affinities for MAG than the parent compound **86**. By far the most beneficial modification was the introduction of an *N*-fluoroacetyl substituent in **101** (entry 17) leading to an 18-fold increase in potency compared to **86**. On the other hand, benzoyl (entry 20) and ethyl squarate (entry 21) residues markedly reduced the affinity.

4.3 Modification in the 9-Position

Substitution of the hydroxy group in the 9-position of neuraminic acid by an amino group (**112**, entry 28) enhanced binding, while substitution by hydrogen (entry 22),

Table 7 Relative inhibitory potencies (rIP) of neuraminic acid derivatives as MAG antagonists – first generation. The rIP of each substance was calculated by dividing the IC_{50} of the reference compound **70** by the IC_{50} of the compound of interest. This results in rIPs above 1.00 for derivatives binding better than **70** and rIPs below 1.00 for compounds with a lower affinity. *n.a.*: not applicable, since less than 50% inhibition at the highest concentration tested




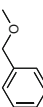












Entry	Compound	R ¹	R ²	R ³	rIP	Ref.
1	70	Gal-β(1-3)-Gal-OSE	Ac	OH	1.00	[66]
<i>Evaluation of substituents at the 2-position</i>						
2	86	-OCH ₃	Ac	OH	0.06	[28, 72]
3	87		Ac	OH	0.62	[28]
4	88	-SCH ₃	Ac	OH	0.08	[72]
5	89		Ac	OH	0.19	[72]
6	90		Ac	OH	0.38	[74]
7	91		Ac	OH	0.27	[74]
8	92		Ac	OH	0.89	[74]
9	93		Ac	OH	1.00	[74]
10	94		Ac	OH	1.10	[74]

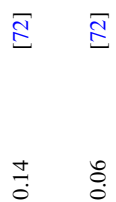

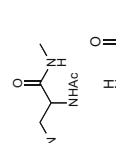
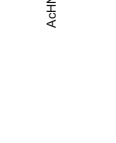
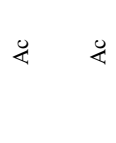



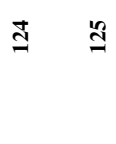
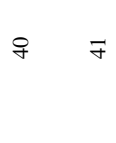



11	95		Ac	OH	0.90	[74]
12	96		Ac	OH	1.70	[74]
<i>Evaluation of substituents at the 5-position</i>						
13	97	-OCH ₃		OH	0.10	[28]
14	98			OH	0.30	[72]
15	99			OH	0.09	[72]
16	100	-OCH ₃		OH	0.24	[28]
17	101	-OCH ₃		OH	1.07	[28, 72]
18	102	-OCH ₃		OH	0.26	[28]
19	103	-OCH ₃		OH	0.44	[72]
20	104			OH	n.a.	[72]
21	105	-OCH ₃		OH	0.03	[72]
<i>Evaluation of substituents at the 9-position</i>						
22	106	-OCH ₃	Ac	H	n.a.	[28]
23	107	-OCH ₃	Ac	Cl	n.a.	[28]
24	108	-OCH ₃	Ac	SH	n.a.	[28]
25	109	-OCH ₃	Ac	SMe	n.a.	[72]
26	110	-OCH ₃	Ac		0.54	[72]

(continued)

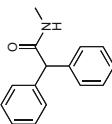
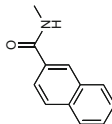
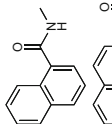
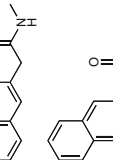
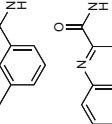
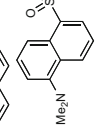
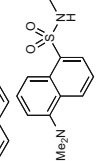
Table 7 (continued)

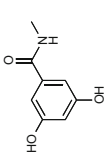
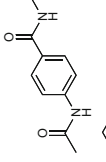
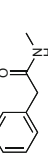
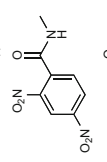
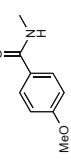
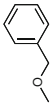
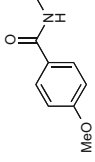
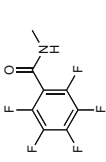
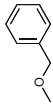
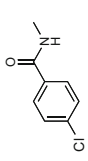
Entry	Compound	R ¹	R ²	R ³	rIP	Ref.
27	111	-OCH ₃	Ac		0.001	[72]
28	112	-OCH ₃	Ac	NH ₂	0.19	[28]
29	113		Ac	N ₃	0.05	[73]
30	114		Ac		2.8	[73]
31	115	-OCH ₃	Ac		0.24	[72]
32	116	-OCH ₃	Ac		0.41	[72]
33	117	-OCH ₃	Ac		0.63	[72]
34	118		Ac		8.6	[73]
35	119	-OCH ₃	Ac		3.42	[72]
36	120		Ac		12	[73]
37	121		Ac		12	[73]
38	122	-OCH ₃	Ac		3.16	[72]
39	123	-OCH ₃	Ac		0.39	[72]

40	124	-OCH ₃	Ac		0.14	[72]
41	125	-OCH ₃	Ac		0.06	[72]
42	126	-OCH ₃	Ac		2.03	[72]
43	127	-OCH ₃	Ac		44.5	[72]
44	128		Ac		690	[73]
45	129	-OCH ₃	Ac		10.7	[72]
46	130	-OCH ₃	Ac		13.8	[72]
47	131		Ac		191	[73]
48	132		Ac		75	[73]

(continued)


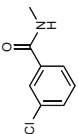

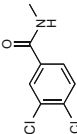

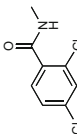

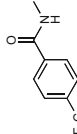

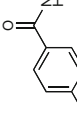

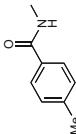

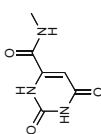
Table 7 (continued)

Entry	Compound	R ¹	R ²	R ³	rIP	Ref.
49	133	-OCH ₃	Ac		1.96	[72]
50	134	-OCH ₃	Ac		14.9	[72]
51	135	-OCH ₃	Ac		5.31	[72]
52	136	-OCH ₃	Ac		0.82	[72]
53	137	-OCH ₃	Ac		1.77	[72]
54	138	-OCH ₃	Ac		3.42	[72]
55	139	-OCH ₃	Ac		1.20	[72]

56	140	-OCH ₃	Ac		10.5	[72]
57	141	-OCH ₃	Ac		4.87	[72]
58	142	-OCH ₃	Ac		0.44	[72]
59	143	-OCH ₃	Ac		1.14	[72]
60	144	-OCH ₃	Ac		24.9	[72]
61	145		Ac		290	[73]
62	146	-OCH ₃	Ac		2.66	[72]
63	147		Ac		1,074	[73]

(continued)

Table 7 (continued)

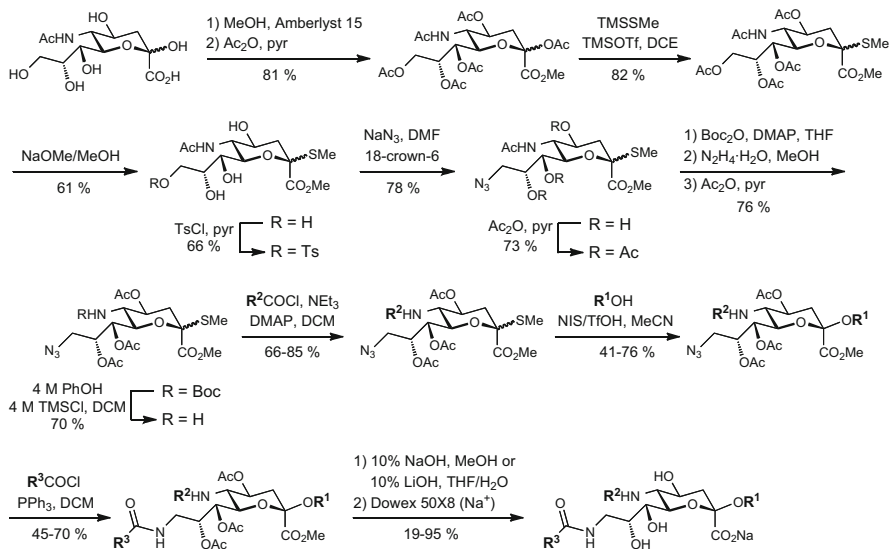
Entry	Compound	R ¹	R ²	R ³	rIP	Ref.
64	148		Ac		483	[73]
65	149		Ac		690	[73]
66	150		Ac		35	[73]
67	151		Ac		207	[73]
68	152		Ac		414	[73]
69	153		Ac		806	[73]
70	154		Ac		6.7	[73]

chlorine (entry 23), SH (entry 24), SMe (entry 25), methylsulfon (entry 27), or azide (entry 29) decreased or abolished binding, suggesting the important role of a hydrogen donor at this position [28, 72, 73]. Additional acylation of the amino group in the 9-position further increased the affinity for MAG significantly. Generally, aromatic amides (entries 43–70) turned out to be significantly more active than their aliphatic counterparts (entries 31–41). A benzamide in the 9-position of benzyl α -sialoside **128** (entry 44) improves binding affinity by almost three orders of magnitude compared to **87**. However, sterically more demanding aromatic residues (entries 46–55) led to decreased activity. Again, derivatives with a benzyl group at the 2-position were found to be approximately tenfold more potent than the corresponding methyl sialosides, as exemplified for the benzamides **127** (rIP 44.5, entry 43) and **128** (rIP 690, entry 44) or the biphenyls **130** (rIP 13.8, entry 46) and **131** (rIP 191, entry 47). A further optimization of the substituent at the 9-position using a non-mathematical Topliss approach delivered the *p*-chlorobenzamide **147** (entry 63) as the most potent compound in this series, exhibiting a 1,000-fold higher affinity than the reference trisaccharide **70** [73]. The Topliss operational scheme [77, 78] is a manual, non-mathematical application of the Hansch analysis [79, 80] to drug design, developed to guide towards the most active analogue of a lead compound with the least synthetic investment. This approach allows for the selection of a limited group of substituents on aromatic rings, which will give good discrimination between π (hydrophobic effects), σ (electronic effects), and E_S (steric effects). A generally applicable synthesis [73, 81] of neuraminic acid derivatives with modifications in the 2-, 5-, and 9-position from sialic acid is shown in Scheme 5.

5 Neuraminic Acid Derivatives as MAG Antagonists – Second Generation

As shown above, a broad optimization effort for the 9-position had led to the identification of *p*-chlorobenzamide as the best substituent (\rightarrow **147**, entry 63 in Table 7). Because the sheer enlargement of the hydrophobic group in the 2-position did not exhibit improved affinities (entries 6–12 in Table 7), the electron densities of the aromatic aglycones were altered in a next step (entries 4–9, Table 8) [81]. Then, with the substituents in the 2- and 9-positions set, a further optimization of the acyl group in the 5-position was conducted (entries 10–16). The binding properties of the neuraminic acid derivatives **155–167** again were evaluated with the hapten binding assay relative to sialoside **147**. In addition, the dissociation constants K_D were determined in a surface plasmon resonance (SPR) based biosensor (Biacore) experiment [70].

Finally, since molecular modeling studies with a homology model of MAG suggested that the hydroxy group in the 4-position is not directly involved in the binding process [70], its impact on the kinetic binding behavior was investigated by



Scheme 5 Synthesis of neuraminic acid derivatives modified in the 2-, 5-, and 9-positions [73, 81]

the introduction of additional hydrophobic substituents in the 4-position or the inversion of the configuration at C-4 (entries 17–27) [82].

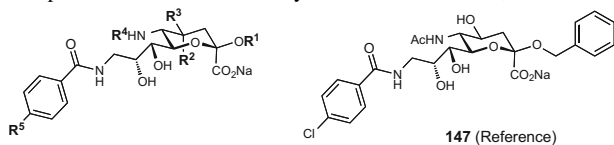
5.1 Modification in the 2-Position

A substantial increase in affinity was achieved when the aromatic aglycone was halogenated in *ortho*- and *meta*-positions (**159** and **160**, entries 8 and 9 in Table 8) [81]. Less effective were halogens in the *para*-position (entries 4 and 5). In addition, with fluorine instead of chlorine, consistently slightly higher affinities (entries 4 and 8 vs entries 5 and 9) were obtained. The observed diminished affinities of the pentafluoro analogue **157** (entry 6) and the 2-naphthylmethyl derivative **158** (entry 7) were attributed to steric hindrance.

5.2 Modification in the 5-Position

As shown above (Table 7), halogenated acetamides in the 5-position of sialic acid derivatives strongly improve binding of MAG antagonists [28]. The application of this beneficial modification to the so far most active antagonist **160** (entry 9) led to the nanomolar fluoroacetamide **161** (entry 10). For chloroacetamide derivative **162** (entry 11), the effect was less pronounced. Interestingly, an equally potent

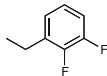

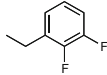

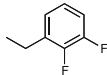
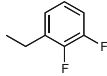
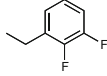
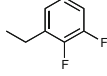
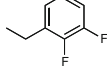

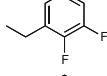
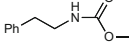
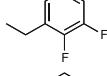
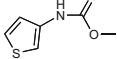
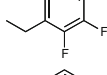
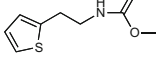
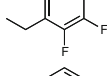
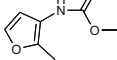
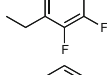
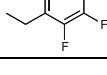
Table 8 Relative inhibitory potencies (rIP) and dissociation constants (K_D) of neuraminic acid derivatives as MAG antagonists – second generation. The rIP of each substance was calculated by dividing the IC_{50} of the reference compound **147** by the IC_{50} of the compound of interest. This results in rIPs above 1.00 for derivatives binding better than **147** and rIPs below 1.00 for compounds with a lower affinity. *n.d.*: not determined; *n.b.*: not binding



Entry	Compound	R ¹	R ²	R ³	R ⁴	R ⁵	rIP	K_D [μ M]	Ref.
1	147		H	OH	Ac	Cl	1.00	17	[73]
<i>Evaluation of substituents at the 2-position</i>									
2	127	CH ₃	H	OH	Ac	H	0.06	137	[72]
3	128		H	OH	Ac	H	0.67	26	[73]
4	155		H	OH	Ac	Cl	0.77	15	[81]
5	156		H	OH	Ac	Cl	0.38	13	[81]
6	157		H	OH	Ac	Cl	3.85	6.1	[81]
7	158		H	OH	Ac	Cl	1.35	11.6	[81]
8	159		H	OH	Ac	Cl	2.00	4.3	[81]
9	160		H	OH	Ac	Cl	3.33	2.4	[81]
<i>Evaluation of substituents at the 5-position</i>									
10	161		H	OH		Cl	50.0	0.5	[81]
11	162		H	OH		Cl	14.3	2.1	[81]
12	163		H	OH		Cl	20.0	1.4	[81]
13	164		H	OH		Cl	1.67	17	[81]
14	165		H	OH		Cl	7.14	2.3	[81]

(continued)

Table 8 (continued)

Entry	Compound	R ¹	R ²	R ³	R ⁴	R ⁵	rIP	K _D [μM]	Ref.
15	166		H	OH		Cl	10.0	4.1	[81]
16	167		H	OH		Cl	7.14	5.8	[81]
<i>Evaluation of substituents at the 4-position</i>									
17	168		H	H	Ac	Cl	n.d.	11.4	[82]
18	169		H	OAc	Ac	Cl	n.d.	11.5	[82]
19	170		H	OBz	Ac	Cl	n.d.	26.0	[82]
20	171		H	OBn	Ac	Cl	n.d.	2.1	[82]
21	172		H		Ac	Cl	n.d.	257	[82]
22	173		H		Ac	Cl	n.d.	114	[82]
23	174		H		Ac	Cl	n.d.	15.6	[82]
24	175		H		Ac	Cl	n.d.	n.b.	[82]
25	176		H		Ac	Cl	n.d.	49.4	[82]
26	177		OH	CH ₃	Ac	Cl	n.d.	30.0	[82]
27	178		OH	H	Ac	Cl	n.d.	9.0	[82]

antagonist was achieved with the nosyl substituent in **163** (entry 12), while sulfone **164** (entry 13) suffers from a drastic loss in activity. Thus, the increase in the size of the acyl substituent (entries 14–16) led to a reduction in affinity [81].

5.3 Modification in the 4-Position

Removal of the 4-OH group (entry 17), inversion of the stereocenter at the C-4 (entry 27), or the introduction of a methyl group in equatorial position (entry 26) led to a substantial drop in binding affinity, indicating a contribution to binding of the equatorial 4-hydroxy, most probably by hydrogen bonding [82]. Acylation (entries 18 and 19) of the 4-OH or modification by carbamates (entries 21–25) also resulted in a pronounced decrease or even complete abolition of binding, which is probably due to the rigid geometry of ester and carbamate substituents, leading to a steric clash with the protein. In contrast, the more flexible benzyl ether in **171** (entry 20) showed the same binding affinity as parent compound **160** (entry 9).

In summary, with sialoside **161** (entry 10) the first nanomolar, low-molecular, and drug-like MAG antagonist was obtained. Its high affinity results from the additivity of three beneficial modifications: (1) a difluorobenzyl substituent at the 2-position, (2) a fluoroacetate at the 5-position, and (3) a *p*-chlorobenzamide at the 9-position. Docking of **161** to the homology model of MAG [66] (see Fig. 7a, b) revealed that the high affinity results from hydrophilic as well as hydrophobic interactions. The most important contribution stems from the salt bridge between the carboxylic acid and Arg118 [26, 28]. Additionally, hydrogen-bond formations between 5-NH and the backbone carbonyl of Gln126, the carboxylate and the OH of Thr128, 8-OH and the backbone NH of Thr128, and 9-NH and the backbone carbonyl of Thr128 were observed. A considerable contribution to the binding affinity results from hydrophobic interactions of the *p*-chlorobenzamide with Ser130 and Glu131, and of the aglycone with a second hydrophobic pocket formed by the side chains of Trp59, Tyr60, and Tyr69.

6 Design of a Non-Carbohydrate Antagonist of MAG

An attempt to increase the metabolic stability of sialosides like **161**, for example against mammalian sialidases, led to the corresponding non-carbohydrate mimic **179** (Fig. 6) [83]. Based on manual docking of **161** to a homology model of MAG (Fig. 7a, b), beneficial modifications of side chains in the 2-, 5-, and 9-position of Neu5Ac [28, 72, 73, 81] combined with the replacement of the pyranose core by a cyclohexane ring, the substitution of the 4-OH by an acetamide, and the simplification of the glycerol side chain in the 6-position led to **179** (Fig. 6).

However, when tested in the fluorescent hapten binding assay, **179** exhibited about 200,000-fold reduced affinity for MAG compared to **161** and in the Biacore experiment compound **179** showed no binding up to a concentration of 800 μmol [83].

Extensive molecular docking studies and molecular-dynamics simulations of **161** and **179** to the homology model of MAG [66] revealed the unexpected loss of binding affinity for mimic **179** (Fig. 7) [83]. In both compounds the carboxylate

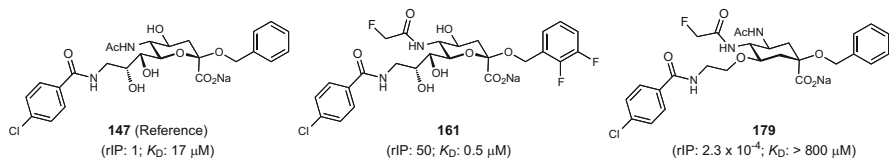


Fig. 6 Structures and affinities of sialoside **161** and non-carbohydrate mimic **179** [83]. The rIPs are relative to reference compound **147**

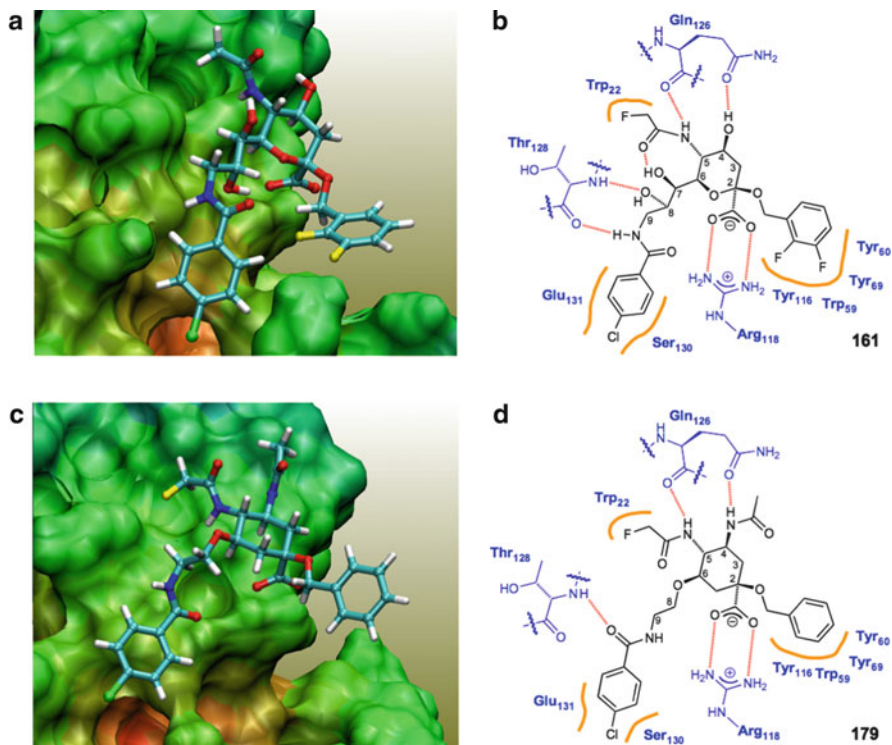


Fig. 7 (a, c) The homology model of MAG complexed with sialoside **161** (a) and mimic **179** (c) [83]. The ligands are depicted colored by atom (C: cyan, H: white, O: red, N: blue, Cl: green, F: yellow); the protein is represented by its van der Waals surface colored according to the depth. (b, d) Binding interactions of **161** (b) and **179** (d); hydrogen bonds are indicated as red dashed lines, hydrophobic interactions are shown as orange bars. For better comparability, the carbon atoms in **179** are numbered in analogy to compound **161**

forms a salt-bridge with Arg118, the hydrogen of the 5-NHFAc group interacts with the backbone carbonyl of Gln126, the 4-OH (in **161**) or 4-NH (in **179**) are engaged in hydrogen bonds with the side chain carbonyl of Gln126, and the (2,3-difluoro) benzyl aglycone is hosted by a hydrophobic pocket formed by Trp59, Tyr60, Tyr69, and Tyr116. However, whereas in **161** the 9-NH of the benzamido group establishes

an H-bond with the backbone carbonyl of Thr128, in **179** the 9-benzamido carbonyl interacts with the backbone NH of Thr128. Consequently, the orientations of the two side chains in the 6-position, the glycerol side chain in **161**, and the 2-benzamido-ethyl side chain in **179** are slightly different, leading to a noticeably reduced interaction of the *p*-chlorobenzamide in **179** with the second hydrophobic pocket formed by Ser130 and Glu131.

Whereas the docking studies were based on a thermodynamic inspection of the binding event, molecular-dynamics simulations gave a more detailed picture of the dynamic stability of the protein–ligand complexes, revealing obvious differences in respect of the hydrogen-bond network [83]. While frequency and strength of the salt bridge to Arg118 and the hydrogen bond of the 5-fluoroacetamido NH with the backbone carbonyl of Gln126 were found to be comparable for **161** and **179**, the interaction patterns of the side chains are markedly different. In **161**, two important hydrogen bonds contribute to binding: (1) the 9-benzamido NH interacts with the backbone carbonyl of Thr128 and (2) the backbone NH of the same amino acid acts as donor in a H-bond with the oxygen of the 8-OH. In mimic **179** the missing hydrogen bond of the former 8-OH leads to an inverted orientation of the benzamidoethoxy side-chain, allowing only the formation of a significantly weaker H-bond between the backbone NH of Thr128 and the 9-benzamido carbonyl. As a consequence, the hydrophobic contact of the *p*-chlorobenzamide with Ser130 and Glu131 is markedly reduced. Obviously, the 8-OH in sialoside **161** represents a key polar group [84] and its absence in mimic **179** results in the dramatic loss in affinity. Interestingly, the hydrogen bond of the 4-OH in **161** with the side chain carbonyl of Gln126, which was identified in the static docking studies, turned out to be irrelevant in the dynamic situation. Although the hydrogen bond of the 4-acetamido NH in **179** with the side chain carbonyl of Gln126 substantially contributes to binding, this additional interaction did not compensate for the loss of affinity caused by the reduced hydrophobic contact of the benzamido group in **179** compared to **161**.

7 Kinetic Studies

In drug discovery the in vitro determined drug-target interactions are classically rated in terms of binding parameters such as IC_{50} s and K_D s. However, an alternative perspective on drug optimization is the residence time of the drug–target binary complex [85], as quantified by the dissociation half-life ($t_{1/2}$). For medical applications, the potential advantages of a long residence time involve extended duration of the pharmacological effect and higher target selectivity [85, 86]. Therefore, an equal therapeutic effect can be reached with a lower dose. In general, high efficacy of a ligand in vivo correlates with fast association and slow dissociation rates [87, 88]. Especially in the field of carbohydrate–lectin interactions, this is a crucial point to address. As a result of the shallow and water-accessible binding sites of lectins, carbohydrates bind with only low affinity and show very fast dissociation rates k_{off} ,

leading to $t_{1/2}$ in the range of seconds, as in the case of the interactions of PSGL-1 and ESL-1 with P- and E-selectin [89, 90] or of sialyl Lewis^a with GSLA-2 mAb [91].

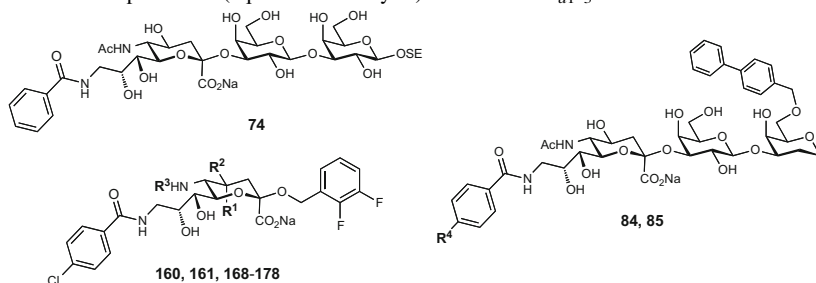
Since the association rate k_{on} is concentration dependent, slow association rates can be compensated by high concentrations. Slow dissociation rates k_{off} , however, must be achieved by optimizing the interaction of ligand and target, for example by introducing additional lipophilic contacts leading to an extended dissociation half-life ($t_{1/2}$) of the drug–target complex [87, 88, 92]. Zanamivir is one of the prominent examples, where a carbohydrate-based lead was optimized to yield a drug with a dramatically improved kinetic behavior, showing a half-life of 33 min of its complex with the B/Memphis/3/89 (H3N2) influenza virus [93].

In order to reveal their binding kinetics, the interaction between several MAG-antagonists and MAG was analyzed by surface plasmon resonance (SPR) experiments [94–96] (Table 9). For this purpose, MAG_{d1-3}-Fc [58] was immobilized on a CM5 dextrane chip containing either a surface of covalently bound Fc-specific goat antihuman IgG antibody or protein A. A reference cell providing only the IgG antibody or protein A was used to compensate for unspecific binding to the matrix. Affinities (K_{D}) as well as association (k_{on}) and dissociation rate constants (k_{off}) were determined [70, 82].

For the oligosaccharide mimics **74**, **84**, and **85** the K_{D} s corresponded well with the results obtained in the hapten inhibition assay (see Table 6). Whereas for trisaccharide **74** (Table 9, entry 1) a K_{D} of 32 μM was measured, the introduction of the lipophilic biphenylmethyl moiety in **84** or **85** (entries 2 and 3) improved the affinities to the low micromolar level. The kinetic analysis of the three oligosaccharides revealed that the increase in lipophilicity in **84** and **85** led to a substantial improvement of the association rate constants k_{on} by a factor of 20 and 30, respectively, whereas the k_{off} was rather stable (0.3–1 s^{-1}), resulting in half-lives $t_{1/2}$ in the range of approximately 1–2 s [70].

As shown before, halogenated acetates at the 5-position of the more drug-like sialosides led to a drastic improvement of the binding affinity (see Tables 7 and 8). The kinetic binding properties of **160** (K_{D} : 2.0 μM) and **161** (K_{D} : 0.55 μM) at 25 °C (entries 5 and 7) exhibit that the introduction of the *N*-fluoroacetyl group doubled the on-rate while the off-rate was reduced approximately threefold, leading to half-lives in the range of 1–5 s. When the Biacore experiment was repeated at a lower temperature (5 °C), a clear slowdown of k_{off} for **161** (entry 6) compared to that for **160** (entry 4) was observed resulting in a $t_{1/2}$ of 10 s [82].

Because an increased lipophilicity often leads to prolonged residence times [87, 88, 92], hydrophobic substituents were introduced in the 4-position of **160** and the influence of the configuration at this position on the kinetic binding behavior was investigated. However, the test compounds **168–178** (entries 8–18) showed decreased binding affinity and no substantial improvement of the kinetic properties could be achieved. All compounds showed fast dissociation rate constants, leading to residence times in the region of seconds. Despite the increased

Table 9 Dissociation constants K_D and kinetic parameters obtained from surface plasmon resonance experiments (equilibrium analysis) with Fc-MAG_{d1-3} immobilized on a CM5 chip

Entry	Compound	R ¹	R ²	R ³	R ⁴	K_D [μM]	k_{on} [$\text{M}^{-1} \text{s}^{-1}$]	k_{off} [s^{-1}]	$t_{1/2}$ [s]	Ref.
1	74					32.6	1.1×10^4	0.30	2.3	[70]
2	84				H	5.0	2.2×10^5	1.05	0.67	[70]
3	85				Cl	2.8	3.1×10^5	0.80	0.87	[70]
4	160 ^{5°C}	H	OH	Ac		1.6	9.4×10^4	0.153	4.5	[82]
5	160 ^{25°C}	H	OH	Ac		2.0	1.5×10^5	0.54	1.2	[82]
6	161 ^{5°C}	H	OH			0.25	2.75×10^5	0.07	9.9	[82]
7	161 ^{25°C}	H	OH			0.55	2.8×10^5	0.154	4.5	[82]
8	168	H	H	Ac		11.4	2.8×10^4	0.28	2.5	[82]
9	169	H	OAc	Ac		11.5	4.6×10^4	0.52	1.3	[82]
10	170	H	OBz	Ac		26.0	1.6×10^4	0.41	1.7	[82]
11	171	H	OBn	Ac		2.1	1.6×10^4	0.33	2.1	[82]
12	172	H		Ac		257	2.7×10^4	7.00	0.1	[82]
13	173	H		Ac		114	5.3×10^6	>100	<0.01	[82]
14	174	H		Ac		15.6	n.d.	n.d.	n.d.	[82]
15	175	H		Ac		n.b.	n.d.	n.d.	n.d.	[82]
16	176	H		Ac		49.4	2.2×10^4	1.10	0.6	[82]
17	177	OH	CH ₃	Ac		30.0	2.4×10^4	0.53	1.3	[82]
18	178	OH	H	Ac		9.0	3.6×10^4	0.33	2.1	[82]

lipophilicity, even the best compound of the series (**171**, entry 11) did not show an extended dissociation half-life. In conclusion, the prolongation of the residence time still remains a challenge and should be a critical issue in further studies on the kinetic properties of MAG antagonists and glycomimetics in general.

8 High-Affinity MAG Antagonists by a *Fragment-Based In Situ Combinatorial Approach*

To further improve the affinity of MAG antagonists, a search for second binding sites was conducted following the concept of fragment-based drug discovery (FBDD)²: When two low-affinity fragments are linked, their individual binding energies are additive. In addition, due to the reduction of translational and rigid body rotational degrees of freedom, the entropy barrier is markedly lowered [100], leading to a new ligand with a substantially improved affinity.

However, the inspection of the homology model of MAG based on sialoadhesin (Siglec-1) [66] revealed a shallow, unstructured binding site not providing obvious additional interaction sites and therefore with little prospect for success by a structure-based approach. Therefore a novel, generally applicable three-step fragment-based in situ combinatorial approach was developed (Fig. 8) [101, 102]. As first site-ligand, the sialoside **127** [72] (K_D 137 μ M, Table 8, entry 2) was selected as starting point (Fig. 9). In order to search for second-site ligands, in the *first step* a fragment library consisting of 80 diverse, drug-like, soluble, and synthetically easily modifiable compounds was screened by NMR. The members of the library binding to MAG were identified based on the change of their transverse magnetization decay which occurs upon complex formation with the protein (Fig. 8a) [103, 104].

In the *second step* those second-site ligands binding in the close vicinity (approximately 10 Å) of the binding site of **127** were identified by their enhanced paramagnetic relaxation caused by the spin-labeled first-site ligand **180*** containing an unpaired electron (Figs. 8b and 9). The method detects the distant-dependent paramagnetic relaxation enhancement exerted on a second-site ligand by a spin-labeled first-site ligand, when both ligands are bound to the target protein simultaneously and at neighboring binding sites. **180*** is a conjugate of first-site ligand **127** and the TEMPO (2,2,6,6-tetramethylpiperidin-1-oxyl) spin-label. TEMPO is a suitable spin-label since the gyromagnetic ratio of an unpaired electron is 658 times as large as that of a proton. Therefore, TEMPO [105–109] causes significant relaxation enhancement of nearby protons. Because **127** (K_D 137 μ M) and the *N*-hydroxy derivative **181** (reduced form of **180***, K_D 96 μ M) exhibited similar affinities in SPR (Biacore) experiments, it was assumed that the TEMPO spin-label does not substantially alter the binding mode of the first-site ligand [101]. The screening led to the identification of 5-nitroindole (**182**) as the most promising second-site ligand (Fig. 9). Since the paramagnetic relaxation enhancement caused by the spin-label is distance-dependent, the spatial orientation of **182** could be determined and the carbons of the pyrrole ring were identified as the optimal linking positions.

² For recent reviews, see [97–99].

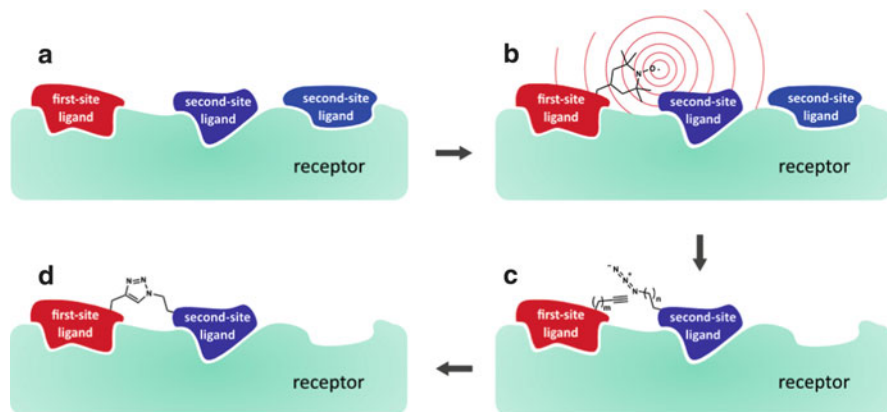


Fig. 8 Fragment-based in situ combinatorial approach to identify ligands for an unknown binding site [101, 102]. (a) Identification of second-site ligands based on their transverse magnetization decay upon binding. (b) Selection of second-site ligands binding adjacent to the spin-labeled first-site ligand. (c) Incubation of the target protein with libraries of first- and second-site ligands. (d) By receptor-mediated triazole formation a high-affinity ligand is generated

In the *third step* MAG was incubated for 72 h with small libraries of four first- (**183a–d**) and three second-site (**184a–c**) fragments equipped with azido or acetyleno groups at the end of flexible methylene chains of variable length, in principle permitting the formation of 12 *syn*- and 12 *anti*-substituted triazoles (Figs. 8c and 10). To minimize the high entropic penalties associated with the introduction of rotatable bonds [110], only short linkers (up to four carbons) were considered. Only in the case of an optimal spatial orientation resulting from the binding of two compatible fragments to MAG can a high-affinity ligand be generated by a receptor-mediated triazole formation (Fig. 8d) [111–118].

Concise LC-MS analysis of the supernatant (SIM- and MS/MS-mode) indicated the formation of the *syn*- and *anti*-triazoles **185** from **183a** and **184c** (Fig. 10). Since a further differentiation between the isomers was not possible, *syn*- and *anti*-**185** were synthesized under thermal conditions and biologically tested. In a Biacore experiment K_{DS} of 190 nM for *anti*-**185** and 2 μ M for *syn*-**185** were obtained, leading to the assumption that *anti*-**185** was formed in the in situ click experiment. With *anti*-**185** a nanomolar MAG antagonist exhibiting a 1,000-fold improved affinity compared to the tetrasaccharide epitope of GQ1b α (**67**, K_D : 180 μ M [68]) was identified.

The analysis of the binding epitope of *anti*-**185** by STD NMR revealed large contributions to binding from both terminal aromatic groups, the benzamide and the 5-nitroindole moiety (Fig. 11) [101]. In addition, the triazole linking the first- and second-site ligand also contributes to the interaction with MAG. This is consistent with previous findings that hydrophobic aglycones on sialosides led to an enhanced affinity (Table 7).

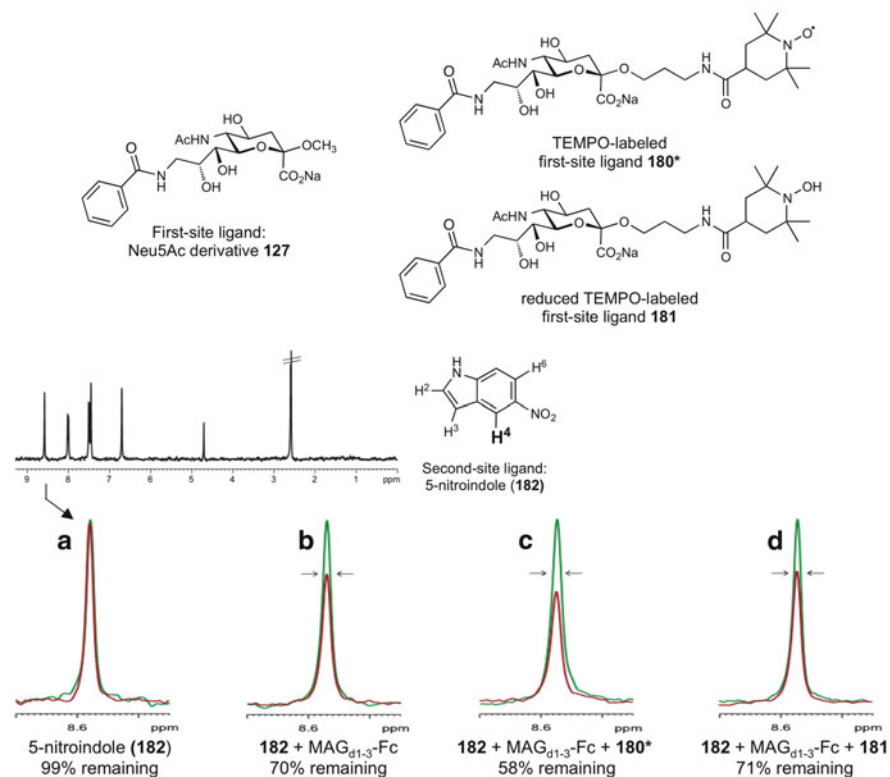


Fig. 9 Principle of second-site screening [101, 102]. (a) ^1H NMR resonance of H-C(4) of 5-nitroindole (**182**) recorded at two relaxation times, 10 ms (*red*) and 200 ms (*green*). (b) Decay of transverse magnetization caused by binding of **182** to MAG. (c) Paramagnetic relaxation enhancement on **182** caused by the spin-labeled first-site ligand **180***. (d) Recovery of the paramagnetic relaxation enhancement by the addition of ascorbic acid (reducing **180*** to **181**), attributing the paramagnetic relaxation effect unambiguously to the spin label

9 Summary of the Structure Affinity Relationship

The considerable effort which was invested in the last two decades in the search for neuraminic acid-based MAG antagonists not only led to the identification of drug-like and high affinity MAG ligands, but also provided a deeper insight into the structure affinity relationship of carbohydrate recognition by MAG. In Fig. 12 the structure-activity data are compiled for the minimal binding epitope, the tetrasaccharide Neu5Ac- α (2-3)-Gal- β (1-3)-[Neu5Ac- α (2-6)]-GalNAc [45].

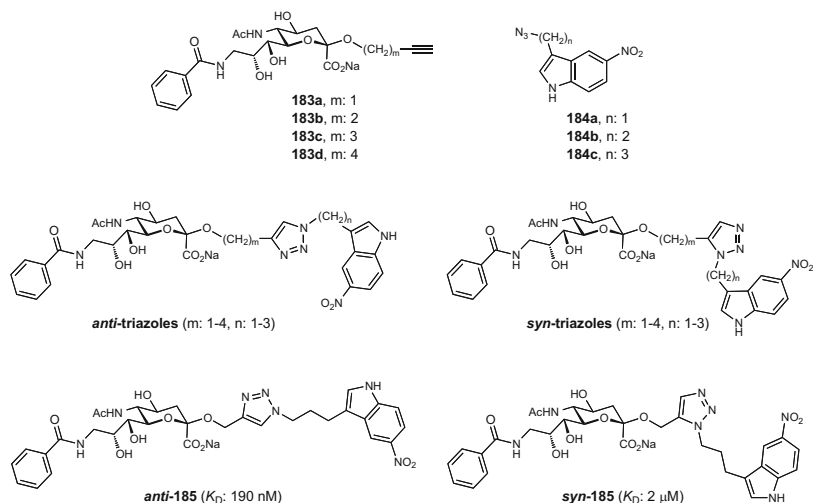


Fig. 10 The library of first-site ligands consisted of four members, **183a–d**, the library of second-site ligands of three, **184a–c**. In total, 24 different triazoles, 12 *anti*-triazoles and 12 *syn*-triazoles could be formed. Concise MS analysis revealed the formation of *anti*-**185** and *syn*-**185** from **183a** and **184c**. Since differentiation between *anti*-**185** and *syn*-**185** was not possible by spectroscopic means, they were both synthesized and biologically evaluated [101]

9.1 The Terminal Neuraminic Acid

The main contribution to binding to MAG stems from the terminal neuraminic acid residue. In general, MAG has a high preference for the Neu5Ac- α (2–3)-Gal epitope [27, 30, 59, 66]. In contrast, a terminal α (2–6)-linked sialic acid only provides for weak interactions, which are obviously not sufficient to mediate stable binding of MAG to glycolipids on cell surfaces [27]. Thus, the spatial orientation of the carboxyl group of the terminal sialic acid is essential for binding. This is further corroborated by docking studies of tetrasaccharide **67** to a homology model of MAG [66] based on the X-ray structure of sialoadhesin (Siglec-1) [26], which show the carboxyl group to interact in a salt bridge with Arg118 in the binding site of MAG.

Several experiments [27, 28, 72, 73, 83] have demonstrated that the hydroxyl groups at C-8 and C-9 represent so-called key polar groups [84], which are essential for binding to MAG, whereas the 7-OH plays a minor role [27]. The 9-OH acts as a donor in a hydrogen bond with Phe129 [70] or Thr128 [83] and its replacement by benzamides [72, 73, 81] strongly enhances affinity through additional interactions of the aromatic substituent with a hydrophobic pocket formed by residues Glu131 and Tyr127 [70, 81]. The hydroxyl group in the 8-position is most probably involved in a hydrogen bond with the backbone NH of Thr128 [81, 83] and its absence leads to a dramatic loss in affinity in oligosaccharides [27] as well as in small molecules like sialoside mimic **179** [83]. For the 4-OH group the results were

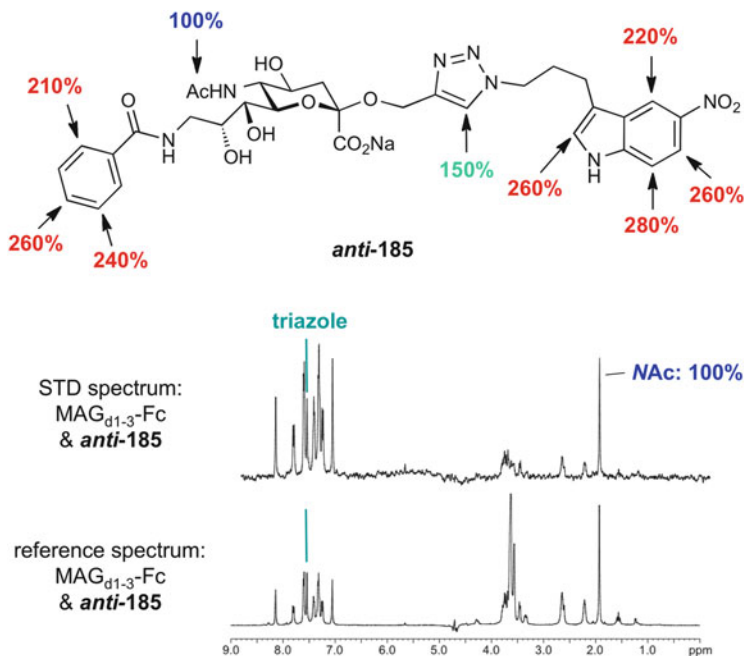


Fig. 11 The binding epitope of *anti-185* was mapped through STD-NMR experiments [101]. The percentages drawn in the structure indicate the level of transfer of magnetization to a particular hydrogen relative to transfer received by the *N*-acetyl methyl group. Large percentages were identified on the benzamide and the nitroindole, thus rationalizing the enhanced affinity of *anti-185*

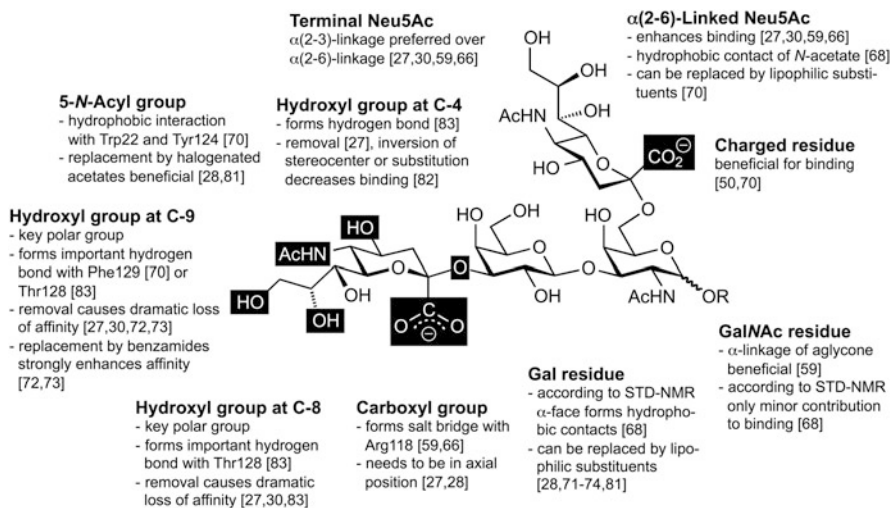


Fig. 12 Functional groups of the binding epitope Neu5Ac-α(2-3)-Gal-β(1-3)-[Neu5Ac-α(2-6)]-GalNAc [45] involved in the binding to MAG. Residues responsible for the main contributions to binding are printed *white* on *black background*

somewhat contradictory: whereas experiments with oligosaccharides showed that the removal of this moiety leads to a strong decrease in affinity [27], docking studies with small sialosides indicated only a minor contribution of the 4-OH [81]. However, the inversion of the stereocenter at C-4, the removal of the 4-OH group, or its substitution by acyl groups or carbamates led to a substantial drop in binding affinity (Table 8) [82].

The investigation of binding of tetrasaccharide **67** to MAG by STD NMR revealed significant interactions originating from the *N*-acetyl function and the protons at C-6 and C-7 indicating a hydrophobic interaction of the α -face of the terminal Neu5Ac with MAG [68]. In addition, by docking **67** to the homology model of MAG, a hydrophobic contact between the acetamido group and the side chains of Trp22 and Tyr124 could be rationalized [70]. Therefore, it is not surprising that acyl substituents have a profound effect on the binding affinity. Generally, sialosides with halogenated acetyl residues at C-5 showed strongly enhanced binding to MAG [28, 81]. The enhanced affinity correlates with the electronegativity of the halogen atom and can be rationalized by halogen-mediated additional contacts with the protein or by electronic effects on the amide.

9.2 The Subterminal Galactose

The results of the MAG binding assays imply that not only the terminal Neu5Ac residue but also the subterminal oligosaccharides contribute to binding, albeit to a lesser extent. Additional substituents at position 6 of the subterminal galactose caused a significant drop in inhibition, whereas the 4- and 6-deoxy compounds showed increased activity (Table 2) [27]. However, these effects may simply be a result of conformational changes of the whole glycan, caused by these modifications, resulting in a less suitable presentation of the terminal Neu5Ac [27]. In addition, the comparably strong STD signals originating from H-3, H-4, and H-5 of the subterminal galactose residue indicated a hydrophobic interaction of the α -face of D-Gal with MAG [68]. Consequently, the whole subterminal saccharide can be replaced by lipophilic aglycones like biphenyl [71] or (substituted) benzyl groups [28, 72–74, 81], giving rise to small and drug-like high-affinity ligands.

9.3 The Third Terminal Monosaccharide

Glycoproteins and glycolipids carry the terminal disaccharide unit, Neu5Ac- α (2–3)-Gal, either β (1–3)-linked to GalNAc or GlcNAc or β (1–4)-linked to GlcNAc. The exact role of the underlying monosaccharide residue in binding to MAG remains unclear, since the available experimental data is partly contradictory. Schnaar and coworkers (Table 1) found a strong decrease of inhibition if the terminal Neu5Ac- α (2–3)-Gal unit was β (1–6)-linked to GalNAc or β (1–4)-linked

to GlcNAc [30]. On the other hand, Kelm et al. (Table 2) observed only minor changes in activity for structures with the terminal disaccharide $\beta(1-3)$ - or $\beta(1-4)$ -linked to GalNAc or GlcNAc [27]. These results are in good agreement with the comparatively low STD values found for the protons of the GalNAc residue of tetrasaccharide **67**, indicating that no significant contribution to binding originates from this part of the molecule [68]. However, according to the results of Paulson et al. [59, 65], α -linkage of aglycones to the GalNAc is highly preferred over β -linkage, leading to a strong increase in affinity for MAG. In fact, the disialyl-T antigen **58** (Tables 3 and 4) bearing an α -linked *N*-acetyl-threonine ester was found to cause statistically significant reversal of MAG-mediated inhibition of axon outgrowth from rat cerebellar granule neurons in vitro [65].

9.4 The Internal $\alpha(2-6)$ -Linked Neuraminic Acid

An additional sialic acid at position 6 of the GalNAc residue significantly enhances inhibition of MAG [27, 30, 59, 66]. This residue may be accommodated by a water-filled cavity, which is visible in the crystal structure of the N-terminal domain of the closely related sialoadhesin (Siglec-1) in complex with 3'-sialyllactose [26]. In addition, docking of tetrasaccharide **67** to a homology model of MAG [66] based on this X-ray structure indicated that the carboxylic acid of the $\alpha(2-6)$ -linked sialic acid forms a salt bridge with Lys67. The importance of a negative charge at this position becomes obvious by the fact that the $\alpha(2-6)$ -linked Neu5Ac can be replaced either by sulfate [30, 50] or by lactate [70] without loss in binding affinity. Because the only sizeable effect detected for the $\alpha(2-6)$ -linked Neu5Ac in the STD NMR of **67** [68] stems from the *N*-acetyl group, the sialic acid residue can also be replaced by spatially demanding lipophilic substituents like biphenylmethyl without losing activity [70]. The inspection of the homology model of MAG revealed the presence of a main hydrophobic pocket lined by the side chains of Trp59, Tyr60, Tyr69, and Tyr116, which can accommodate this kind of substituent [70]. This hydrophobic pocket is also exploited by the 5-nitroindole moiety of the high-affinity ligand *anti*-**185** [101].

10 Conclusions

From the first description of MAG in the early 1970s, two decades went by before its binding activity for sialylated glycoconjugates was discovered. Now, almost 20 years later, the biological role of the sialic acid binding activity of MAG still remains unclear. Certainly, the availability of high affinity inhibitors binding selectively to the sialic acid binding site of Siglec-4 will allow one to address this question much more directly. It will be very interesting to see, which effects these compounds have on biological systems where MAG is involved. Furthermore, the

route leading from the first description of the selectivity for $\alpha(2-3)$ -linked Neu5Ac to the synthetic sialosides binding with low nanomolar affinity described in this review provides ample information regarding the binding mode of Siglec-4 and innovative approaches, which are also expected to foster the development of high affinity antagonists of other siglecs. Since all siglecs share a similar primary binding site, besides high affinity the selectivity for the targeted siglec will also play an important role to obtain compounds addressing specific siglecs in basic research on their function as well as in the development of drugs.

References

1. Cao H, Crocker PR (2010) *Immunology* 132:18–26
2. Crocker PR, Redelinghuys P (2008) *Biochem Soc Trans* 36:1467–1471
3. Crocker PR, Clark EA, Filbin MT, Gordon S, Jones Y, Kehrl JH, Kelm S, Le Douarin N, Powell L, Roder J, Schnaar RL, Sgroi DC, Stamenkovic K, Schauer R, Schachner M, van den Berg TK, van der Merwe PA, Watt SM, Varki A (1998) *Glycobiology* 8:v
4. Quarles RH, Everly JL, Brady RO (1972) *Biochem Biophys Res Commun* 47:491–497
5. Quarles RH, Everly JL, Brady RO (1973) *Brain Res* 58:506–509
6. Sternberger NH, Quarles RH, Itoyama Y, Webster HD (1979) *Proc Natl Acad Sci U S A* 76:1510–1514
7. Itoyama Y, Sternberger NH, Webster HD, Quarles RH, Cohen SR, Richardson EP (1980) *Ann Neurol* 7:167–177
8. Kruse J, Mailhammer R, Wernecke H, Faissner A, Sommer I, Goridis C, Schachner M (1984) *Nature* 311:153–155
9. Martini R, Schachner M (1988) *J Cell Biol* 106:1735–1746
10. McGarry RC, Helfand SL, Quarles RH, Roder JC (1983) *Nature* 306:376–378
11. Poltorak M, Sadoul R, Keilhauer G, Landa C, Fahrigr T, Schachner M (1987) *J Cell Biol* 105:1893–1899
12. Frail DE, Braun PE (1984) *J Biol Chem* 259:14857–14862
13. Lai C, Brow MA, Nave KA, Noronha AB, Quarles RH, Bloom FE, Milner RJ, Sutcliffe JG (1987) *Proc Natl Acad Sci U S A* 84:4337–4341
14. Salzer JL, Holmes WP, Colman DR (1987) *J Cell Biol* 104:957–965
15. Li C, Tropak MB, Gerlai R, Clapoff S, Abramow-Newerly W, Trapp B, Peterson A, Roder J (1994) *Nature* 369:747–750
16. Montag D, Giese KP, Bartsch U, Martini R, Lang Y, Blüthmann H, Karthigasan J, Kirschner DA, Wintergerst ES, Nave KA (1994) *Neuron* 13:229–246
17. Umemori H, Sato S, Yagi T, Aizawa S, Yamamoto T (1994) *Nature* 367:572–576
18. Kelm S, Pelz A, Schauer R, Filbin MT, Tang S, de Bellard M-E, Schnaar RL, Mahoney JA, Hartnell A, Bradfield P, Crocker PR (1994) *Curr Biol* 4:965–972
19. McKerracher L, David S, Jackson DL, Kottis V, Dunn RJ, Braun PE (1994) *Neuron* 13:805–811
20. Mukhopadhyay G, Doherty P, Walsh FS, Crocker PR, Filbin MT (1994) *Neuron* 13:757–767
21. Schnaar RL (2010) *FEBS Lett* 584:1741–1747
22. Quarles RH, Barbarash GR, Figlewicz DA, McIntyre LJ (1983) *Biochim Biophys Acta* 757:140–143
23. Kelm S, Ravindran R (2007) *Comprehensive Glycoscience* 3:523–538
24. Tropak MB, Roder JC (1997) *J Neurochem* 68:1753–1763
25. Lehmann F (2004) *Glycobiology* 14:959–968
26. May AP, Robinson RC, Vinson M, Crocker PR, Jones EY (1998) *Mol Cell* 1:719–728

27. Strengé K, Schauer R, Bovin N, Hasegawa A, Ishida H, Kiso M, Kelm S (1998) *Eur J Biochem* 258:677–685
28. Kelm S, Brossmer R, Isecke R, Gross H-J, Strengé K, Schauer R (1998) *Eur J Biochem* 255:663–672
29. Kelm S, Schauer R, Manuguerra J-C, Gross H, Crocker PR (1994) *Glycoconjugate J* 11:576–585
30. Collins BE, Ito H, Sawada N, Ishida H, Kiso M, Schnaar RL (1999) *J Biol Chem* 274:37637–37643
31. Collins BE, Yang LJS, Mukhopadhyay G, Filbin MT, Kiso M, Hasegawa A, Schnaar RL (1997) *J Biol Chem* 272:1248–1255
32. Strengé K, Brossmer R, Ihrig P, Schauer R, Kelm S (2001) *FEBS Lett* 499:262–267
33. Strengé K, Schauer R, Kelm S (1999) *FEBS Lett* 444:59–64
34. Vyas AA, Patel HV, Fromholt SE, Heffer-Lauc M, Vyas KA, Dang J, Schachner M, Schnaar RL (2002) *Proc Natl Acad Sci U S A* 99:8412–8417
35. Akbik F, Cafferty WBJ, Strittmatter SM (2012) *Exp Neurol* 235:43–52
36. Lee JK, Zheng B (2012) *Exp Neurol* 235:33–42
37. Cao Z, Gao Y, Deng K, Williams G, Doherty P, Walsh FS (2010) *Mol Cell Neurosci* 43:1–14
38. Quarles RH (2007) *J Neurochem* 100:1431–1448
39. Quarles RH (2008) *Neurochem Res* 34:79–86
40. Weiss MD, Hammer J, Quarles RH (2000) *J Neurosci Res* 62:772–780
41. Yin X, Crawford TO, Griffin JW, Tu P-h, Lee VMY, Li C, Roder J, Trapp BD (1998) *J Neurosci* 18:1953–1962
42. Mehta NR, Nguyen T, Bullen JW Jr, Griffin JW, Schnaar RL (2010) *ACS Chem Neurosci* 1:215–222
43. Lopez PHH, Ahmad AS, Mehta NR, Toner M, Rowland EA, Zhang J, Doré S, Schnaar RL (2011) *J Neurochem* 116:900–908
44. Mountney A, Zahner MR, Lorenzini I, Oudega M, Schramm LP, Schnaar RL (2010) *Proc Natl Acad Sci U S A* 107:11561–11566
45. Crocker PR, Kelm S, Hartnell A, Freeman S, Nath D, Vinson M, Mucklow S (1996) *Biochem Soc Trans* 24:150–156
46. Vyas AA, Schnaar RL (2001) *Biochimie* 83:677–682
47. Yang LJS, Zeller CB, Shaper NL, Kiso M, Hasegawa A, Shapiro RE, Schnaar RL (1996) *Proc Natl Acad Sci U S A* 93:814–818
48. Collins BE, Kiso M, Hasegawa A, Tropak MB, Roder JC, Crocker PR, Schnaar RL (1997) *J Biol Chem* 272:16889–16895
49. Sheikh KA, Sun J, Liu Y, Kawai H, Crawford TO, Proia RL, Griffin JW, Schnaar RL (1999) *Proc Natl Acad Sci U S A* 96:7532–7537
50. Ito H, Ishida H, Collins BE, Fromholt SE, Schnaar RL, Kiso M (2003) *Carbohydr Res* 338:1621–1639
51. Prabhanjan H, Kameyama A, Ishida H, Kiso M, Hasegawa A (1991) *Carbohydr Res* 220:127–143
52. Prabhanjan H, Aoyama K, Kiso M, Hasegawa A (1992) *Carbohydr Res* 233:87–99
53. Hotta K, Ishida H, Kiso M, Hasegawa A (1995) *J Carbohydr Chem* 14:491–506
54. Ito H, Ishida H, Kiso M, Hasegawa A (1998) *Carbohydr Res* 306:581–585
55. Sawada N, Ishida H, Collins BE, Schnaar RL, Kiso M (1999) *Carbohydr Res* 316:1–5
56. Richardson PJ, Walker JH, Jones RT, Whittaker VP (1982) *J Neurochem* 38:1605–1614
57. Hirabayashi Y, Nakao T, Irie F, Whittaker VP, Kon K, Ando S (1992) *J Biol Chem* 267:12973–12978
58. Bock N, Kelm S (2006) Binding and inhibition assay for siglecs. In: Brockhausen I (ed) *Methods in molecular biology*, vol 347. Humana, Totowa, pp 359–376
59. Blixt O, Collins BE, van den Nieuwenhof IM, Crocker PR, Paulson JC (2003) *J Biol Chem* 278:31007–31019
60. Blixt O, Brown J, Schur MJ, Wakarchuk W, Paulson JC (2001) *J Org Chem* 66:2442–2448
61. Blixt O, Allin K, Pereira L, Datta A, Paulson JC (2002) *J Am Chem Soc* 124:5739–5746
62. Blixt O, Paulson JC (2003) *Adv Synth Catal* 345:687–690

63. Lüning B, Norberg T, Tejbrant J (1989) *Glycoconjugate J* 6:5–19
64. Lüning B, Norberg T, Rivera-Baeza C, Tejbrant J (1991) *Glycoconjugate J* 8:450–455
65. Vyas AA, Blixt O, Paulson JC, Schnaar RL (2005) *J Biol Chem* 280:16305–16310
66. Bhunia A, Schwardt O, Gäthje H, Gao G-P, Kelm S, Benie AJ, Hricovini M, Peters T, Ernst B (2008) *ChemBioChem* 9:2941–2945
67. Schwardt O, Gao G-P, Visekruna T, Rabbani S, Gassmann E, Ernst B (2004) *J Carbohydr Chem* 23:1–26
68. Shin S-Y, Gäthje H, Schwardt O, Gao G-P, Ernst B, Kelm S, Meyer B (2008) *ChemBioChem* 9:2946–2949
69. Sato S, Fujita N, Kurihara T, Kuwano R, Kakahashi Y, Miyatake T (1989) *Biochem Biophys Res Commun* 163:1473–1480
70. Schwardt O, Gäthje H, Vedani A, Mesch S, Gao G-P, Spreafico M, von Orelli J, Kelm S, Ernst B (2009) *J Med Chem* 52:989–1004
71. Schwizer D, Gäthje H, Kelm S, Porro M, Schwardt O, Ernst B (2006) *Bioorg Med Chem* 14:4944–4957
72. Kelm S, Brossmer R (2003) PCT patent WO 03/000709A2
73. Shelke SV, Gao G-P, Mesch S, Gäthje H, Kelm S, Schwardt O, Ernst B (2007) *Bioorg Med Chem* 15:4951–4965
74. Gao G, Smiesko M, Schwardt O, Gäthje H, Kelm S, Vedani A, Ernst B (2007) *Bioorg Med Chem* 15:7459–7469
75. Powell LD, Varki A (1995) *J Biol Chem* 270:14243–14246
76. Sjöberg ER, Powell LD, Klein A, Varki A (1994) *J Cell Biol* 126:549–562
77. Topliss JG (1972) *J Med Chem* 15:1006–1011
78. Topliss JG (1977) *J Med Chem* 20:463–469
79. Hansch C, Fujita T (1964) *J Am Chem Soc* 86:1616–1626
80. Hansch C (1971) Quantitative structure-activity relationships in drug design. In: Ariens EJ (ed) *Drug design*, vol 1. Academic, New York, pp 271–342
81. Mesch S, Moser D, Strasser DS, Kelm A, Cutting B, Rossato G, Vedani A, Koliwer-Brandl H, Wittwer M, Rabbani S, Schwardt O, Kelm S, Ernst B (2010) *J Med Chem* 53:1597–1615
82. Mesch S, Lemme K, Koliwer-Brandl H, Strasser DS, Schwardt O, Kelm S, Ernst B (2010) *Carbohydr Res* 345:1348–1359
83. Schwardt O, Koliwer-Brandl H, Zimmerli R, Mesch S, Rossato G, Spreafico M, Vedani A, Kelm S, Ernst B (2010) *Bioorg Med Chem* 18:7239–7251
84. Lemieux RU (1989) *Chem Soc Rev* 18:347–374
85. Swinney DJ (2008) *PharmMed* 22:23–34
86. Copeland RA, Pompliano DL, Meek TD (2006) *Nat Rev Drug Disc* 5:730–739
87. Markgren P-O, Schaal W, Hämäläinen M, Karlén A, Hallberg A, Samuelsson B, Danielson UH (2002) *J Med Chem* 45:5430–5439
88. Shuman CF, Hämäläinen MD, Danielson UH (2004) *J Mol Recognit* 17:106–119
89. Mehta P, Cummings RD, McEver RP (1998) *J Biol Chem* 273:32506–32513
90. Wild MK, Huang MC, Schulze-Horsel U, van der Merwe PA, Vestweber D (2001) *J Biol Chem* 276:31602–31612
91. Herfurth L, Ernst B, Öhrlein R, Wagner B, Ricklin D, Strasser DS, Magnani JL, Benie AJ, Peters T (2005) *J Med Chem* 48:6879–6886
92. Regan J, Pargellis CA, Cirillo PF, Gilmore T, Hickey ER, Peet GW, Proto A, Swinamer A, Moss N (2003) *Bioorg Med Chem Lett* 13:3101–3104
93. Kati WM, Montgomery D, Carrick R, Gubareva L, Maring C, McDaniel K, Steffy K, Molla A, Hayden F, Kempf D, Kohlbrenner W (2002) *Antimicrob Agents Chemother* 46:1014–1021
94. Rich RL, Myszka DG (2000) *Curr Opin Biotechnol* 11:54–61
95. Morton TA, Myszka DG (1998) *Methods Enzymol* 295:268–294
96. Nagata K, Handa H (2000) *Real-time analysis of biomolecular interactions: applications of Biacore*. Springer, Berlin

97. Hajduk PJ, Greer J (2007) *Nat Rev Drug Discov* 6:211–219
98. Chessari G, Woodhead AJ (2009) *Drug Discov Today* 14:668–675
99. Murray CW, Rees DC (2009) *Nature Chem* 1:187–192
100. Finkelstein AV, Janin J (1989) *Protein Eng* 3:1–3
101. Shelke SV, Cutting B, Jiang X, Koliwer-Brandl H, Strasser DS, Schwardt O, Kelm S, Ernst B (2010) *Angew Chem Int Ed* 49:5721–5725
102. Ernst B, Cutting B, Shelke SV (2006) PCT/WO2007/105094A1
103. Otting G (1993) *Curr Opin Struct Biol* 3:760–768
104. Hajduk PJ, Olejniczak ET, Fesik SW (1997) *J Am Chem Soc* 119:12257–12261
105. Jahnke W, Perez LB, Paris CG, Strauss A, Fendrich G, Nalin CM (2000) *J Am Chem Soc* 122:7394–7395
106. Jahnke W, Rüdissler S, Zurini M (2001) *J Am Chem Soc* 123:3149–3150
107. Vazquez J, Tautz L, Ryan JJ, Vuori K, Mustelin T, Pellecchia M (2007) *J Med Chem* 50:2137–2143
108. Jain NU, Venot A, Umemoto K, Leffler H, Prestegard JH (2001) *Protein Sci* 10:2393–2400
109. Bertini I, Fragai M, Lee Y-M, Luchinat C, Terni B (2004) *Angew Chem Int Ed* 43:2254–2256
110. Mammen F, Shakhnovich EI, Whitesides GM (1998) *J Org Chem* 63:3168–3175
111. Lewis WG, Green LG, Grynszpan F, Radic Z, Carlier PR, Taylor P, Finn MG, Sharpless KB (2002) *Angew Chem* 114:1095–1099
112. Lewis WG, Green LG, Grynszpan F, Radic Z, Carlier PR, Taylor P, Finn MG, Sharpless KB (2002) *Angew Chem Int Ed* 41:1053–1057
113. Manetsch R, Krasinski A, Radic Z, Raushel J, Taylor P, Sharpless KB, Kolb HC (2004) *J Am Chem Soc* 126:12809–12818
114. Mocharla VP, Colasson B, Lee LV, Roeper S, Sharpless KB, Wong C-H, Kolb HC (2004) *Angew Chem* 116:118–122
115. Mocharla VP, Colasson B, Lee LV, Roeper S, Sharpless KB, Wong C-H, Kolb HC (2005) *Angew Chem Int Ed* 44:116–120
116. Krasinski A, Radic Z, Manetsch R, Raushel J, Taylor P, Sharpless KB, Kolb HC (2005) *J Am Chem Soc* 127:6686–6692
117. Whiting M, Muldoon J, Lin Y-C, Silverman SM, Lindstrom W, Olson AJ, Kolb HC, Finn MG, Sharpless KB, Elder JH, Fokin VV (2006) *Angew Chem* 118:1463–1467
118. Whiting M, Muldoon J, Lin Y-C, Silverman SM, Lindstrom W, Olson AJ, Kolb HC, Finn MG, Sharpless KB, Elder JH, Fokin VV (2006) *Angew Chem Int Ed* 46:1435–1439

Chemical Approach to a Whole Body Imaging of Sialo-*N*-Linked Glycans

Katsunori Tanaka and Koichi Fukase

Abstract PET and noninvasive fluorescence imaging of the sialo-*N*-linked glycan derivatives are described. To establish the efficient labeling protocol for *N*-glycans and/or glycoconjugates, new labeling probes of fluorescence and ^{68}Ga -DOTA, as the positron emission nucleus for PET, through rapid 6π -azaelectrocyclization were designed and synthesized, (*E*)-ester aldehydes. The high reactivity of these probes enabled the labeling of lysine residues in peptides, proteins, and even amino groups on the cell surfaces at very low concentrations of the target molecules ($\sim 10^{-8}$ M) within a short reaction time (~ 5 min) to result in “selective” and “non-destructive” labeling of the more accessible amines. The first MicroPET of glycoproteins, ^{68}Ga -DOTA-orosomuroid and asialoorosomuroid, successfully visualized the differences in the circulatory residence of glycoproteins, in the presence or absence of sialic acids. In vivo dynamics of the new *N*-glycoclusters, prepared by the “self-activating” Huisgen cycloaddition reaction, could also be affected significantly by their partial structures at the non-reducing end, i.e., the presence or absence of sialic acids, and/or sialoside linkages to galactose. Azaelectrocyclization chemistry is also applicable to the engineering of the proteins and/or the cell surfaces by the oligosaccharides; lymphocytes chemically engineered by sialo-*N*-glycan successfully target the tumor implanted in BALB/C nude mice, detected by noninvasive fluorescence imaging.

Keywords 6π -Azaelectrocyclization • Asparagine-linked glycan (*N*-glycan) • Cell surface engineering • Fluorescence • Glycocluster • Glycoprotein • Labeling, PET (positron emission tomography) • Lymphocytes • Lysine • Tumor targeting

K. Tanaka (✉) and K. Fukase

Department of Chemistry, Graduate School of Science, Osaka University, 1-1
Machikaneyama, Toyonaka, Osaka 560-0043, Japan

e-mail: ktzenori@chem.sci.osaka-u.ac.jp; koichi@chem.sci.osaka-u.ac.jp

Contents

1	Introduction	202
2	Azaelectrocyclization-Based Labeling of Lysines; New Microgram-Scale Procedure for Molecular Imaging of Glycan Derivatives	203
3	PET of Orosomucoid and Asialoorosomucoid: First Visualization of Sialic Acid-Dependent Circulatory Residence of Glycoproteins	209
4	Noninvasive Imaging of Dendrimer-Type <i>N</i> -Glycan Clusters: Remarkable In Vivo Dynamics Dependence on Non-Reducing End Structures of <i>N</i> -Glycans	210
5	Noninvasive Fluorescence Imaging of <i>N</i> -Glycan-Engineered Living Cells: Effects of <i>N</i> -Glycans on Lymphocytes for Tumor Targeting	218
5.1	Fluorescence Labeling of Living Cell Surfaces by Azaelectrocyclization	218
5.2	Chemical Engineering of Cell Surfaces by <i>N</i> -Glycans and Other Functional Molecules	221
5.3	Noninvasive Fluorescence Imaging of Lymphocytes Trafficking and Effects of Cell-Surface Engineering by <i>N</i> -Glycan for Tumor Targeting	222
6	Conclusion	225
	References	227

1 Introduction

Among the various types of oligosaccharide structures, asparagine-linked oligosaccharides (*N*-glycans) are the most prominent in terms of diversity and complexity. In particular, *N*-glycans containing sialic acid residues are involved in a variety of important physiological events, including cell-cell recognition, adhesion, signal transduction, and quality control [1]. Moreover, it has long been known that sialic acids in *N*-glycans on soluble proteins or peptides enhance circulatory residence [2–4], i.e., *N*-glycan-engineered erythropoietin (EPO) [3] or insulin [4] exhibits a remarkably higher stability in serum, which effects prolonged bioactivity. Antibody-dependent cellular cytotoxicity (ADCC) and/or complement-dependent cytotoxicity (CDC) has also been proposed to be modulated by the sialic acids of *N*-glycans in immunoglobulin (IgG) through Siglec interactions by glycosylating or removing the sialic acids [5]. However, these important findings and previous efforts in investigating *N*-glycan functions have been mostly based on in vitro experiments using isolated lectins, cultured cells, and dissected tissues. Recently, interest has shifted to the dynamics of these glycans in vivo; molecular imaging is the most promising tool to visualize the “on-time” *N*-glycan dynamics in vivo [6, 7]. Recently, biomolecule-based tracers such as peptides and antibodies have actively been investigated to visualize in vivo dynamics and target specific organs [8, 9]. Although fluorescence imaging is the method of choice because of the convenient experimental and detection procedures at small animal levels, magnetic resonance (MR) and positron emission tomography (PET) imaging, which have technologically improved sensitivity, are well-suited for diagnostic applications. A successful example is ^{18}F -FDG (2-deoxy2- ^{18}F]fluoro-D-glucose), technically a

monosaccharide, which is a Food and Drug Administration-approved PET-radio-pharmaceutical [10, 11].

Nevertheless, the challenge in efficient glycan imaging in living animals is to obtain the structurally pure oligosaccharides either from nature or by synthetic methods [12 and references cited therein]. In addition, the bioactivity of oligosaccharides might be derived from the multivalency and/or heterogeneous environment, i.e., on cell surfaces that are composed of oligosaccharide clusters along with other biomolecules. Therefore, efficiently mimicking and labeling such an *N*-glycan-involved bio-environment, e.g., by conjugating the *N*-glycans to the liposomes, to the clusters, or even to the surface of the cells, may provide information on the “in vivo dynamics” of *N*-glycans; a single molecule of the *N*-glycan, obtained from natural or synthetic sources, is readily excreted from the body [13, 14]. Recent successful non-invasive imaging and biodistribution study of glycans and/or glycoconjugates dealt with natural- and neo-glycoproteins, liposomes, and nanoparticles. For representative examples of noninvasive imaging and biodistribution experiments, see [15–27].

In this personal account, we want to describe our recent trials on the in vivo imaging of the sialo-*N*-glycans; we first describe (1) the new chemistry-based labeling of the amino groups [28–30], e.g., lysines or ethanol amines, which is quite useful not only for the labeling of the glycans and/or the glycoconjugates, but also generally for the peptides, proteins, as well as the cell surfaces [31]. Then, PET and the noninvasive fluorescence images of (2) sialo- and asialo-glycoproteins [8, 9, 28–30], (3) dendrimer-type glycoclusters consisting of 16 molecules of *N*-glycans [32], and (4) chemically engineered lymphocytes by sialo-*N*-glycans [33] are discussed. These imaging studies clearly visualize the remarkable dependence of in vivo dynamics and organ-specific accumulation on the partial structures of the non-reducing end of the *N*-glycans.

2 Azaelectrocyclization-Based Labeling of Lysines; New Microgram-Scale Procedure for Molecular Imaging of Glycan Derivatives

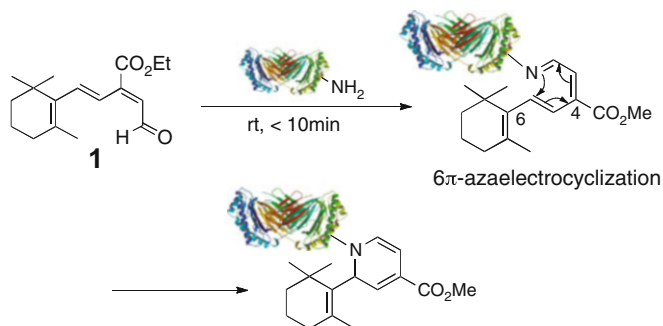
Because of the ability to visualize and analyze the in vivo dynamics of the target molecules in living animals, molecular imaging is an important topic which has garnered significant attention in the fields of chemical biology, drug discovery, and diagnosis [6–11]. As noted previously, positron emission tomography (PET) is a common noninvasive and diagnostically applicable method, which quantitatively visualizes the locations and levels of radiotracer accumulation with high imaging contrast. Generally, PET imaging of the “biomolecules,” including glycoproteins and other glycoconjugates, is achieved by conjugation to metal chelating agents such as DOTA (1,4,7,10-tetraazacyclododecane-1,4,7,10-tetraacetic acid) or DTPA (diethylenetriaminepentaacetic acid) prior to radiolabeling with a suitable β^+

emitting metal [8, 9]. Chelating agents can be introduced, if possible, during the synthesis of peptides or oligosaccharides, or more directly, and in most cases, via a reaction of lysine residues or *N*-terminal amino groups with DOTA-*O*-sulfosuccinimidyl ester. However, the latter method, which uses activated DOTA ester, usually proceeds slowly and requires a large quantity of sample, in worst cases as much as a few milligrams, to maintain high reaction concentrations. Additionally, the conjugate efficiency is not high, and, more importantly, labeling under such high reagent concentrations (10^{-3} – 10^{-4} M) and long reaction times (a few hours) indiscriminately modifies the key lysines for the activity. Because PET experiments require only small amounts of tracers, i.e., picograms to micrograms, and important biomolecules are occasionally obtained only in small amounts, a subpicogram-scale and non-destructive conjugation methodology would greatly expand the applicability of PET more generally to other imaging research.

Under such circumstances in the molecular imaging and chemical biology fields, various bioorthogonal methods employing advanced organic reactions, including Sharpless/Meldal click reaction (for selected examples of protein labeling, see [34–50]) and Bertozzi's Staudinger [51–60] and strain-releasing click ligation [25, 61–71], have been reported. Nevertheless, an efficient, rapid, and mild protocol not requiring bioengineering techniques and/or time-consuming procedures [63, 66–68] and is widely applicable for anyone in various buffer solutions had yet to be discovered. It is most preferable if the labeling can be performed simply by mixing a minuscule amount of the target with the probe solution for a few minutes at ambient temperature; bioorthogonality is not involved unless the native activity of the labeled molecules is affected.

We have recently found that unsaturated (*E*)-ester aldehydes **1** (Scheme 1) quantitatively reacted with primary amines, including lysine, within a very short time (~5 min) in an extremely diluted organic buffer solution (10^{-7} – 10^{-8} M) to yield 1,2-dihydropyridines as irreversible products through the “accelerated” 6- π -azaelectrocyclization of the intermediary Schiff base [72, 73]. To the best of our knowledge, our reaction where unsaturated aldehyde **1**, for instances, reacts with enzyme to modify several Lys residues within 15 min at room temperature is the fastest conjugation reaction with lysines in water (Fig. 1). We wondered if this reaction would work as a new protocol for Lys labeling.

Inspired by the notable electrocyclization reactivity during our enzyme inhibition study [72, 73] and in synthetic applications [74–81], a new DOTA-labeling probe, DOTA-(*E*)-ester aldehyde **1a**, was designed and synthesized (Scheme 2) [28–30]. Thus, THP-protected (*Z*)-vinyl bromide **2** and glycine linked (*E*)-stannane **3** were heated to 110°C in the presence of Pd₂(dba)₃, P(2-furyl)₃, and LiCl in DMF to provide the Stille coupling product, which was subsequently treated with 3 M hydrochloric acid to give aminoalcohol **4** in 73% yield in two steps. Currently, compound **4** can be prepared on a 50-g scale. This scaled up reaction allowed us to develop a complete electrocyclization-labeling process as a labeling kit “*Stella*⁺” [30], which can be utilized for bio-directed research with broader applications (see below). Selective acylation of the amino group in **4** was achieved in 58% yield by the reaction with DOTA-*O*-succinimidyl ester in the presence of Et₃N in DMF.



Scheme 1 Lysine modification by 6 π -aza-electrocyclization

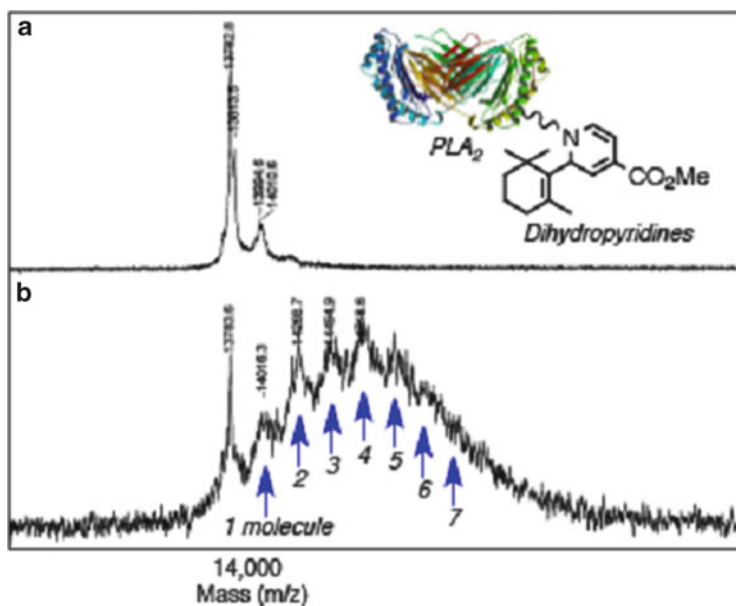
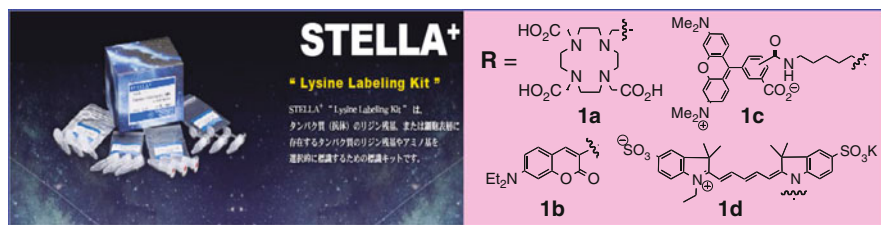
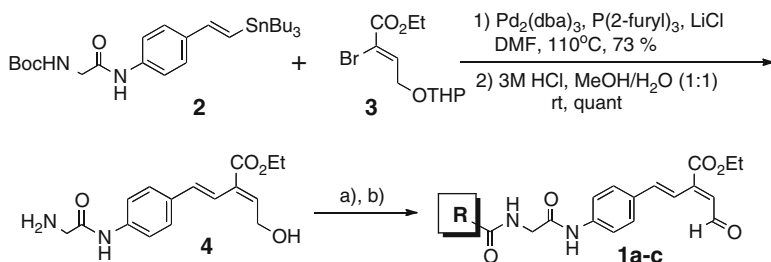


Fig. 1 Formation of dihydropyridines with lysines of PLA₂ through aza-electrocyclization. (a) Intact PLA₂. (b) Modified PLA₂ by **1** for 15 min at 40°C

Finally, the allylic alcohol was oxidized by Dess–Martin periodinane in DMF/CH₂Cl₂ (3:10) to afford the desired aldehyde **1a**, which was immediately used for labeling in a DMF solution after size-partitioning filtration and removal of CH₂Cl₂. Moreover, the established route in Scheme 2 can also introduce various functionalities on the amino group in **4** to provide general probes easily. For example, coumarin, TAMRA, and Cy5 fluorescence probes **1b–d** as well as other imaging reporters or biologically useful tags have been prepared [28–30].



Scheme 2 Preparation of lysine-labeling probes. DOTA-OSu, Et₃N, DMF, rt, 58%; coumarin-OSu, CH₂Cl₂, rt, 68%; TAMRA-OSu, DMF/CH₂Cl₂ (1:1), rt, 74%; Cy5-OSu, DMF/CH₂Cl₂ (1:1), rt, 60%; Dess-Martin periodinane, DMF/CH₂Cl₂ (3:10), rt or polystyrene-supported IBX, DMF/CH₂Cl₂ (1:1), rt

With probes **1a–d** in hand, reactivity toward lysine residues of various biomolecules, namely, somatostatin as a peptide, albumin, orosomucoïd, and anti-GFP antibody as proteins, were tested (Table 1) [28–30]. To our satisfaction, HPLC analysis indicated that the reaction of somatostatin (170 µg) with 100 equiv. of DOTA probe **1a** at room temperature for 30 min provided mono-DOTA-labeled somatostatin in 96% yield (entry 1). Interestingly, of the two lysines in somatostatin, only the lysine not involved in receptor binding was modified under these labeling conditions. In contrast, conventional labeling conditions using the DOTA-activated ester required a much longer reaction time (over 24 h) compared to the conditions in entry 1, and employing the DOTA-activated ester indiscriminantly modified both lysine residues and the *N*-terminus, and the intact peptide was also recovered. Thus, the high reactivity of **1a** enabled lysine-azaelectrocyclization modification to be completed in a very short time, resulting in selective labeling of the more accessible lysine residue. Similarly, incubation of human serum albumin (100 µg) with 25 equiv. of probe **1a** for just 10 min followed by quick size-partitioning gel filtration successfully provided the modified protein that incorporated five DOTA units (entry 2). The present labeling via azaelectrocyclization proceeded in nearly quantitative yield at the lysine residues of all target protein in the reaction mixture at both high ($>10^{-3}$ M) and low (10^{-5} – 10^{-6} M) concentrations of probe **1a**; namely, the target protein in the solution can homogeneously be labeled. Consequently, the number of DOTA units introduced into a protein can be precisely controlled by adjusting the number of equivalents of **1a**. For example,

Table 1 Labeling efficiency by azaelectrocyclization

1a : R = DOTA **1c** : R = TAMRA
1b : R = coumarin

Entry	Probe	Biomolecule (μg)	Conc. of biomolecule (M)	Equiv. of probe	t (min)	Number of labeled Lys ^b
1 ^c	1a	Somatostatin (170)	5.5×10^{-4}	100	30	1
2	1a	Albumin (100)	5.0×10^{-5}	25	10	5
3	1a	Albumin (100)	5.0×10^{-5}	10	10	2
4	1a	Albumin (100)	5.0×10^{-5}	500	10	20
5	1a	Orosomucoïd (62)	4.5×10^{-6}	10	30	2
6	1a	Asialo-orosomucoïd (62)	4.5×10^{-6}	10	30	3
7 ^c	1a (Gd) ^d	Somatostatin (32)	3.6×10^{-4}	31	30	1
8	1b	Albumin (120)	1.7×10^{-5}	7	30	1 ^e
9	1c	Albumin (12)	2.1×10^{-5}	7	30	1 ^e
10	1c	Anti-GFP mAb (2.0)	1.1×10^{-6}	20	30	2 ^e

^aOtherwise noted, reactions were performed in 0.1 M phosphate buffer (pH 7.4) at 24°C

^bCalculated by g-counter of ⁵⁷Co introduced to DOTA according to the method of Meares

^cIn H₂O

^dGd was introduced by the reaction of with 0.1 M GdCl₃ in H₂O

^eEstimated by emission spectra at 470 nm (coumarin) and 555 nm (TAMRA)

under 10 min reaction conditions, the use of 10 equiv. of **1a** provided the labeled protein with two molecules of DOTA, whereas employing 500 equiv. introduced 20 DOTA molecules (entries 3 and 4).

Furthermore, the different efficiencies between our electrocyclization and conventional succinimidyl ester method could be evaluated by labeling various concentrations of HSA (10 min at rt and 10 equiv. of the reagents with respect to the protein). Although one and two molecules of DOTA were introduced into the HSA protein at concentrations of 10^{-6} M and 10^{-7} M, respectively, the conventional method failed with labeled HSA even at a concentration of 10^{-5} M under identical

incubation conditions. Hence, our electrocyclization protocol is 100–1,000 times more efficient.

To display further the performance of this labeling method, proteins, which are available in only small amounts (62 μg of glycoproteins: orosomucoid and asialoorosomucoid), were labeled successfully and $\sim 2\text{--}3$ units of DOTA were incorporated by incubating the respective protein with 10 equiv. of **1a** for 30 min (entries 5 and 6). Additionally, paramagnetic metal Gd^{3+} could be chelated in the DOTA unit of **1a** prior to peptide labeling (entry 7). This new process should allow valuable and/or unstable materials, which might decompose during metal-incorporation to DOTA afterwards, to be labeled.

The present electrocyclization protocol is also applicable to rapid fluorescent labeling; 10–100 μg of albumin and anti-GFP mAb were labeled with coumarin **1b** and TAMRA probes **1c**, respectively (entries 8 and 9). As little as 2 μg of anti-GFP antibody (~ 10 pmol) was successfully labeled by **1c** in 30 min (entry 10); 20 equiv. of **1c** preferentially labeled the more accessible Fc fragment with two fluorophore molecules, while retaining its GFP recognition activity with 90% of the intact mAb.

We would especially like to mention that the entire process of the labeling procedure has been developed into a “labeling-kit” [30]. For a broader and more general applicability of labeling reagents to biology-directed research, an easy-to-use oxidation (activation) method was also developed by using the polystyrene-bound IBX (see Scheme 2); a complete oxidation/labeling process was therefore realized by just mixing, concentrating, and filtrating the stable and therefore “storable” alcohol precursors **1a–d** even at a scale levels of a few milligrams in Eppendorf-tube (see detailed procedure in [31]). Using a simplified “kit” procedure, which is now available as “Stella +”, probe **1a–d** as well as other unsaturated aldehyde probes can be conveniently and reproducibly prepared for labeling use in both organic chemistry and biology laboratories.

Thus, we successfully developed a new lysine-based labeling of biomolecules on the basis of rapid 6π -azaelectrocyclization. Without inactivating the biomolecules, both DOTA as a metal chelating agent (for MRI, PET, or other radiopharmaceutical purposes, such as SPECT with γ -emitters) and fluorescent groups can be efficiently and selectively introduced to lysine residues of samples at levels of a few micrograms in either phosphate or Tris-HCl buffer solutions within 30 min at room temperature. The efficiency of our rapid azaelectrocyclization protocol depends on the steric accessibility of the primary amino groups. Although reactions with internal lysines in protein tertiary structures and *N*-terminal amines (secondary amine) are very slow (>5 h at 24°C) [79], lysines at the protein’s surface react rapidly (10–30 min at 24°C) and, therefore, labeling occurs preferentially at surface positions. Moreover, dihydropyridines as the electrocyclization products, which retain cationic charges as those of the inherent lysines, should contribute to the retention of the protein activity. With efficient labeling protocol to hand, we have investigated PET and noninvasive fluorescence imaging of *N*-glycans and glycoconjugates.

3 PET of Orosomuroid and Asialoorosomuroid: First Visualization of Sialic Acid-Dependent Circulatory Residence of Glycoproteins

MicroPET imaging of orosomuroid and asialoorosomuroid in rabbits as the representative *N*-glycoproteins was then performed to investigate the effects of sialic acid residues on the *in vivo* dynamics of the glycoproteins (Fig. 2) [8, 9, 28–30]. The DOTA-glycoprotein adducts obtained in the above labeling study in Table 1 were treated with $^{68}\text{Ga}^{3+}$ as a β^+ emitter for PET. Approximately, 200 pmol (corresponding to 1 μg) of the $^{68}\text{Ga}^{3+}$ -labeled glycoproteins were then administered from the tail vein of the rabbit prior to whole body scanning over 4 h. The radioactivity of each ^{68}Ga -DOTA-labeled glycoclusters was adjusted to 10 MBq, prior to the injection. PET of these two proteins successfully visualized asialo-glycoprotein being cleared through the kidney and through the liver to the gallbladder faster than orosomuroid (Fig. 2), thus, achieving the first visualization of sialic acid-dependent circulatory residence of glycoproteins. The results agree well with the well-known hypothesis about the clearance of the proteins through the asialoglycoprotein receptor in the liver [2]. That is, the sialic acid residue on the non-reducing end of the *N*-glycans contributes to the stability of glycoproteins in the blood, i.e., EPO [3] or insulin [4], whereas the sialidase-mediated production of

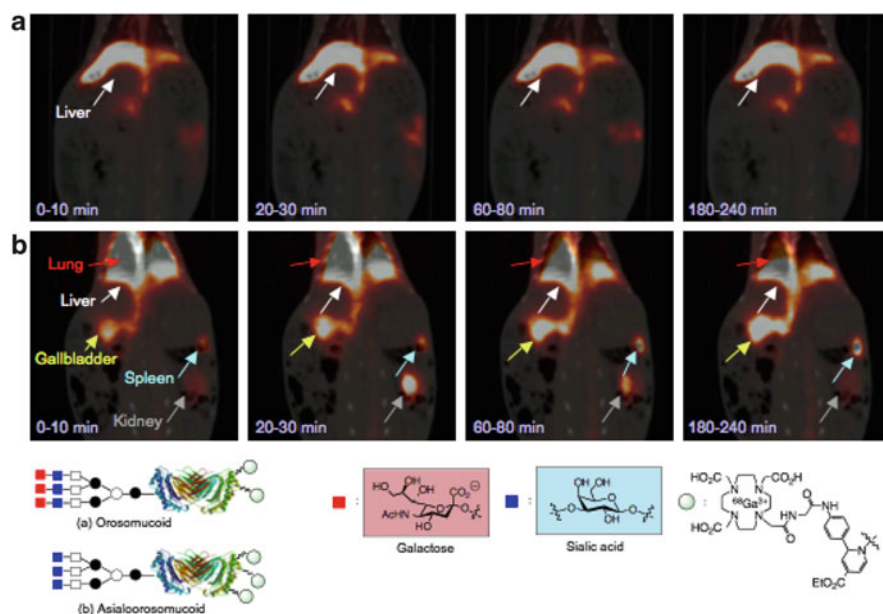


Fig. 2 Dynamic microPET images of [^{68}Ga]DOTA-glycoproteins in rabbits (*axial view*, overlapped with CT). Time course of accumulation of (a) ^{68}Ga -DOTA-orosomuroid and (b) ^{68}Ga -DOTA-asialoorosomuroid (*lower*) in some peripheral organs (*axial views*)

asialoglycoproteins bearing a galactose residue at the non-reducing end of the *N*-glycans is responsible for the metabolic pathway. Furthermore, we observed that the asialoglycoprotein accumulated in the lung and, more interestingly, in the spleen in a time-dependent manner, thus showing a significant difference in biodistribution in the presence and absence of sialic acid residues in the glycoproteins. It is noted that the spleen is an organ located in the abdomen, and it functions in the destruction of old, aged red blood cells as well as holds a reservoir of blood. Moreover, it can function as part of the immune system, such as the reticuloendothelial system (RES) or the production of antigen-specific antibodies, by interacting the T-cells with the mature B-cells. Although it awaits further investigation to rationale the accumulation of asialoglycoprotein in the spleen, the imaging results in Fig. 2 may suggest the negative regulation effects of the sialo-*N*-glycoproteins on the immune system [82–88]. Overall, when combined with our efficient electrocyclization/PET imaging protocol, the protein engineering by *N*-glycans would give a promising opportunity for pharmacological and/or clinical applications.

4 Noninvasive Imaging of Dendrimer-Type *N*-Glycan Clusters: Remarkable In Vivo Dynamics Dependence on Non-Reducing End Structures of *N*-Glycans

Having visualized the *N*-glycoproteins in living animals, which turned out to depend significantly on the non-reducing sialic acids, our next concern was the dynamics of *N*-glycans themselves. A single molecule of *N*-glycan (monomeric glycan), however, is generally not suited as the imaging tracer because of the small size and the weak interaction with the sugar-binding proteins (~millimolar levels). Consequently, they will not accumulate in the specific organs but will be rapidly excreted from the kidney via biofiltration [13, 14]. As noted in the Introduction, most of the glycans in nature enhance their interactions by forming glycoclusters (termed cluster effects and/or multivalency effects) [89–96]. Many previous examples utilize proteins, polymers, dendrimers, and microarrays as platforms for chemically mimicking the natural glycocluster environment. For representative reviews of glycodendrimers and clusters see [97–101]. In addition, natural glycoclusters often consist of several kinds of glycans (termed glycoforms), thus highly diversifying for the *glycobioenvironment*; such heterogeneous glycoclusters could be responsible for enhancing the glycan-dependent interactions and also for regulating and/or adjusting biological responses in vivo [1]. Therefore, efficiently mimicking an *N*-glycan-involved bioenvironment, is necessary to obtain detailed information on the “in vivo dynamics” of *N*-glycans by molecular imaging [15–27].

In order to mimic the *glycobioenvironment*, we designed polylysine-based dendrimer-type glycoclusters (Fig. 3);[32] the different generations of clusters, namely, those consisting of four (4-mer 5), eight (8-mer 6), and sixteen (16-mer 7)

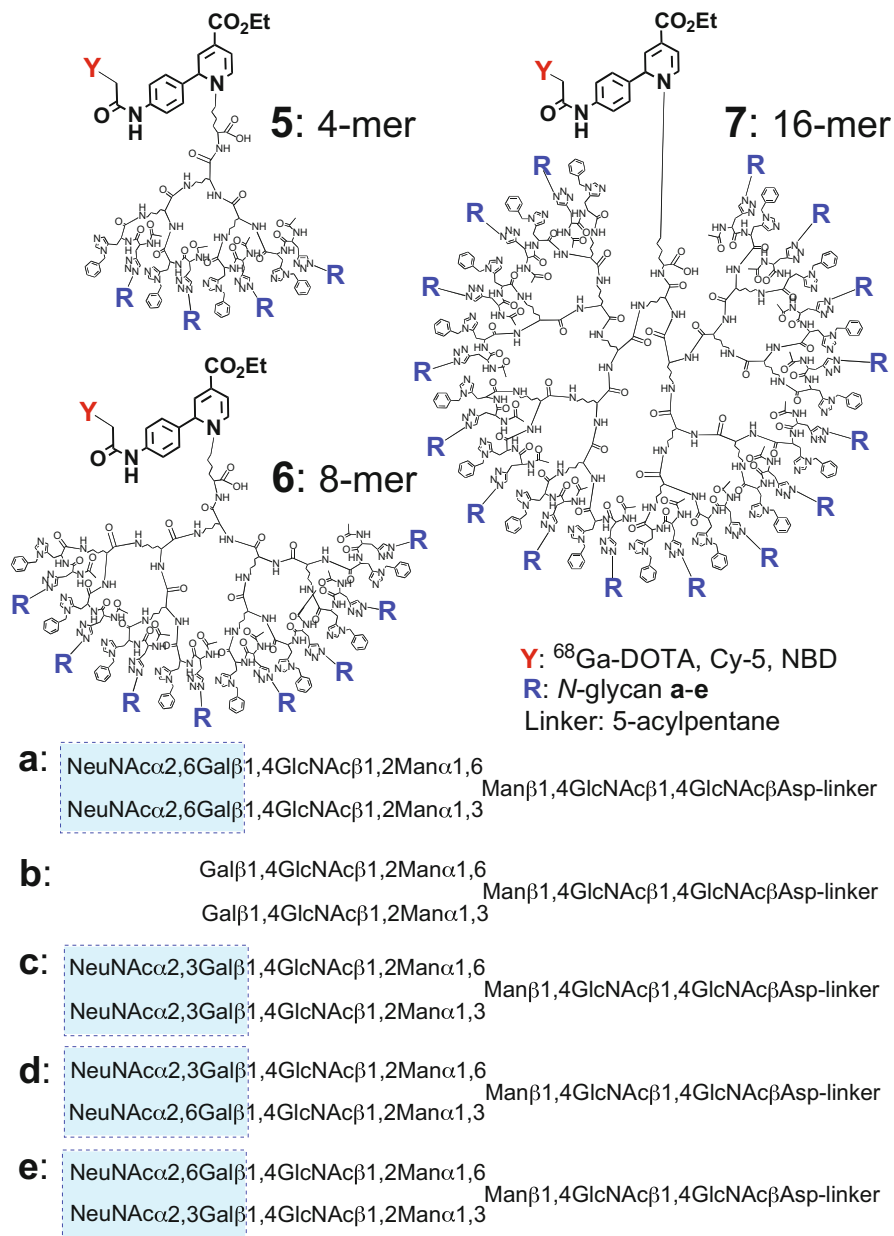
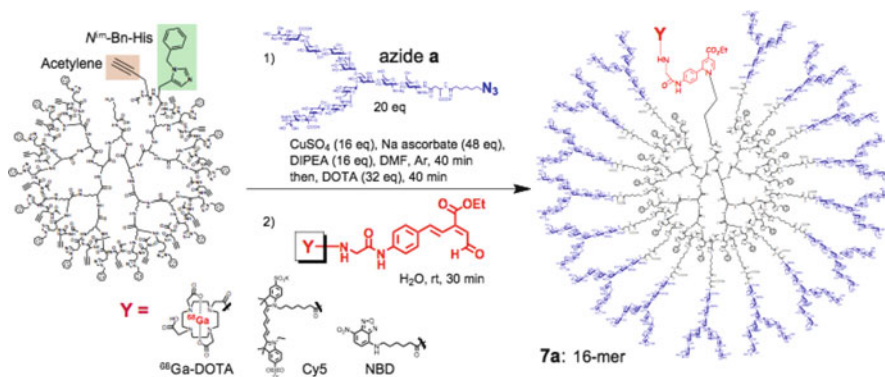


Fig. 3 Generation and structures of glycoclusters



Scheme 3 Preparation of *N*-glycan clusters through histidine-accelerated Cu(I)-mediated Huisgen 1,3-dipolar cycloaddition and labeling by 6 π -azaelectrocyclization

molecules of *N*-glycan derivatives (**a–e**) (kind gifts from Otsuka Chemical Co., Ltd., <http://tansaku.otsukac.co.jp/en/oligo04.html>). These structurally pure *N*-glycans, isolated from egg yolk according to the procedure developed by Kajihara and co-workers [102], were investigated to examine cluster (multivalency) effects [97–101] on in vivo dynamics. The clusters were designed to have a terminal lysine ϵ -amino group so that they could be efficiently labeled by fluorescent groups or ^{68}Ga -DOTA as the PET radiolabel in the presence of numerous hydroxyls through our 6 π -azaelectrocyclization protocol under mild conditions [28–30].

The polylysine-based dendrimer core with the terminal histidine and propargyl glycine residues could be synthesized by a solid-supported protocol. *N*-Glycans were subsequently introduced by “self-activating” Huisgen 1,3-dipolar cycloaddition (reaction between acetylene on the polylysine template and azide on the *N*-glycans, Scheme 3) [103]. Examples include the terminal acetylene of polylysine-based dendrimer (Scheme 3), which was smoothly reacted with azide-partner **a**, bis- $\alpha\text{Neu}(2-6)\text{Gal}$ containing complex-type *N*-glycan in the presence of equimolar amounts of copper sulfate and sodium ascorbate, and diisopropylethylamine (each relative to the *N*-glycan molecules) at room temperature for 40 min. The residual copper ions were removed by chelation with DOTA and size-partitioning centrifugal filtration using Microcon[®] (MW = 100,000, Millipore). Subsequent HPLC purification gave tetra-glycocluster **5a** (4-mer), octa-glycocluster **6a** (8-mer), and hexadeca-glycoclusters **7a–e** (16-mers) with a molecular weight of ca. 50 kDa in almost quantitative yields (Scheme 3, Fig. 3). Detection of the mother ions by MALDI-TOF analyses (as with the previous reports on mass analysis of the sialosides, the molecular ions of the sialoglycocluster **7a** were observed as more broadening peak consisting of partially desialylated derivatives [104], HPLC patterns of size-partitioning gel filtration and the integration of the representative sugars, triazole, His, and Lys signals in their ^1H NMR spectra all confirmed the desired clusters.

These clusters were subsequently labeled by DOTA or NBD- and Cy5-fluorophores through rapid 6π -azaelectrocyclization (Scheme 3). The incorporation of ^{68}Ga to DOTA-labeled glycoclusters **5a–7e** was performed by slightly modifying the procedures previously described [105] by reacting with ^{68}Ga in HEPES buffer (pH = 3.5, 95°C, 15 min) and subsequent HPLC purification.

Throughout the following imaging studies by both PET and fluorescence detection, 500 pmol of labeled glycoclusters **5a–7e** by ^{68}Ga -DOTA and Cy5 (excitation at 646 nm, emission at 663 nm) were administered from the tail vein of the BALB/c nude mice ($n = 3$) prior to whole body scanning over 4 h. We initially examined the PET of glycoclusters **5–7**, which consisted of bis-Neu α (2-6)Gal-containing glycan (**a**), asialo-glycan (**b**), and bis-Neu α (2-3)Gal-glycan (**c**). Prior to injection, the radioactivity of each ^{68}Ga -DOTA-labeled glycoclusters was adjusted to 10 MBq.

First of all, the *in vivo* dynamics between the generations of the clusters, namely those between 4-mer (**5a**), 8-mer (**6a**), and 16-mer (**7a**), differed remarkably (Fig. 4a–c). Although 4-mer **5a** and 8-mer **6a** were rapidly and almost completely cleared through the kidney (then to urinary bladder) over 1 h (Fig. 4a, b), the radioactivity derived from 16-mer **7a** was still retained in the body after 4 h (Fig. 4c), but was excreted slowly from the kidney/urinary bladder and from the gallbladder (intestinal excretion pathway); the biodistribution study of the dissected tissues after 4 h detected the ^{68}Ga -radioactivity of **7a** mostly in the liver, gallbladder, and blood, then in the order of the lung, kidney, colon, pancreas, and spleen (Fig. 5). These results clearly show the significant cluster and/or multivalency effects on the *in vivo* dynamics. Hence, the 16-mer cluster should be well suited for further biodistribution analyses by *N*-glycan imaging, but the 32-mer glycocluster should be unsuitable for *in vivo* dynamics because of the high molecular weight (ca. 100,000), and, thus, was not considered for these experiments. The differences in the clearance properties between the 4-mer, 8-mer, and 16-mer may be due to their molecular size [106, 107]. In fact, the glomerular capillary wall in the kidney is highly permeable to water, electrolytes, and substances with relatively small molecular weights, while larger molecules such as myoglobin are filtered to a lesser degree [106, 107]. A molecule size of 3 nm in diameter is known to clear smoothly through the kidney. Clusters **5** (MW = 10,000) and **7** (MW = 50,000) in the unfolded conformation are estimated to be roughly 3 and 7 nm in size, respectively. Therefore, the smaller sizes of the 4-mer and the 8-mer can easily clear through bio-filtration in the kidney.

Alternatively, the degree of negative charge (the number of sialic acid residues) and/or that of the hydrophilicity of the clusters, i.e., hydrophilic *N*-glycans vs the hydrophobic poly-lysine backbone occupying the surface on the clusters may affect rapid clearance through the kidney. For example, the 4-mer and 8-mer, which have more hydrophobic natures, may be trapped by the scavenger receptors in the serum, similar to the degradation process for the “misfolded” hydrophobic glycoproteins inside the endoplasmic reticulum (ER) (quality control of glycoproteins) [108].

We subsequently examined the *in vivo* dynamics and the biodistribution of asialo-glycan **7b** and bis-Neu α (2-3)Gal glycan **7c** using the 16-mer poly-lysine template (Fig. 3). Unlike bis-Neu α (2-6)Gal-case **7a** (Fig. 4c), asialo-glycan cluster

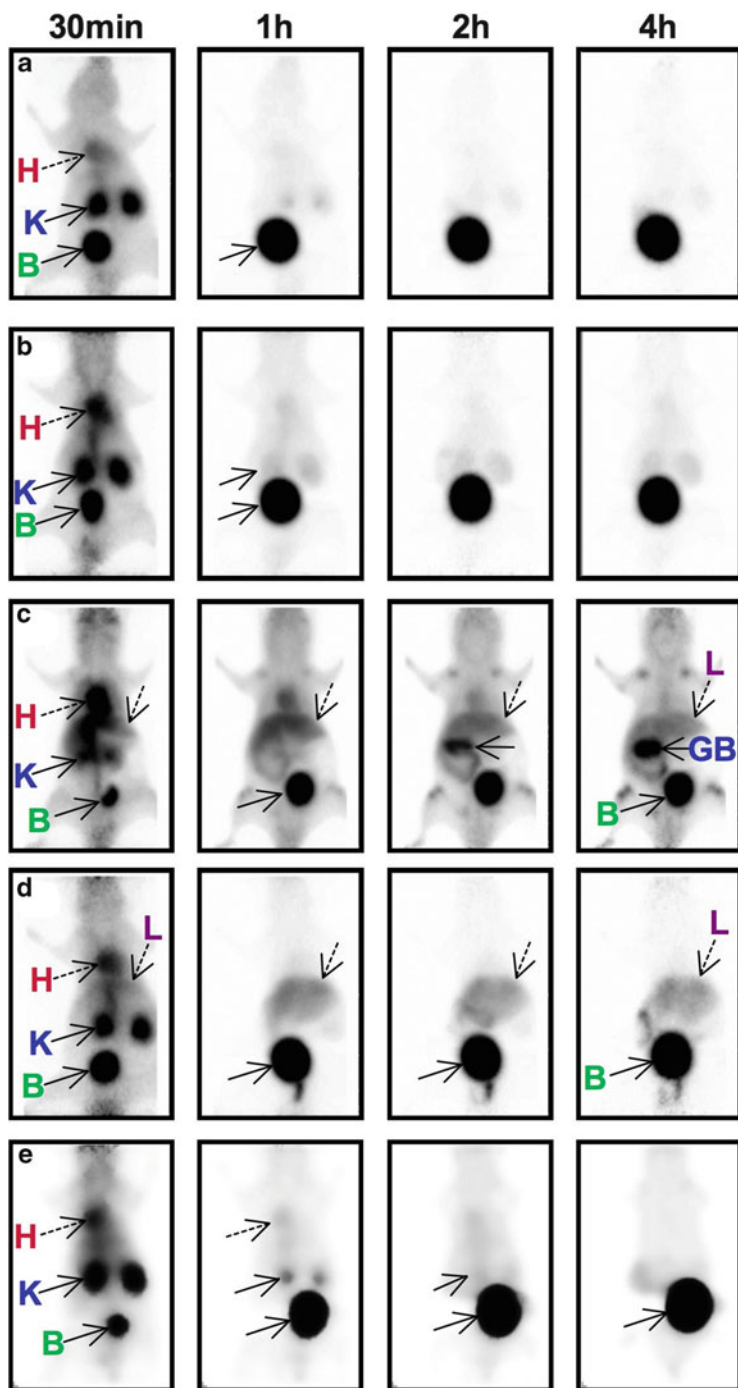


Fig. 4 Dynamic PET imaging of glycoclusters **5a**, **6a**, and **7a-c** in normal BALB/c nude mice. ^{68}Ga -DOTA-labeled glycoclusters (10 MBq) were administrated from the tail vein of the mice

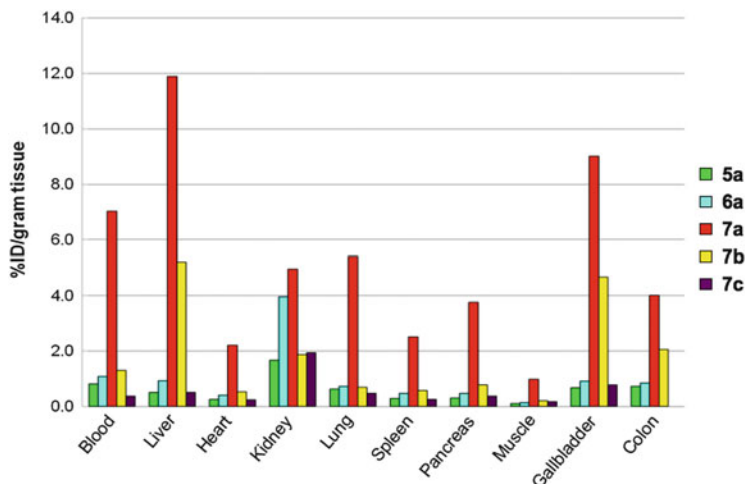


Fig. 5 Standardized uptake value (SUV) of ^{68}Ga -DOTA-labeled glycoclusters **5a**, **6a**, and **7a–c** in normal BALB/c nude mice. SUV was normalized by radioactivity in the tissue region, injected dose, and weight of the subjects. These values represent the means and S.Ds

7b rapidly cleared through the kidney to the bladder (Fig. 4d), although some accumulation was observed in the liver because the asialoglycoprotein receptors are highly expressed in this organ [2]. Dissection experiments after 4 h also found the strongest accumulation in the liver followed by the gallbladder and slight accumulation in the colon and kidney (Fig. 5). The results are consistent with the PET analyses of glycoproteins, orosomucoid, and its asialoorosomucoid as described in the previous section [8, 9, 28–30], where the asialo-congener is more rapidly excreted than orosomucoid through the kidney and the liver/gallbladder pathways (leading to intestinal excretion). However, the α -linking to the 3-OH of galactose in glycocluster **7c**, which also contains sialic acid, was readily excreted through the kidney/urinary bladder as shown in Fig. 4e. These PET results on the 16-mer glycoclusters **7a–c** suggest that the specific sialoside linkage to galactose, i.e., Neu α (2-6)Gal linkage, in *N*-glycan structures plays an important role in the circulatory residence of *N*-glycans, which in turn results in uptake of **7a** in the liver (Figs. 4c and 5). In addition, this specific sialoside linkage markedly differentiates the excretion mechanism from those of the asialo- and Neu α (2-3)Gal cases, which proceed via bio-filtration pathway through the kidney.



Fig. 4 (continued) ($n = 3$, 500 pmol, 100 μL /mouse) and the whole body was scanned by a small animal PET scanner, microPET Focus 220 (Siemens Medical Solutions Inc., Knoxville, TN, USA), over 0–4 h after injection. *H* heart, *K* kidney, *L* liver, *B* urinary bladder, *GB* gallbladder. (a) glycocluster **5a**; (b) glycocluster **6a**; (c) glycocluster **7a**; (d) glycocluster **7b**; (e) glycocluster **7c**

The prolonged “in vivo lifetime” for the sialic acid-containing glycoclusters agrees well with the well-known hypotheses of the clearance of the asialoglycoproteins through the asialoglycoprotein receptors [2, 109]. However, the notable difference in serum stability produced by the sialoside bond linkages to the galactose, i.e., the $\alpha(2-6)$ - vs $\alpha(2-3)$ -linkages, is an intriguing observation. Very interestingly, Unverzagt and co-workers have reported the opposite serum stability of neoglycoproteins in mice, i.e., $\alpha(2-3)$ -sialylated derivatives show longer half-lives than the $\alpha(2-6)$ -sialylated congeners, reported when glycans were attached to albumin [110]. Our dynamic PET images of the *N*-glycoclusters therefore suggest a new receptor-mediated excretion mechanism for Neu $\alpha(2-3)$ Gal-containing glycans. Namely, Neu $\alpha(2-3)$ Gal-cluster **7c**, which usually cannot be found in serum, is probably recognized as an invader and smoothly excreted by the vascular endothelial cells, erythrocytes, leucocytes, and via the phagocytosis by a macrophage; the smaller sized degradation products may be filtered in the kidney. Alternatively, the “excretion-escaping” mechanism, involving stimulating the immunosuppressive signals through the ITIM (immunoreceptor tyrosine-based inhibitory motif) molecules via Siglec families [82–85], may account for the higher stability of Neu $\alpha(2-6)$ Gal-glycan (see below). It is reported that the Neu $\alpha(2-6)$ Gal-containing BSA reduces but does not prevent binding to the asialoglycoprotein receptor, while the Neu $\alpha(2-3)$ Gal congener abolishes the binding [111]. Therefore, the prolonged half-life coupled to uptake by the asialoglycoprotein receptor accounts for the high accumulation of **7a** in the liver (Fig. 4c). Biantennary Neu $\alpha(2-6)$ Gal-chains especially reduce binding in comparison with tri- and tetra-antennary glycans [112]. Nevertheless, the combination of high valency and long circulatory half-life (reduced clearance) likely leads to the uptake of **7a** by hepatocytes via the Gal/GalNAc lectin receptor. Note that the slow clearance of bis-Neu $\alpha(2-6)$ Gal cluster **7a** through the gallbladder may be caused by the “polar transport mechanism” [113, 114], which “tags” the fucose to specific *N*-glycans in the liver. Elucidation of a detailed mechanism will be the subject of future investigations.

In light of the strikingly different in vivo dynamics and excretion pathways between the *N*-glycan clusters **7a–c**, which are sensitively regulated by the truncated structures at the non-reducing end, namely, in the presence or the absence of sialic acid and their glycosyl bond linkages to the galactose, a biodistribution study of the glycoclusters **7a–e** was performed in further detail using fluorescence imaging (Fig. 6, upper images (a–e)) [23]. Clusters **7d** and **7e**, which contain mixed Neu $\alpha(2-6)$ Gal and Neu $\alpha(2-3)$ Gal moieties, were particularly interesting because the circulatory stability was notably enhanced by Neu $\alpha(2-6)$ Gal moiety. Although the fluorescence-based in vivo dynamics and biodistribution of clusters **7a–e** were almost identical to those detected by PET, fluorescence detection more clearly visualized the time-course accumulations in specific organs. Thus, as visualized by PET (Fig. 4), glycoclusters **7a**, **7d**, and **7e**, which contain at least one Neu $\alpha(2-6)$ Gal non-reducing end motif, showed higher stabilities in vivo than asialo and Neu $\alpha(2-3)$ Gal congeners **7b** and **7c**. The Cy5-fluorescence derived from clusters **7a**, **7d**, and **7e** was eventually accumulated mostly in the liver, presumably because of the interaction with asialoglycoprotein receptor [2, 109], but was also

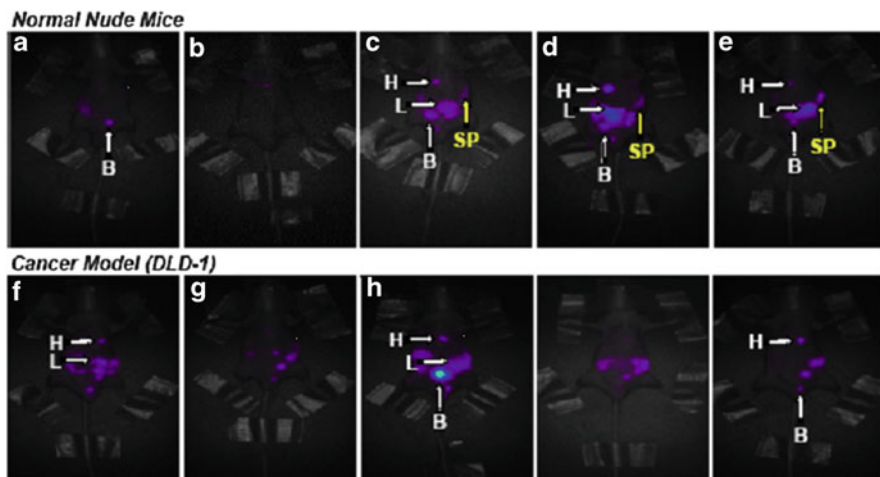


Fig. 6 (a–e) Dynamic fluorescence imaging of glycoclusters **6a–e** in BALB/c nude mice: (a) glycocluster **6b**; (b) glycocluster **6c**; (c) glycocluster **6a**; (d) glycocluster **6d**; (e) glycocluster **6e**. (f–j) Images in DLD-1 implanted nude mice (*left femoral region*) as a cancer model: (f) glycocluster **6b**; (g) glycocluster **6c**; (h) glycocluster **6a**; (i) glycocluster **6d**; (j) glycocluster **6e**. Cy5-labeled glycoclusters **6a–e** were administered from the tail vein of the mice ($n = 3$, 500 pmol, 100 μ L/mouse) and whole body scans were performed from the front side by eXplore Optix, GE Healthcare, Bioscience (excitation at 646 nm, emission 663 nm), 4 h after injection. Data was normalized. *H* heart, *L* liver, *B* urinary bladder, *SP* spleen

observed in the spleen after 4 h (Fig. 6c–e). Out of the three glycoclusters, glycocluster **7e**, which contains both “Neu α (2-6)Gal” and “Neu α (2-3)Gal” non-reducing end structures (from 6- and 3-hydroxyls of branching mannose, respectively), showed the highest fluorescent intensity in the spleen (Fig. 6e).

As noted previously, the spleen is part of the immune system, where the antigen-specific antibodies are produced by the interaction between T-cells and mature B-cells. Because Siglec 2 (CD22) [82–88], which is a Neu α (2-6)Gal-specific lectin, is overexpressed in mature B-cells and is responsible for immune-negative regulations, we examined the possibility of an interaction with our clusters **7a–e**. A20 B-cell line (expressed murine CD22), and Daudi B-cell line (expressed human CD22) [115–117] only showed a weak interaction with these clusters; cluster **7d** exhibited the weakest interaction. Therefore, Neu α (2-6)Gal-containing clusters **7a**, **7d**, and **7e** are most likely captured by the RES, which consists of phagocytic cells such as monocytes and macrophages which traffic to the spleen and liver as the reticular connective tissues. Yet during the capture by the RES (reticuloendothelial system), stimulation by “immune suppressive signal” through the Neu α (2-6)Gal/Siglec interactions on the phagocytic cells [82–85] cannot be ruled out to explain the serum stability of these clusters. Overall, these *in vivo* images clearly visualize the importance of “at least” one Neu α (2-6)Gal moiety in the circulatory residence of the *N*-glycans and the precise regulation of the biodistribution by a combination of the α (2-6)- and α (2-3)-sialoside linkages.

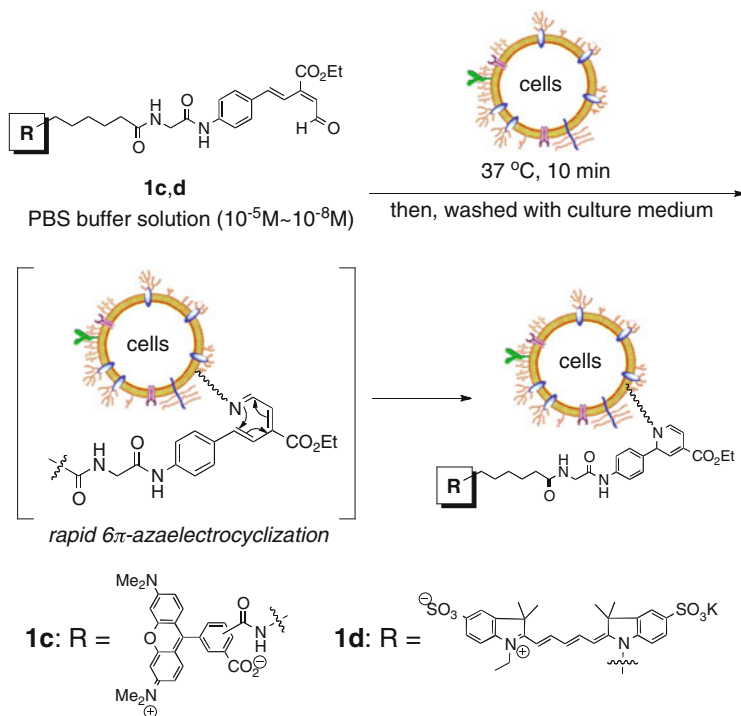
Additionally, an examination of glycoclusters **3a–3e** in a cancer model was performed (Fig. 6, lower images (f–j)). Although these clusters did not target the tumor tissue (DLD-1 implanted to the left femoral regions), markedly different *in vivo* dynamics were observed from those in normal mice. For example, excretion rates of glycoclusters **7d** and **7e** were accelerated in the cancer mice and accumulation was not detected in the spleen (Fig. 6i, j), whereas excretion rate of **7b**, which lacks sialic acid, was considerably suppressed in the cancer mice (Fig. 6f). Although the phenomena could not be explained by the currently available data, these differences make *N*-glycan clusters applicable to a new class of diagnostic tracers.

Thus, we have developed new dendrimer-type *N*-glycan clusters. Their structure-specific “*in vivo* dynamics” and “biodistribution” were clearly visualized by PET and fluorescence imaging. This study demonstrated for the first time the marked difference in the *in vivo* dynamics and biodistributions between $\alpha(2-6)$ or $\alpha(2-3)$ sialosides, although multivalent effects were shown to allow high selectivity and affinity in ligand–protein interactions [89–101]. Totally different dynamics of the *N*-glycans between the normal and the tumor models could also be discovered by the present investigation. These results indicated the importance and validity of the *in vivo* molecular imaging in living animals. Because various *N*-glycan clusters can easily be prepared via histidine-accelerated “self-activating” Huisgen 1,3-dipolar cycloaddition and efficiently labeled by imaging agents through the 6- π -electrocyclization protocol, the present method has potential in developing glycocluster-based imaging tracers.

5 Noninvasive Fluorescence Imaging of *N*-Glycan-Engineered Living Cells: Effects of *N*-Glycans on Lymphocytes for Tumor Targeting

5.1 Fluorescence Labeling of Living Cell Surfaces by Azaelectrocyclization

Our azaelectrocyclization chemistry developed in Sect. 2 can also covalently and selectively label most of the accessible amino groups on cell surface constituents via a *one-step* procedure under mild and the physiological conditions, i.e., lysines of membrane proteins and/or ethanol amine derivatives, without interfering with their native functions [31]. Although the recently developed protocols, including the Bertozzi’s Staudinger ligation [51–60] and strain-accelerated Huisgen 1,3-cycloaddition reaction [25, 61–71], are noteworthy by their bioorthogonality and now widely applied to the chemical biology research, they require a “long-time” double procedure consisting of metabolic cell modification and subsequent chemical reaction [63, 66–68]. Our method, on the other hands, is more straightforward, easy-handling, and therefore widely applicable, i.e., by directly treating a whole



Scheme 4 Fluorescence labeling of the living cells

cell by unsaturated aldehydes [30, 31]. For instances, the C6 glioma cells [118], which were either coated on a dish or used as a suspension, were treated with a PBS buffer solution of TAMRA-probe **1c** (1.0×10^{-5} M) at 37 °C for 10 min (Scheme 4). After the cells were washed with a culture medium a few times, they were fixed with formaldehyde. The fluorescence labeled cells were further treated with DAPI as a blue dye to visualize the nuclei, and analyzed by confocal laser scanning microscopy (Fig. 7).

Red fluorescence derived from the TAMRA dye (excitation at 525 nm, emission at 555 nm) was observed exclusively on the surface of the glioma cells treated with **1c** (Fig. 7a), whereas a weaker fluorescence was distributed over the entire cell when the cells were treated with the succinimidyl ester of TAMRA (NHS-TAMRA) as control, i.e., not localized on the cell membrane (Fig. 7b). These results clearly show that TAMRA-probe **1c** rapidly reacts with the amino groups of the lysines and/or other amino-containing cell membrane constituents, such as phosphatidyl ethanolamines, to anchor the fluorescence label. On the other hand, the succinimidyl ester reagent was internalized into the cell before reacting with the amino groups on the cell membrane because of the much lower reactivity of this reagent than **1c** under such mild labeling conditions. Selective labeling on the cell membrane was also supported by co-staining with the cell-surface marker

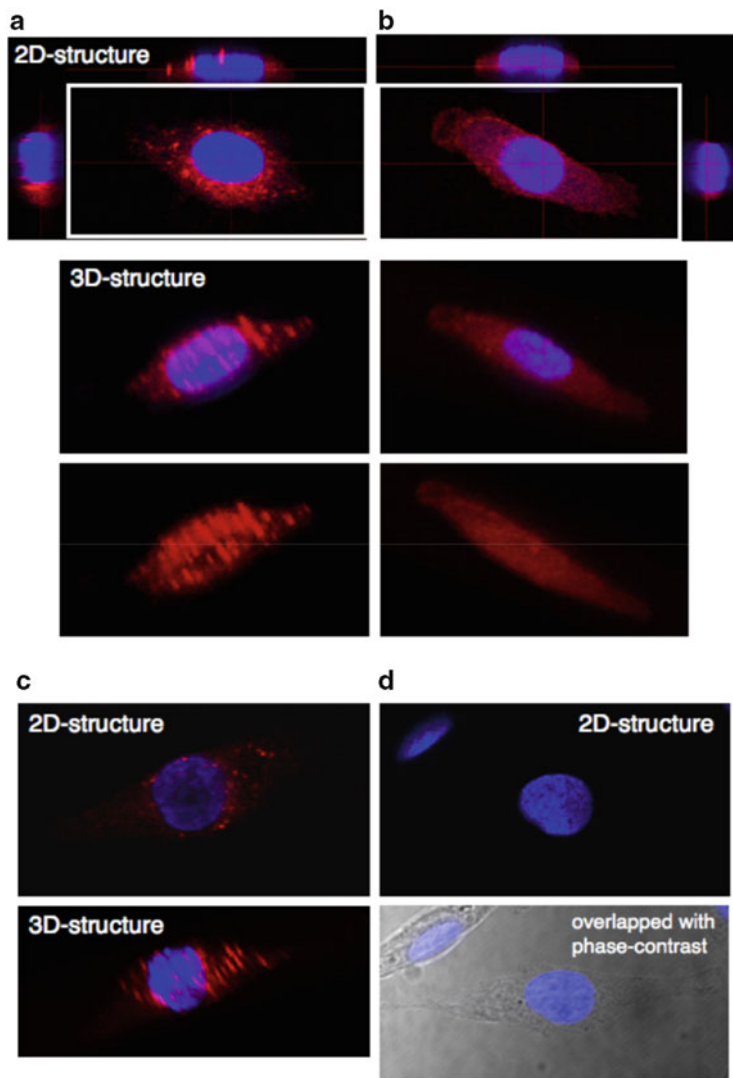


Fig. 7 Confocal microscopy of the TAMRA-labeled C6 glioma cells (excitation at 525 nm). As a *blue dye*, DAPI is introduced after fixing the TAMRA-labeled cell by paraformaldehyde treatments. 2D- and 3D-pictures of a labeled cell by (a) unsaturated aldehyde probe **1c** and (b) TAMRA-succinimidyl ester at 1×10^{-5} M, 37°C for 10 min. Labeling performed at 1×10^{-8} M, 37°C for 10 min by (c) **1c** and (d) TAMRA-succinimidyl ester

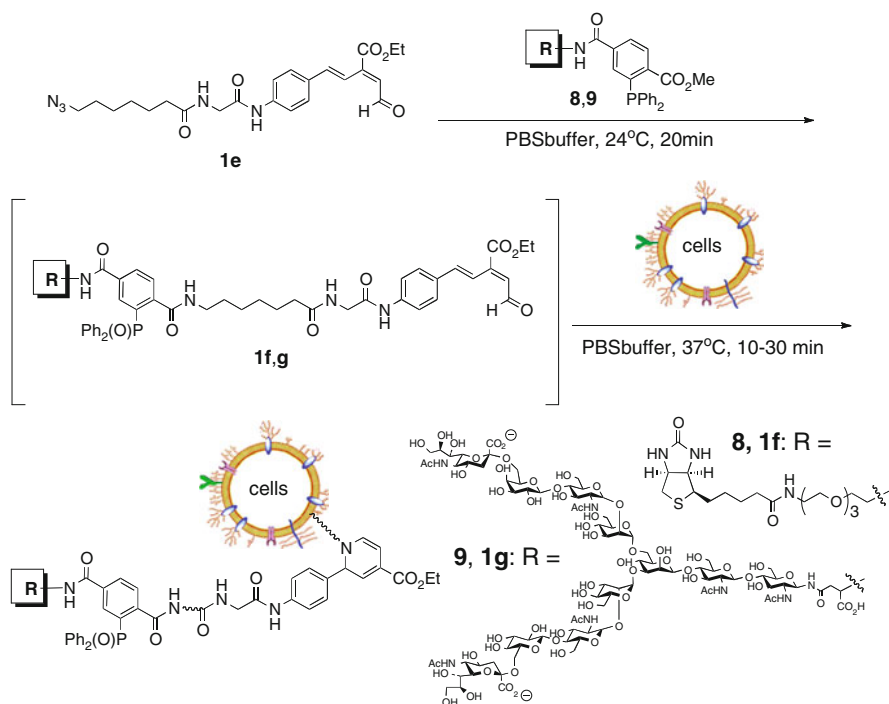
[31] and treatment with Triton-X100, a detergent which damages the lipid bilayer of the cells [119].

It should be noted that the fluorescence labeling by **1c** was effective even at a 10^{-8} M concentration of **1c**, while no fluorescence was detected when treated by

TAMRA-NHS ester at this highly diluted concentration (Fig. 7c, d at 37°C for 10 min). To date, this is the most diluted concentration to be realized by the chemistry-based bioconjugation through covalent bond formation. Interestingly, confocal microscopy in Fig. 7c detected specific localization of the TAMRA-fluorescence on the cell membrane, such that the fluorescence surrounded the nucleus and then radiated outward. Based on previous results on the reactivity of the unsaturated aldehyde probes [28–30, 75, 77], azaelectrocyclization might proceed selectively at the protein assembly (high concentration of Lys) and/or the other ethanolamine functions on the cell surface, such as raft or microdomains, under such a diluted concentration of **1c**. Because of such mild labeling conditions performed with **1c**, 95% of the glioma cells were alive according to trypan blue tests; the labeled glioma cells therefore divided over 2–44 h without incident [31]. Thus, the excellent reactivity of unsaturated aldehyde **1c** with the amino functions enables simple *one-step* chemical labeling of living cells under extremely mild conditions.

5.2 Chemical Engineering of Cell Surfaces by *N*-Glycans and Other Functional Molecules

Complex-type *N*-glycans and other functional molecules, such as biotin, could also be efficiently introduced on cell surfaces using current azaelectrocyclization chemistry (Scheme 5); [33] thus, azide-containing aldehyde **1e**, which is readily prepared from the corresponding aminoalcohol, was initially treated with 2-methoxycarbonylphenyldiphenylphosphines **8** and **9** [59, 102] in a PBS buffer at 24°C. Intermediary products **1f** and **1g** were subsequently reacted with the C6 glioma cell *in a one-pot procedure* at 37°C for 10–30 min (1.0×10^{-5} M for **1f** and **1g**). The modified cell surfaces were then evaluated by treatments with TRITC (rhodamine)-labeled avidin and TRITC-labeled SNA (Sambucus nigra, Elderberry, Neu(α 2-6)Gal-specific lectin) [120] (Fig. 8). It will be noted that, for the C6 cells engineered by *N*-glycan probe **1g**, TRITC-fluorescence was localized in specific regions of the nucleus outward on the cell surfaces, i.e., artificial *N*-glycan might be introduced in the same regions where native glycans are expressed. Thus, the possible modification site by azaelectrocyclization chemistry was again suggested to be the rafts or microdomains, the proteins and/or other amino groups-rich domains. We would like to emphasize that the opposite reaction order, i.e., first azaelectrocyclization with the cells and then Staudinger ligation and/or the other conjugation methods, such as the strain-accelerated Huisgen 1,3-cycloaddition reaction [25, 61–71], might provide a choice of method, and, therefore, the rapid azaelectrocyclization chemistry can significantly expand the labeling and/or modification of living cell surfaces.



Scheme 5 Cell surface engineering by combined Staudinger/electrocyclization

5.3 Noninvasive Fluorescence Imaging of Lymphocytes Trafficking and Effects of Cell-Surface Engineering by N-Glycan for Tumor Targeting

The whole cell-based labeling and *N*-glycan engineering protocols were then applied to the *in vivo* fluorescence imaging of the lymphocytes (Fig. 9). The lymphocytes (1 mL, 1.0×10^5 cells) directly extracted from the abdominal cavity of wild-type mice were fluorescently labeled by reacting with Cy5-fluorescence probe **1d** [31] (Scheme 4, excitation at 646 nm, emission at 663 nm) in a 1.0×10^{-5} M PBS buffer solution at 37°C for 10 min. Although the labeling was also possible at 10^{-8} M, the decay of the fluorescence, caused by cell division, decomposition, and/or metabolization, during the *in vivo* imaging over 1 week [31] caused a problem in fluorescence detection in sensitivity (see Fig. 9). The Cy5-labeled lymphocytes (100 μ L/mouse, 10^4 cells) were then administered from the tail vein of the nude mouse and the dynamic fluorescence images were recorded over a week by using eXplore Optix, GE Healthcare, Bioscience (Fig. 9) [23]. The lymphocytes gradually accumulated over 6 h in the immunology-related organs, namely spleen and intestinal lymph nodes, from which the fluorescence intensity in the spleen disappeared (Fig. 9a). These are the marked results which the trafficking

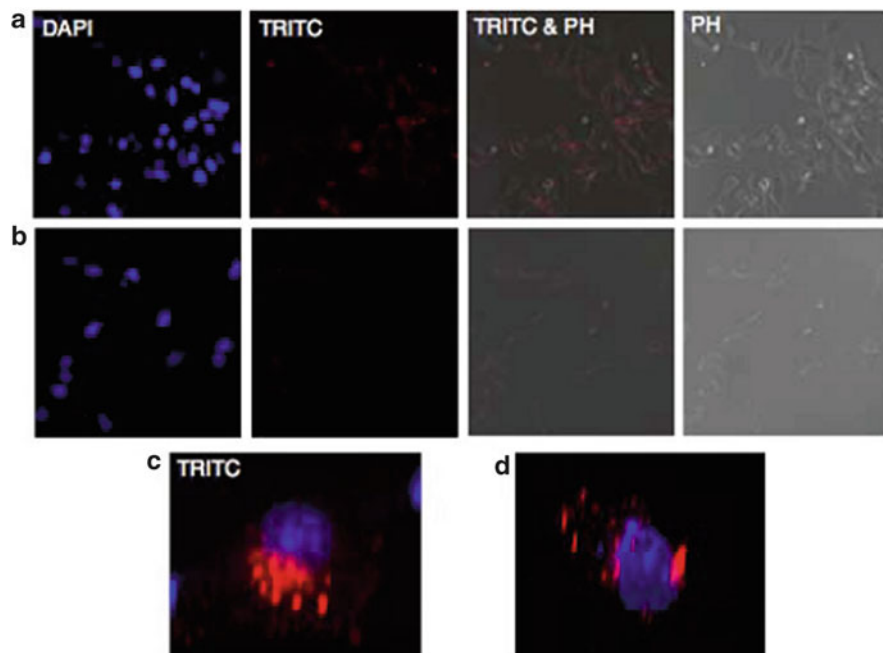


Fig. 8 Confocal microscopy of biotin- and *N*-glycan-modified C6 glioma cells (DAPI-, TRITC, and PH-detections). (a) **1f**-Modified cell on TRITC-avidin treatments. (b) Control cell treated by biotin-containing phosphine **8** on TRITC-avidin treatments. (c) **1g**-Modified cell on SNA lectin (α Neu(2-6)Gal-specific) treatments. (d) Intact cell on SNA lectin treatments; *PH* Phase-contrast

of the cells, especially into the spleen, clearly visualized with exceptionally high imaging contrasts compared with those reported under the antibody-based labeling protocols [8, 9, 121, 122]. Thus, the observation highlights the “direct” electrocyclization-based labeling to a whole cell-based *in vivo* imaging. For leading examples of MRI using whole cells, see [123–126]. The fluorescence-labeled lymphocytes were then injected to the mouse of the cancer model, where the DLD-1 human colon carcinoma was implanted to the dorsal division of the mouse (Fig. 9b). The trafficking properties were similar to those observed in the normal nude mouse, although the accumulation to the spleen was slightly slower over 6 h. No fluorescence was detected in the tumor tissues in a week. On the other hands, we made a remarkable observation when the lymphocytes were simultaneously labeled (by **1d**) and engineered by the *N*-glycan (by **1g**) under the conditions established in Schemes 4 and 5; namely, the fluorescence intensity of the engineered cells still remained strong in the spleen after 48 h (Fig. 9c), while they gradually began to accumulate into DLD-1 (TM). The fluorescence intensity in the tumor regions became even stronger over a week.

To explain the imaging results in Fig. 9, further investigations are needed to examine the mechanisms at the molecular level. Nevertheless, since neither the native lymphocytes nor the complex-type *N*-glycan clusters (see Sect. 4) [32] traffic

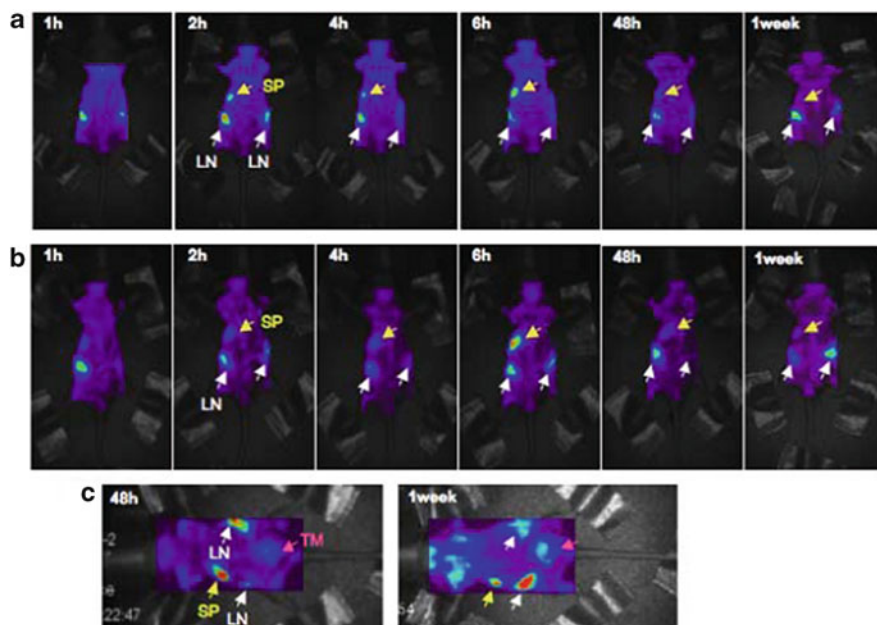


Fig. 9 Fluorescence imaging of lymphocytes in mice. Labeled and/or engineered cells were administrated intravenously ($n = 3$, $100 \mu\text{L}/\text{mouse}$, 10^4 cells) and whole body was scanned from the back side by eXplore Optix, GE Healthcare, Bioscience (excitation at 646 nm, emission 663 nm), 1 h, 2 h, 4 h, 6 h, 48 h, and 1 week after injection. Data were normalized. *SP* spleen, *LN* lymph node of epidermal intestinal tract; *TM* DLD-1 human colon carcinoma. (a) Cy5-labeled cells to the nude mice. (b) Cy5-labeled cells to the DLD-1 implanted nude mice at dorsal division. (c) Both Cy5-labeled and *N*-glycan-engineered lymphocytes into the tumor model

to the cancer regions on their own, the engineering of the lymphocyte surfaces is key to tumor targeting. The low efficiency of lymphocyte-based cancer immunotherapy, such as using NK, LAK, CTL, and TIL as the immune effector cells, is well documented, and it has been widely applied to circumvent the problems of (1) co-injection with cytokines, co-stimulatory molecules, or cancer antigens, and more recently, (2) engineering of the lymphocytes by the chimeric antigen receptors [127]. The reason for *N*-glycan-enriched lymphocyte trafficking to tumor is unclear. There are two possible explanations: (1) the *N*-glycan-enriched lymphocytes might interact with the tumor via additional interaction with lectins on cancer cells and (2) the interaction of Siglec and internal sialoglycans on the lymphocytes, which stimulate the immunosuppressive signals through the ITIM molecules [82–85], could be interrupted by the *N*-glycan **1g** externally introduced by chemical engineering. This might eventually activate the lymphocytes to target the tumor.

These preliminary results show the potential of the current engineering method to tune the functions of the living cells very rapidly and conveniently, may we say, *in tailor-made fashion*, in adjusting to dynamic responses in vivo. This is just as the natural glycans are diversifying their *glycobioenvironment* through the dynamic

process, by creating the heterogenic clusters on the cell surfaces. Based on this chemistry, it may also be possible to label living cells with metal-incorporated DOTA, which would greatly expand the method to targeting, imaging, or clinical applications, and therefore highlight the protocol.

6 Conclusion

Thanks to azaelectrocyclization chemistry, we can now visualize the *in vivo* dynamics of any glycans on proteins, dendrons, as well as on living cells by PET and/or noninvasive fluorescent imaging. We have so far imaged the remarkable effects of sialic acids at the non-reducing end of *N*-glycans on the *in vivo* dynamics of the glycoconjugates. These imaging results give the quite promising possibility for the future of developing glycan-based imaging tracers, which selectively target specific tumors and/or inflammation. As discussed in this review, natural glycoclusters often consist of several kinds of glycans (termed glycoform), thus highly diversifying the *glycobioenvironment*. We therefore need to synthesize and mimic such heterogeneous glycoclusters for efficient glycan-based imaging in the future. Wong, Wu, and co-workers have recently prepared heterogeneous clusters consisting of two different glycans on the microarray and also on the dendrimer template, and have experimentally for the first time exemplified enhanced and/or decreased reactivity towards antibodies, which are sensitively affected by the neighboring glycans in the heterogeneous structures [128]. By applying their glycocluster preparation on the microarray and/or on dendrimeric templates, the various heterogeneous glycocluster environments will be constructed in a combinatorial fashion, leading to a promising strategy in developing not only glycan-based diagnostic tracers, but also novel glycan-based vaccine adjuvants or vaccine-neutralizing pharmaceuticals. However, the previously reported heterogeneous glycocluster synthesis, including Wong and Wu, can approximately estimate the “average ratio” between a few different glycans on the array and/or within the dendron structures, but cannot precisely define the spatial relationship between the neighboring glycans in the clusterized structures. To evaluate in detail the *heterogeneity* in natural glycoclusters and then mimic them to develop useful glycan-based imaging tracers, we glycochemists should consider how structurally defined *homogeneous* heterogeneous glycoclusters could be prepared efficiently and rapidly, i.e., by a combinatorial approach. In other words, these issues depend on how many organic reactions, which are truly *orthogonal* to the functional groups, can be developed in the future.

Acknowledgments The authors thank Professor James C. Paulson and Dr. Weihsu C. Chen, The Scripps Research Institute, for cell-based assays for the glycoclusters. The authors also thank Hiroaki Asai, Dr. Michio Sasaoka, Kazuhiro Fukae, and Azusa Hashimoto, Otsuka Chemical Co., Ltd. for supplying *N*-glycans. This work was supported in part by Grants-in-Aid for Scientific Research No. 19681024 and 19651095 from the Japan Society for the Promotion of Science,

Collaborative Development of Innovative Seeds from Japan Science and Technology Agency (JST), New Energy and Industrial Technology Development Organization (NEDO, project ID: 07A01014a), Research Grants from Yamada Science Foundation, and Molecular Imaging Research Program, Grants-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology (MEXT) of Japan.

Biographical Sketches



Koichi Fukase received his Ph.D. in 1987 from the Department of Chemistry, Graduate School of Science, Osaka University under the direction of Prof. Tetsuo Shiba, and continued his research as a postdoctoral fellow in the same group. Then he became a Research Associate (1988–1996) at the Department of Chemistry, Faculty of Science, Osaka University (Prof. Shoichi Kusumoto). From 1994 to 1995, he joined the group of Prof. W. Clark Still at Columbia University as a postdoctoral fellow. He was promoted to Assistant Professor (1996), Associate Professor (1998) (Prof. Shoichi Kusumoto), and became a full Professor in

2004. Research projects within his group focus on (1) chemistry in innate immunity, (2) efficient and selective methods for glycosylation and oligosaccharide synthesis, (3) new labeling methods for PET imaging, and (4) combinatorial chemistry, solid-phase synthesis, and high-throughput synthesis using affinity separation.



Katsunori Tanaka received his BS (1996) and Ph.D. (2002) from Kwansei Gakuin University in Japan under the direction of Professor Shigeo Katsumura. After a post-doc with Professors Koji Nakanishi and Nina Berova at Columbia University, New York (2002–2005), he joined Professor Koichi Fukase's group at Osaka University as an Assistant Professor (2005–present). His research interests include exploring new methods for total synthesis, configurational analysis, biological evaluation, molecular imaging, and molecular recognition of natural products.

References

1. Kamerling JP, Boons G-J, Lee YC, Suzuki A, Taniguchi N, Voragen AGJ (eds) (2007) Analysis of glycans, polysaccharide functional properties and biochemistry of glycoconjugate glycans, carbohydrate-mediated interactions: comprehensive glycoscience, from chemistry to systems biology, vol II & III. Elsevier, UK
2. Morell AG, Irvine RA, Sternlieb I, Scheinberg IH, Ashwell G (1968) *J Biol Chem* 243:155–159
3. Elliott S, Lorenzini T, Asher S, Aoki K, Brankow D, Buck L, Busse L, Chang D, Fuller J, Grant J, Hernday N, Hokum M, Hu S, Knudten A, Levin N, Komorowski R, Martin F, Navarro R, Osslund T, Rogers G, Rogers N, Trail G, Egrie J (2003) *Nat Biotechnol* 21:414–421
4. Sato M, Furuike T, Sadamoto R, Fujitani N, Nakahara T, Niikura K, Monde K, Kondo H, Nishimura SI (2004) *J Am Chem Soc* 126:14013–14022
5. Kaneko Y, Nimmerjahn F, Ravetch JV (2006) *Science* 313:670–673
6. Suzuki M, Noyori R, Langstom B, Watanabe Y (2000) *Bull Chem Soc Jpn* 73:1053–1070
7. Hosoya T, Sumi K, Doi H, Wakao M, Suzuki M (2006) *Org Biomol Chem* 4:410–415
8. Tanaka K, Fukase K (2008) *Org Biomol Chem* 6:815–828
9. Tanaka K, Fukase K (2008) *Mini-Rev Org Chem* 5:153–162
10. Ido T, Wan CN, Cassela V, Fowler JS, Wolf AP (1978) *J Label Compds Radiopharm* 14:175–183
11. Gambhir SS (2002) *Nat Rev Cancer* 2:683–693
12. Tanaka K, Fujii Y, Tokimoto H, Mori Y, Tanaka S, Bao G-m, Siwu ERO, Nakayabu A, Fukase K (2009) *Chem Asian J* 4:574–580
13. Vyas SP, Singh A, Sihorkar V (2001) *Crit Rev Ther Drug Carrier Syst* 18:1–76
14. Willis M, Forssen E (1998) *Adv Drug Deliv Rev* 29:249–271
15. Andre S, Unverzagt C, Kojima S, Dong X, Fink C, Kayser K, Gabius H-J (1997) *Bioconjug Chem* 8:845–855
16. Andre S, Unverzagt C, Kojima S, Frank M, Seifert J, Fink C, Kayser K, von der Lieth C-W, Gabius H-J (2004) *Eur J Biochem* 271:118–134
17. Andre S, Kojima S, Prahl I, Lensch M, Unverzagt C, Gabius H-J (2005) *FEBS J* 272:1986–1998
18. Andre S, Kojima S, Gundel G, Russwurm R, Schrott X, Unverzagt C, Gabius H-J (2006) *Biochim Biophys Acta* 1760:768–782
19. Andre S, Kozar T, Schubert R, Unverzagt C, Kojima S, Gabius H-J (2007) *Biochemistry* 46:6984–6995
20. Andre S, Kozar T, Kojima S, Unverzagt C, Gabius H-J (2009) *Biol Chem* 390:557–565
21. Pimm MV, Perkins AC, Strohal J, Ulbrich K, Duncan R (1996) *J Drug Target* 3:385–390
22. Prata MIM, Santos AC, Torres S, Andre JP, Martins JA, Neves M, Garcia-Martin ML, Rodrigues TB, Lopez-Larrubia P, Cerdan S, Geraldes CFGC (2006) *Contrast Media Mol Imaging* 1:246–258
23. Hirai M, Minematsu H, Kondo N, Oie K, Igarashi K, Yamazaki N (2007) *Biochem Biophys Res Commun* 353:553–558
24. Chen WC, Completo GC, Paulson JC (2008) *Glycobiology* 18:963–963
25. Laughlin ST, Baskin JM, Amacher SL, Bertozzi CR (2008) *Science* 320:664–667
26. Serres S, Anthony DC, Jiang Y, Broom KA, Campbell SJ, Tyler DJ, van Kasteren SI, Davis BG, Sibson NR (2009) *J Neurosci* 29:4820–4828
27. van Kasteren SI, Campbell SJ, Serres S, Anthony DC, Sibson NR, Davis BG (2009) *Proc Natl Acad Sci U S A* 106:18–23
28. Tanaka K, Masuyama T, Hasegawa K, Tahara T, Mizuma H, Wada Y, Watanabe Y, Fukase K (2008) *Angew Chem Int Ed Engl* 47(1):102–105.
29. Tanaka K, Fujii Y, Fukase K (2008) *ChemBioChem* 9:2392–2397

30. Lysine labeling & engineering kit “STELLA[™]” from Kishida Chemical Co., Ltd., <http://www.kishida.co.jp/>
31. Tanaka K, Minami K, Tahara T, Fujii Y, Siwu ERO, Nozaki S, Onoe H, Yokoi S, Koyama K, Watanabe Y, Fukase K (2010) *ChemMedChem* 5:841–845
32. Tanaka K, Siwu ERO, Minami K, Hasegawa K, Nozaki S, Kanayama Y, Koyama K, Chen WC, Paulson JC, Watanabe Y, Fukase K (2010) *Angew Chem Int Ed* 49:8195–8200
33. Tanaka K, Minami K, Tahara T, Siwu ERO, Koyama K, Nozaki S, Onoe H, Watanabe Y, Fukase K (2010) *J Carbohydr Chem* 29:118–132
34. Wang Q, Chan TR, Hilgraf R, Fokin VV, Sharpless KB, Finn MG (2003) *J Am Chem Soc* 125:3192–3193
35. Speers AE, Adam GC, Cravatt BF (2003) *J Am Chem Soc* 125:4686–4687
36. Link AJ, Tirrell DA (2003) *J Am Chem Soc* 125:11164–11165
37. Chin JW, Cropp TA, Anderson JC, Mukherji M, Zhang ZW, Schultz PG (2003) *Science* 301:964–967
38. Deiters AD, Cropp TA, Mukherji M, Chin JM, Anderson JC, Schultz PG (2003) *J Am Chem Soc* 125:11782–11783
39. Kho Y, Kim SC, Barma JD, Kwon SW, Cheng JK, Jaunbergs J, Weinbaum C, Tamanori F, Falck J, Zhao YM (2004) *Proc Natl Acad Sci U S A* 101:12479–12484
40. Speers AE, Cravatt BF (2004) *Chem Biol* 11:535–546
41. Zhan W-h, Barnhill HN, Sivakumar K, Tian H, Wang Q (2005) *Tetrahedron Lett* 46:1691–1695
42. Gierlich J, Burley GA, Gramlich PME, Hammond DM, Carell T (2006) *Org Lett* 8:3639–3642
43. Link AJ, Vink MKS, Agard NJ, Prescher JA, Bertozzi CR, Tirrell DA (2006) *Proc Natl Acad Sci U S A* 103:10180–10185
44. Yamaguchi M, Kojima K, Hayashi N, Kakizaki I, Kon A, Takagaki K (2006) *Tetrahedron Lett* 47:7455–7458
45. Lin P-C, Ueng S-H, Yu S-C, Jan M-D, Adak AK, Yu C-C, Lin C-C (2007) *Org Lett* 9:2131–2134
46. van Kasteren SI, Kramer HB, Jensen HH, Campbell SJ, Kirkpatrick J, Oldham NJ, Anthony DC, Davis BG (2007) *Nature* 446:1105–1109
47. Lin YA, Chalker JM, Floyd N, Bernardes GJL, Davis BG (2008) *J Am Chem Soc* 130:9642–9643
48. Becker CFW, Liu X, Olschewski D, Castelli R, Seidel R, Seeberger PH (2008) *Angew Chem Int Ed* 43:8215–8219
49. Song W, Wang Y, Qu J, Lin Q (2008) *J Am Chem Soc* 130:9654–9655
50. Chang PV, Chen X, Smyrniotis C, Xenakis A, Hu T, Bertozzi CR, Wu P (2009) *Angew Chem Int Ed* 48:4030–4033
51. Sivakumar K, Xie F, Cash BM, Long S, Barnhill HN, Wang Q (2004) *Org Lett* 6:4603–4606
52. Hanson SR, Hsu TL, Weerapana E, Kishikawa K, Simon GM, Cravatt BF, Wong CH (2007) *J Am Chem Soc* 129:7266–7267
53. Paulick MG, Forstner MB, Groves JT, Bertozzi CR (2007) *Proc Natl Acad Sci U S A* 104:20332–20337
54. Rabuka D, Forstner MB, Groves JT, Bertozzi CR (2008) *J Am Chem Soc* 130:5947–5953
55. Tanaka Y, Kohler J (2008) *J Am Chem Soc* 130:3278–3279
56. Saxon E, Bertozzi CR (2000) *Science* 287:2007–2010
57. Kiick KL, Saxon E, Tirrell DA, Bertozzi CR (2002) *Proc Natl Acad Sci U S A* 99:19–24
58. Vocado DJ, Hang HC, Kim EJ, Hanover JA, Bertozzi CR (2003) *Proc Natl Acad Sci U S A* 100:9116–9121
59. Chang PV, Prescher JA, Hangauer MJ, Bertozzi CR (2007) *J Am Chem Soc* 129:8400–8401
60. Hangauer MJ, Bertozzi CR (2008) *Angew Chem Int Ed* 47:2394–2397
61. Agard NJ, Prescher JA, Bertozzi CR (2004) *J Am Chem Soc* 126:15046–15047
62. van Berkel SS, Dirks ATJ, Debets MF, van Delft FL, Cornelissen JLM, Nolte RJM, Rutjes FPJT (2007) *ChemBioChem* 8:1504–1508
63. Lutz JF (2008) *Angew Chem Int Ed* 47:2182–2184

64. Fernandez-Suarez M, Baruah H, Martinez-Hernandez L, Xie KT, Baskin JM, Bertozzi CR, Ting AY (2007) *Nat Biotechnol* 25:1483–1487
65. Sletten EM, Bertozzi CR (2008) *Org Lett* 10:3097–3099
66. Luchansky SJ, Bertozzi CR (2004) *ChemBioChem* 5:1706–1709
67. Agard NJ, Baskin JM, Prescher JA, Lo A, Bertozzi CR (2006) *ACS Chem Biol* 1:644–648
68. Prescher JA, Bertozzi CR (2006) *Cell* 126:851–854
69. Baskin JM, Prescher JA, Laughlin ST, Agard NJ, Chang PV, Miller IA, Lo A, Codelli JA, Bertozzi CR (2007) *Proc Natl Acad Sci U S A* 104:16793–16797
70. Codelli JA, Baskin JM, Agard NJ, Bertozzi CR (2008) *J Am Chem Soc* 130:11486–11493
71. Ning X, Guo J, Wolfert MA, Boons G-J (2008) *Angew Chem Int Ed* 47:2253–2255
72. Tanaka K, Kamatani M, Mori H, Fujii S, Ikeda K, Hisada M, Itagaki Y, Katsumura S (1998) *Tetrahedron Lett* 39:1185–1188
73. Tanaka K, Kamatani M, Mori H, Fujii S, Ikeda K, Hisada M, Itagaki Y, Katsumura S (1999) *Tetrahedron* 55:1657–1686
74. Tanaka K, Katsumura S (2000) *Org Lett* 2:373–375
75. Tanaka K, Mori H, Yamamoto M, Katsumura S (2001) *J Org Chem* 66:3099
76. Tanaka K, Katsumura S (2002) *J Am Chem Soc* 124:9660
77. Tanaka K, Kobayashi T, Mori H, Katsumura S (2004) *J Org Chem* 69:5906
78. Kobayashi T, Nakashima M, Hakogi T, Tanaka K, Katsumura S (2006) *Org Lett* 8:3809–3812
79. Kobayashi T, Hasegawa F, Tanaka K, Katsumura S (2006) *Org Lett* 8:3813–3816
80. Sakaguchi T, Kobayashi T, Hatano S, Tsuchikawa H, Fukase K, Tanaka K, Katsumura S (2009) *Chem Asian J* 4:1573–1577
81. Tanaka K, Fukase K, Katsumura S (2010) *Chem Rec* 10:119–139
82. Crocker PR, Varki A (2001) *Immunology* 103:137–145
83. Crocker PR, Varki A (2001) *Trends Immunol* 22:337–342
84. Angata T (2006) *Mol Divers* 10:555–566
85. Varki A, Angata T (2006) *Glycobiology* 16:1R–27R
86. Cyster JG, Goodnow CC (1997) *Immunity* 6:509–517
87. Collins BE, Blixt O, DeSieno AR, Bovin N, Marth JD, Paulson JC (2004) *Proc Natl Acad Sci U S A* 101:6104–6109
88. Crocker PR, Paulson JC, Varki A (2007) *Nat Rev Immunol* 7:255–266
89. Manning DD, Hu X, Beck P, Kiessling LL (1997) *J Am Chem Soc* 119:3161–3162
90. Reuter JD, Myc A, Hayes MM, Gan Z, Roy R, Qin D, Yin R, Piehler LT, Esfand R, Tomalia DA, Baker JR Jr (1999) *Bioconjug Chem* 10:271–278
91. Andre S, Kaltner H, Furuike T, Nishimura S-I, Gabius H-J (2004) *Bioconjug Chem* 15:87–98
92. Ladmiral V, Mantovani G, Clarkson GJ, Cauet S, Irwin JL, Haddleton DM (2006) *J Am Chem Soc* 128:4823–4830
93. Carlson CB, Mowery P, Owen RM, Dykhuizen EC, Kiessling LL (2007) *ACS Chem Biol* 2:119–127
94. Wolfenden ML, Cloninger MJ (2006) *Bioconjug Chem* 17:958–966
95. Kim J, Ahn Y, Park KM, Kim Y, Ko YH, Oh DH, Kim K (2007) *Angew Chem Int Ed* 46:7393–7395
96. Lee MJ, Pal K, Tasaki T, Roy S, Jiang Y, An JY, Banerjee R, Kwon YT (2008) *Proc Natl Acad Sci U S A* 105:100–105
97. Roy R, Beak M-G (2002) *Rev Mol Biotechnol* 90:291–309
98. Leeuwenburgh MA, van der Marel GA, Overkleeft HS (2003) *Curr Opin Chem Biol* 7:757–765
99. Renaudet O (2008) *Mini-Rev Org Chem* 5:274–286
100. Chabre YM, Roy R (2008) *Curr Top Med Chem* 8:1237–1285
101. Chabre YM, Roy R (2010) *Advances in carbohydrate chemistry and biochemistry*, vol 63. Elsevier, UK, pp 165–393

102. Kajihara Y, Suzuki Y, Yamamoto N, Sasaki K, Sakakibara T, Juneja LR (2004) *Chem Eur J* 10:971–985
103. Tanaka K, Kageyama C, Fukase K (2007) *Tetrahedron Lett* 48:6475–6479
104. Fukuyama Y, Nakaya S, Yamazaki Y, Tanaka K (2008) *Anal Chem* 80:2171–2179
105. Maেকে HR, Hofmann M, Haberkorn U (2005) *J Nucl Med* 46:172S–178S
106. Haraldsson B, Nyström J, Deen WM (2008) *Physiol Rev* 88:451–487
107. Edwards A, Daniels BS, Deen WM (1999) *Am J Physiol Renal Physiol* 276:F892–F902
108. Wyatt AR, Yerbury JJ, Poon S, Wilson MR (2009) *Curr Med Chem* 16:2855–2866
109. Tozawa R, Ishibashi S, Osuga J, Yamamoto K, Yagyu H, Ohashi K, Tamura Y, Yahagi N, Iizuka Y, Okazaki H, Harada K, Gotoda T, Shimano H, Kimura S, Nagai R, Yamada N (2001) *J Biol Chem* 276:12624–12628
110. Unverzagt C, Andre S, Seifert J, Kojima S, Fink C, Srikrishna G, Freeze H, Kayser K, Gabius H-J (2002) *J Med Chem* 45:478–491
111. Park EI, Mi Y, Unverzagt C, Gabius H-J, Baenziger JU (2005) *Proc Natl Acad Sci U S A* 102:17125–17129
112. Lee YC, Townsend RR, Hardy MR, Lonngren J, Arnarp J, Haraldsson M, Lonn H (1983) *J Biol Chem* 258:199–202
113. Nakagawa T, Uozumi N, Nakano M, Mizuno-Horikawa Y, Okuyama N, Taguchi T, Gu J, Kondo A, Taniguchi N, Miyoshi E (2006) *J Biol Chem* 281:29797–29806
114. Miyoshi E, Noda K, Yamaguchi Y, Inoue S, Ikeda Y, Wang W, Ko JH, Uozumi N, Li W, Taniguchi N (1999) *Biochim Biophys Acta* 1473:9–20
115. Comelli EM, Head SR, Gilmartin T, Whisenant T, Haslam SM, North SJ, Wong NK, Kudo T, Narimatsu H, Esko JD, Drickamer K, Dell A, Paulson JC (2006) *Glycobiology* 16:117–131
116. Tateno H, Li HY, Schur MJ, Bovin N, Crocker PR, Wakarchuk WW, Paulson JC (2007) *Mol Cell Biol* 27:5699–5710
117. O'Reilly MK, Collins BE, Han S, Liao L, Rillahan C, Kitov PI, Bundle DR, Paulson JC (2008) *J Am Chem Soc* 130:7736–7745
118. Grobбен B, De Deyn PP, Slegers H (2002) *Cell Tissue Res* 310:257–270
119. Hannah MJ, Weiss U, Huttner WB (1998) *Methods* 16:170–181
120. Broekaert WF, Nsimba-Lubaki M, Peeters B, Peumans WJ (1984) *Biochem J* 221:163–169
121. Matsui K, Wang Z, McCarthy TJ, Allen PM, Reichert DE (2004) *Nucl Med Biol* 31:1021–1031
122. Radu CG, Shu CJ, Nair-Gill E, Shelly SM, Barrio JR, Satyamurthy N, Phelps ME, Witte ON (2008) *Nat Med* 14:783–788
123. Yeh T-C, Zhang W, Ildstad ST, Ho C (1993) *Magn Reson Med* 30:617–625
124. Bulte JWM, Duncan ID, Frank JAJ (2002) *J Cereb Blood Flow Metab* 22:899–907
125. Hoehn M, Kustermann E, Blunk J, Wiedermann D, Trapp T, Wecker S, Focking M, Arnold H, Hescheler J, Fleischmann BK, Schwindt W, Buhle C (2002) *Proc Natl Acad Sci U S A* 99:16267–16272
126. Toyoda K-I, Tooyama I, Kato M, Sato H, Morikawa S, Hisa Y, Inubushi T (2004) *NeuroReport* 15:589–593
127. Baxevanis CN, Papamichail M (2004) *Cancer Immunol Immunother* 53:893–903
128. Liang C-H, Wang S-K, Lin C-W, Wang C-C, Wong C-H, Wu C-Y (2011) *Angew Chem* 123:1646–1650

Trypanosomal Trans-sialidases: Valuable Synthetic Tools and Targets for Medicinal Chemistry

Sebastian Meinke and Joachim Thiem

Abstract In contrast to the general hydrolases, trans-sialidase from *Trypanosoma cruzi* (TcTS) shows excellent regio- and stereoselectivity as well as high yields in transfer reactions. Discussed are the occurrence of trans-sialidases and studies on the transfer mechanism. In detail, the preparative use by chemoenzymatic syntheses with TcTS are outlined with emphasis on the design of modified donor and acceptor substrates. Another section focuses on attempts to develop inhibitors for TcTS, and these endeavors are based on donor- and acceptor-inspired modifications as well as on some completely different structures.

Keywords Chemoenzymatic synthesis · Donor and acceptor substrates · Donor- and acceptor-based inhibitors · Inhibitor design · Modifications · Preparative use · Trans-sialidase

Contents

1	Trypanosomes	232
1.1	Introduction	232
1.2	<i>Trypanosoma cruzi</i>	232
1.3	Other Trypanosomes	233
2	Trans-sialidases	234
2.1	Sialylation with TcTS	235
2.2	One-Pot Syntheses	236
2.3	Modified Substrates	236

3	Inhibitors	244
3.1	Donor-Based Inhibitors	244
3.2	Acceptor-Based Inhibitors	245
3.3	Combined Approaches	246
3.4	Further Structures	247
	References	248

1 Trypanosomes

1.1 Introduction

Trypanosomes are protozoic parasites transmitted by insects and are responsible for a number of endemic zoonotic diseases in Africa and South America. The South American trypanosomiasis (Chagas disease) [1] is caused by *Trypanosoma cruzi*. In Africa *Trypanosoma brucei gambiense* causes the chronic form and *T. brucei rhodesiense* causes the acute form of sleeping sickness. *T. brucei brucei*, the pathogenic agent of acute Nagana (a cattle disease), is not infectious for humans. There is also a chronic form of Nagana, caused by *T. congolense*.

1.2 *Trypanosoma cruzi*

The South American trypanosomiasis was discovered by the Argentine physician Carlos Chagas [2]. The causative agent *T. cruzi* lives in the gut of blood-sucking insects of the species *Triatoma*, *Rhodnius*, and *Panstrongylus* (Fig. 1) and is transmitted by rubbing their feces into the injection site or mucous membranes (conjunctiva of the eyes) after a blood-meal. Other possible routes of infection are blood transfusion and prenatal infection. Once in the blood, the Trypanosomes penetrate tissue cells, in which they proliferate before being shed into the bloodstream again. Possible hosts are vertebrates. Rodents play a particular role as intermediate hosts.

Chagas disease is a widespread disease in Central and South America. According to the *Pan-American Health Organization* in 2005 there were 7.7 million infected persons, 42,200 new infections by insect bite, and 14,385 new cases by transmission in the womb. These were distributed over 21 affected countries with 108,600,000 people at risk of infection. A critical country is Bolivia, where 6.8% of the population are infected with *T. cruzi*.

An infection leads to a first acute form of Chagas disease with fever, inflammation of lymph nodes, enlargement of liver and spleen, and possibly myocarditis. However, in most cases the acute form proceeds without symptoms and is therefore not recognized. Afterwards the chronic form develops over 10–20 years and leads to chronic cardiopathy, damage to the digestive system, and possibly neurological symptoms [4].

T. cruzi is unable to synthesize sialic acid de novo [5]. Therefore, the trypomastigote form of *T. cruzi* existing in the bloodstream of the host (Fig. 2) expresses a trans-sialidase (TcTS), which either cleaves or transfers terminal sialic acid residues from host cells and in the transfer process sialylates the surface



Fig. 1 *Rhodnius prolixus* [3]

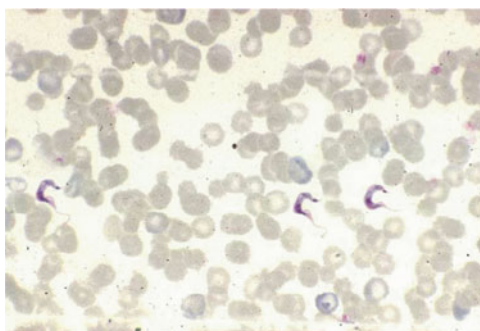


Fig. 2 *Trypanosoma cruzi* trypomastigotes in blood [6]

glycans of the pathogen. These sialic acid residues are responsible for a number of interactions of the pathogen with host cells, which are related to cell invasion and protection against the host immune system [7–12].

1.3 Other Trypanosomes

Though TcTS is by far the most studied enzyme of its class, *T. cruzi* is not the only trypanosome species that expresses a trans-sialidase. While some trypanosomes like the American *T. rangeli* possess only a sialidase that merely hydrolyses terminal sialyl residues in spite of highly sequential similarity to TcTS [13, 14], the trans-sialidase of African trypanosome *T. brucei brucei* (causing nagana cattle disease) was also investigated and revealed differences to TcTS in the overall structure but a highly conserved catalytic center [15]. Recently, a trans-sialidase of *T. congolense* was recombinantly expressed and intensively studied [16]. Whereas *T. cruzi* expresses a trans-sialidase while in the human blood stream, in African trypanosomes this enzyme seems to facilitate survival in the vector, the tsetse fly [17, 18]. Interestingly, there is also a known trans-sialidase in the fish parasite *T. carassii* [19].

2 Trans-sialidases

Trans-sialidases belong to the *exo- α -sialidase* family (EC 3.2.1.18), but instead of hydrolysis they show a strong preference for transfer of sialic acids. The trans-sialidase of *T. cruzi* (TcTS), which is by far the most investigated, was discovered and characterized as an *exo- α -trans-sialidase* [20–23], which showed weak hydrolysis in the absence of a suitable acceptor molecule. By papain degradation a 70-kD fragment with full catalytic activity was obtained [24] and, based on this work, recombinant expression of TcTS and finally crystallization [25, 26] was facilitated.

Investigating the mechanism of TcTS revealed the role of a tyrosine as the catalytic nucleophile [27, 28]. The enzyme works by a ping-pong-mechanism [29] involving a covalent intermediate and results in retention of configuration (Fig. 3) [30]. The preference for transfer over hydrolysis is facilitated by the aromatic amino acids tyrosine (Tyr₁₁₉) and tryptophan (Trp₃₁₂), which form the acceptor binding site (Fig. 4) [13, 31–33].

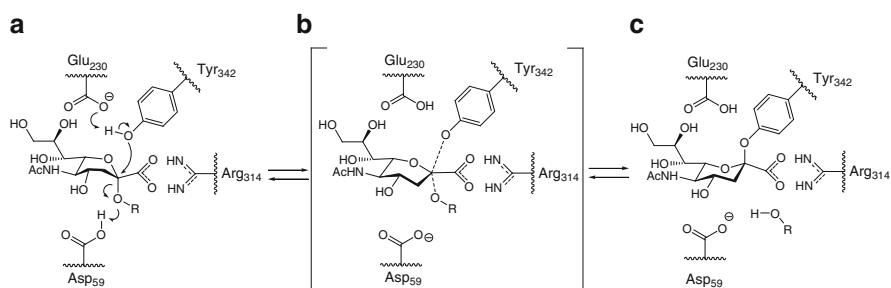


Fig. 3 Ping-pong mechanism of TcTS

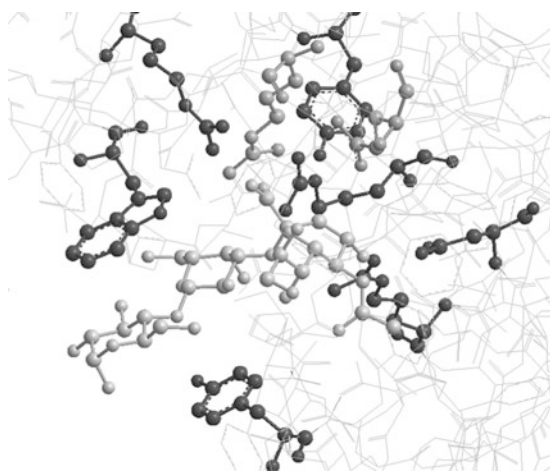


Fig. 4 Active site of TcTS [26]

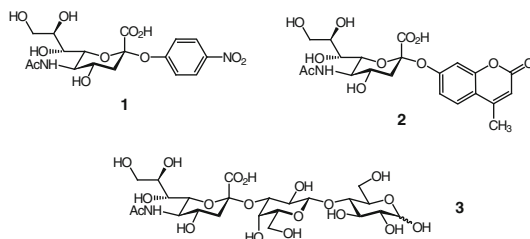


Fig. 5 Potential donor substrates for chemoenzymatic synthesis with TcTS [35]

2.1 Sialylation with TcTS

2.1.1 Donor Substrates

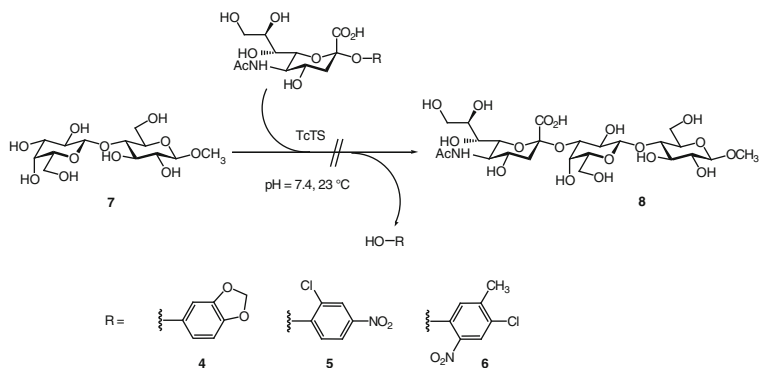
For application of TcTS as a synthetic tool there are different potential donor substrates. As natural donors, oligosaccharides with a pre-terminal galactopyranosyl structure terminally 3α -sialylated represent highly active donor substrates. Thus, even sialyllactose (Neu5Ac α 2-3Gal β 1-4Glc, **3**) could be successfully employed. Sialylated glycoproteins such as fetuin also proved to be useful as rather active donors [34]. Earlier various phenyl α -sialosides were proposed by Scudder et al. [23] to be quite useful in TcTS sialylations. These are *p*-nitrophenyl α -sialoside (pNP-Neu5Ac, **1**) as well as 4-methylumbelliferyl α -sialoside (MU-Neu5Ac, **2**), both of which proved advantageous even though their transfer rates are rather low. However, in contrast to the fast reaction employing sialyllactose **3**, the remaining cleavage product lactose enjoys re-sialylation. Thus the transfer is difficult to control in that case, whereas none of the leaving phenols will be re-sialylated (Fig. 5).

Following the classical syntheses of aromatic α -sialosides by phase transfer catalysis starting with the per-*O*-acetylated sialosyl chloride methyl ester, further aromatic α -sialosides could be prepared with sesamol (**4**), 2-chloro-4-nitrophenol (**5**), and 4-chloro-5-methyl-4-nitrophenol (**6**) [36]. In TcTS-sialylations none of these modified donor substrates showed any sialylation of methyl β -lactoside (**7**) to give the methyl sialyllactoside (**8**). It may be assumed that both steric and electronic factors affect interactions with aromatic and/or hydrophobic amino acid residues in the enzyme (Scheme 1).

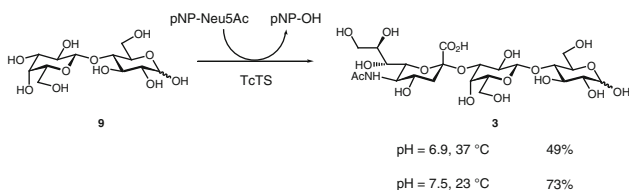
2.1.2 Conditions

As reported earlier by Scudder et al. [23], TcTS shows an apparent pH optimum at 7.9 and a temperature optimum at 13 °C. In general, slightly modified conditions – mostly due to solubility issues – would be followed throughout TcTS sialylations with a large number of substrates.

It was shown that sialylation of lactose (**9**) employing pNP-Neu5Ac at pH 6.9 and 37 °C gave sialyllactose (**3**) in 49% yield [37]. By changing to pH 7.5 and 23 °C the yield could be enhanced to 73% (Scheme 2) [36].



Scheme 1 Further phenyl sialosides investigated with TcTS [36]



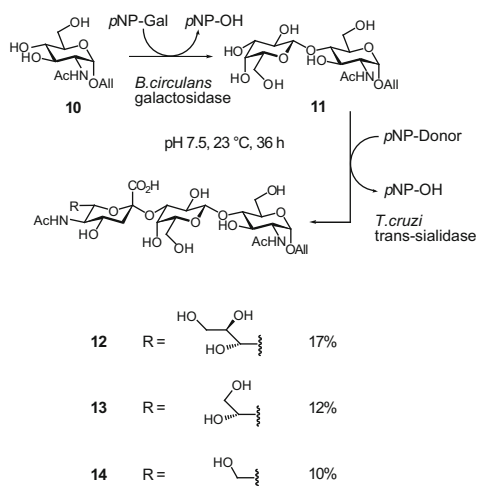
Scheme 2 Conditions for trans-sialylation with TcTS [36]

2.2 One-Pot Syntheses

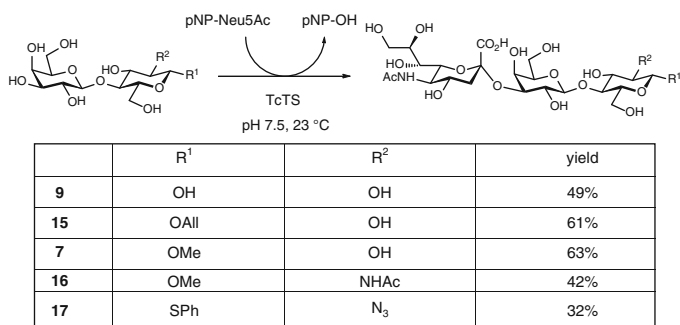
Previously, it could be demonstrated that the sequential use of different glycosidases was efficiently used for construction of some complex lower oligosaccharides. Thus, e.g., following a galactosylation of GlcNAc with β -galactosidase (bovine testes, *Bacillus circulans*) an α -sialylation with various α -sialidases (*Vibrio cholerae*, *Clostridium perfringens*, *Salmonella typhimurium*, Newcastle disease virus) was employed to construct several sialyl oligosaccharides [38]. As a follow-up it was of interest to check a corresponding approach for combination of β -galactosidase (*B. circulans*) with TcTS. First allyl *N*-acetyl- α -glucosaminide **10** was galactosylated with pNP-Gal, and then three different pNP-donors (see Sect. 2.3.2.) were employed with TcTS to give the corresponding trisaccharides **12–14** in 10–17% yield (Scheme 3).

2.3 Modified Substrates

For both mechanistic and preparative reasons it is of interest to prove the structural requirements and preparative limits for employing modified substrates in enzymatic transfers.



Scheme 3 One-pot chemoenzymatic transformations employing a β -galactosidase and TcTS [36]



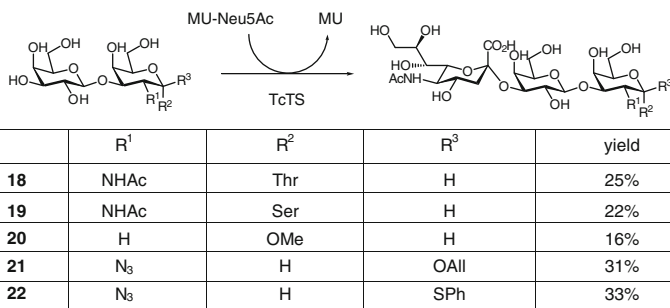
Scheme 4 Lactose derived acceptor substrates

2.3.1 Acceptor Substrates

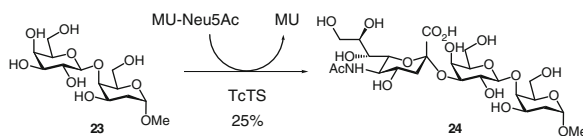
Thus, a series of lactose-derived acceptor substrates were successfully sialylated with TcTS and pNP-Neu5Ac under optimum conditions to give the trisaccharide derivatives (Scheme 4) in up to more than 60% yield. In all cases modifications were in the pre-terminal glucopyranose unit, and apparently no particularly negative effect could be noted, thus allowing considerable broad acceptor variations.

Along similar lines sialylations with TcTS and MU-Neu5Ac were shown to proceed smoothly with Gal β 1-3GalNac-modified acceptor substrates (Scheme 5). Even amino acid glyco derivatives, as well as deoxy and azido functionalized structures, could be sialylated. Furthermore, Gal β 1-4Gal structures were also readily transformed (Scheme 6) [35, 36].

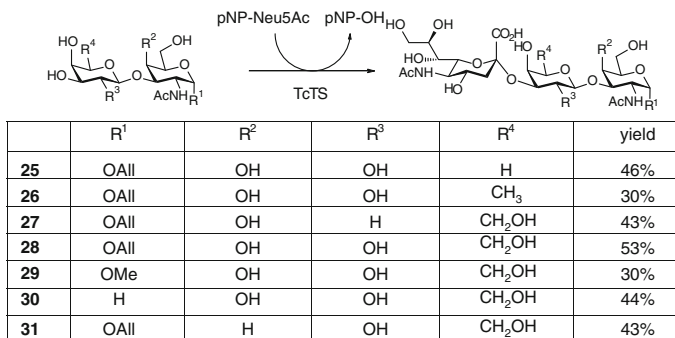
In further studies both the terminal Gal and the pre-terminal GalNac units could be modified [39] and these in turn were nicely sialylated by TcTS and pNP-Neu5Ac to



Scheme 5 Gal β (1-3)-GalNAc derived acceptor substrates



Scheme 6 A Gal β (1-4)-Gal-R derived acceptor substrate



Scheme 7 Gal β (1-3)-GalNAc derived acceptor substrates

give the Neu5Ac α 2-3Gal β 1-4GalNAc-modified trisaccharides in 30–50% yield comprising modified Thomsen-Friedenreich antigen components (Scheme 7) [36, 40, 41].

A further recent work in a corresponding direction was by Harrison et al. [42] in which TcTS-catalyzed sialylations by pNP-Neu5Ac of both widely modified octyl galactopyranoside (at positions 3, 4, and 6) and octyl *N*-acetylglucosaminide (at positions 3, 6, 2', 3', 4', and 6') were studied. Apparently, as expected, some modifications in the acceptor substrates were tolerated, whereas other more drastic alterations did not show any turnover (Fig. 6) [42].

In another nice contribution the sialylation of cyclic pseudo-oligosaccharides, derived by click chemistry, by use of TcTS and MU-Neu5Ac were studied to give di- as well as tri-sialylated products. In all cases α -sialylation was obtained at all terminal Gal-3-positions (Fig. 7) [43].

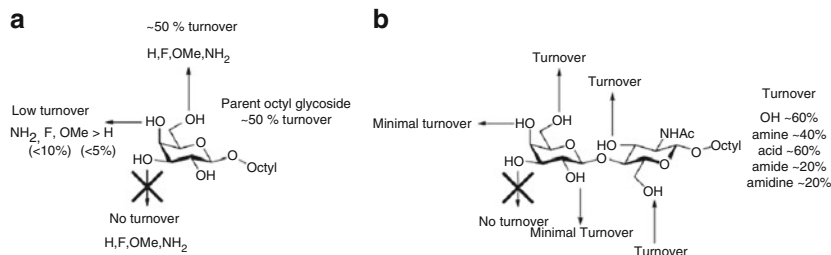


Fig. 6 TcTS-catalyzed sialylation with modified octyl galactopyranosides (a) and octyl LacNAc (b) derivatives [42]

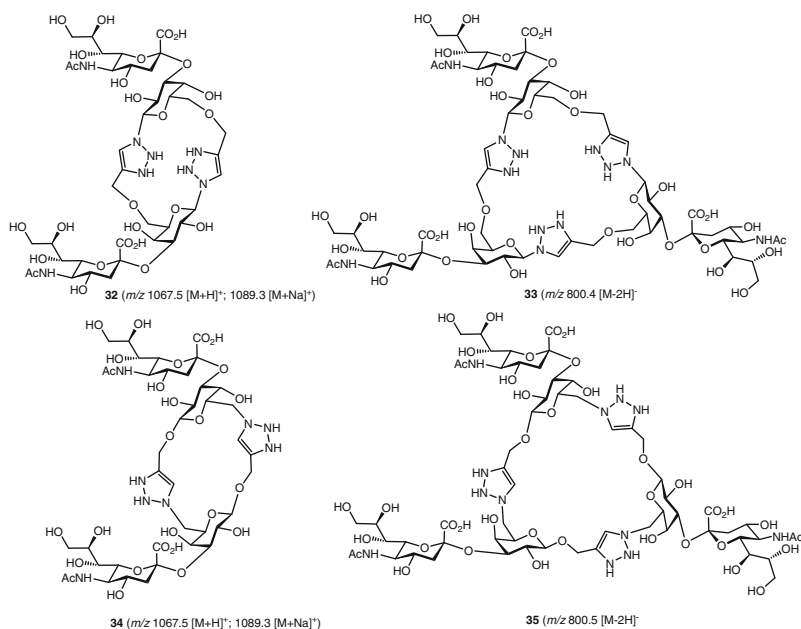
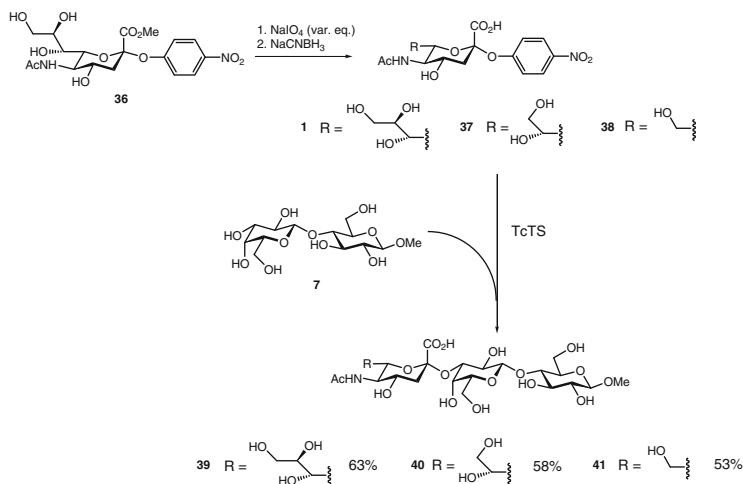


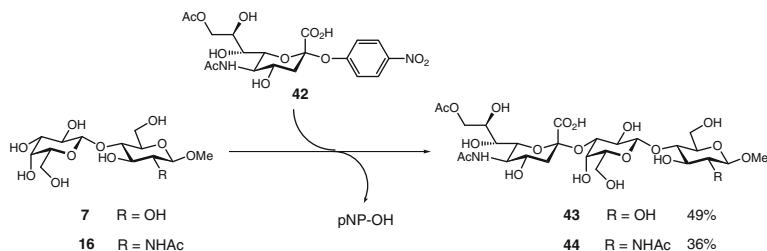
Fig. 7 Cyclic pseudo-oligosaccharides sialylated by TcTS [43]

2.3.2 Modified Donor Substrates

It would be of considerable interest to understand the donor substrate alignment within the enzyme's active site pocket. Therefore it was considered of value to synthesize modified donors that resemble the well known donor substrates. Initially, the glycerol side chain C7-C9 in Neu5Ac represented an interesting target. By chemoselective periodate cleavage of pNP-Neu5Ac and subsequent reduction, both the corresponding octoside and heptoside derivatives were obtained in good yield. They were used in TcTS-catalyzed reactions with methyl lactoside as acceptor. Pleasingly, the enzyme recognized these truncated donor substrates just like the original pNP-Neu5Ac to give



Scheme 8 TcTS-catalyzed glycosylation with Neu5Ac-derived truncated pNP-octoside and heptoside [44]

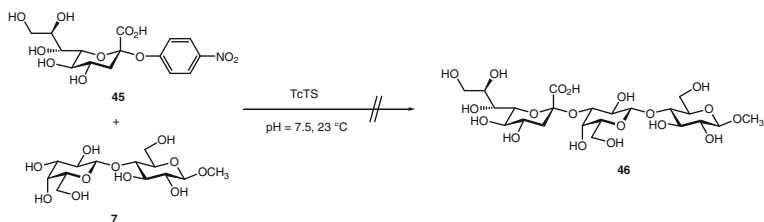


Scheme 9 TcTS-catalyzed sialylation of methyl β -lactoside and methyl β -lactosaminide with pNP-Neu5,9Ac₂ [36]

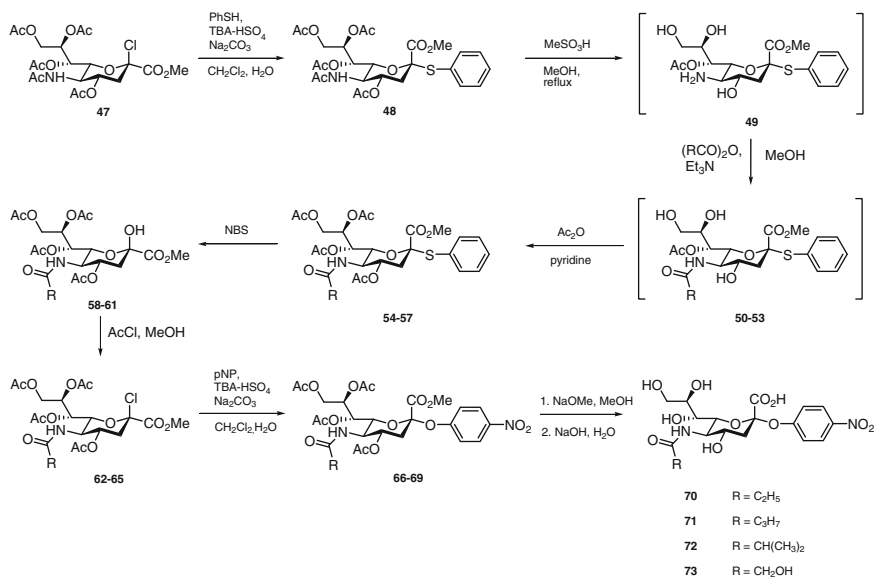
the corresponding trisaccharides in yields of 53–63% (Scheme 8). Furthermore, it would seem that the glycerol side chain is not required for the enzymatic reaction. In contrast, STD-NMR studies showed that this structural unit is a key determinant in the binding process of sialylated oligosaccharides in their interaction with the myelin-associated glycoprotein MAG [44].

A minor modification of pNP-Neu5Ac, that is the regiospecific acetylation at position 9, gave donor **42** which could be used under normal conditions to sialylate methyl β -lactoside as well as methyl β -*N*-acetylactosaminide to give the corresponding trisaccharides in 49% and 36% yield, respectively (Scheme 9) [36].

Structurally close to Neu5Ac is the deaminated form ketodeoxy nonulosonic acid (KDN). The synthesis was conducted according to the Cornforth method [45] following Ogura et al. [46] from mannose and oxalacetate. Per-*O*-acetylation was followed by formation of the benzyl ester and transformation into the β -chloride. Then phase transfer glycosylation with *p*-nitrophenol followed by deacetylation and ester cleavage gave pNP-KDN **45**. As in the above cases, glycosylation of methyl β -lactoside



Scheme 10 Attempts for TcTS-catalyzed glycosylation with pNP-KDN [36]



Scheme 11 Synthesis of *N*-acylated pNP-Neu5N donor derivatives [47, 48]

with this donor substrate and TcTS was attempted but did not result in the expected trisaccharide. Obviously this deviation at position 5 of the donor was not tolerated by the enzyme (Scheme 10) [36].

It would seem that the 5-*N*-acyl group in Neu5Ac donors represents a critical structural element. Thus, it was of interest to check whether the *N*-acetyl function could at least be modified if not omitted. Starting with the phenylthio glycoside of Neu5Ac, treatment with methane sulfonic acid cleaved all the acetates. Then re-*N*-acylation with acyl anhydrides was followed by *O*-acetylation and *N*-bromosuccinimide cleavage of the thiophenyl group. Further, the classical four-step formation of the modified *p*-nitrophenyl sialosides could be achieved [47].

Thus a series of *N*-acyl derivatives was obtained (Scheme 11). In addition the 9-*O*-butyryl derivative as well as the 5-amino- and 5-azido-modified pNP-Neu5Ac donor derivatives could be prepared (Schemes 12 and 13).

All of these donor substrates could be studied in trans-sialylations with TcTS and methyl β -allolactoside as an efficient acceptor. Interestingly, next to the good and

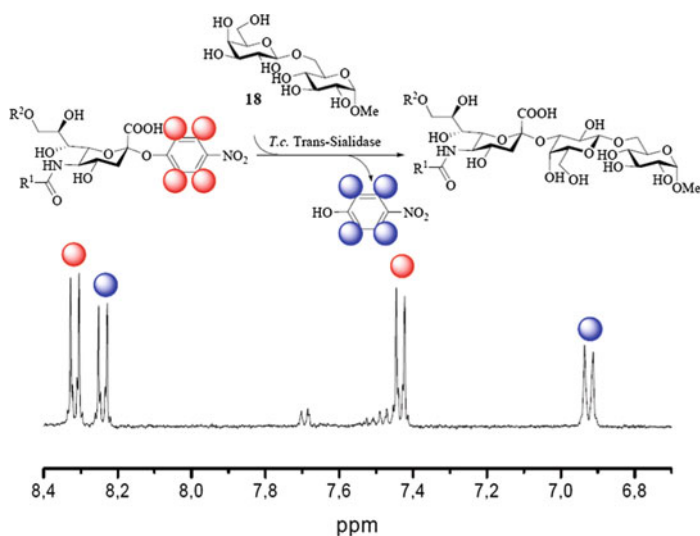


Fig. 8 NMR assay for measurement of transfer rates [47, 48]

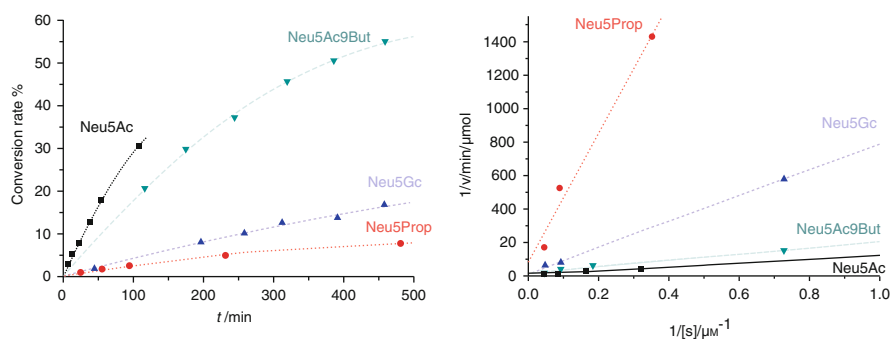


Fig. 9 Trans-sialylation of *N*-acylated pNP-Neu5N derivatives and Lineweaver–Burke plot [47, 48]

Table 1 Michaelis–Menten parameters of accepted donor substrates in D₂O (9But = 9-*O*-butyryl, Gc = *N*-glycolyl, Prop = *N*-propionyl) [47, 48]

Donor molecule	K_M (μM)	v_{max} (nmol/min)
<i>p</i> NP-Neu5Ac	10	91
<i>p</i> NP-Neu5Ac9But	10	51
<i>p</i> NP-Neu5Gc	31	41
<i>p</i> NP-Neu5Prop	45	12

N-glycolyl derivative Neu5NGly turned out to be quite a good donor substrate, and even the Neu5NProp derivative was still accepted. Again, a considerable tolerance was observed for derivatives modified in the glycerol side chain (Table 1) [47, 48].

3 Inhibitors

Presently, the only therapeutic actions available in the case of Chagas disease are treatment with Nifurtimox (**86**) and Benznidazole (**87**), both of which show severe side effects such as cytotoxicity as well as genotoxicity (Fig. 10) [4]. They are only effective in the acute phase of the disease, their efficiency is limited since the infection remains intact in 80% of all cases [4], and the chronic phase cannot be inhibited [49]. Thus, up until now the most successful method to confine Chagas disease has been limited to the fight against the vectors (*Triatoma infestans*, *Rhodnius prolixus*, and *Panstrongylus megistus*) employing insecticides [4, 50].

Based on the invasion mechanism for mammals by *Trypanosoma cruzi*, a promising approach to tackle Chagas disease would be to modulate or better inhibit trans-sialidase as recently discussed by Neres et al. [51].

3.1 Donor-Based Inhibitors

As previously shown, 3-fluoro sialosyl fluorides could be employed as donor substrates for TcTS. Due to the fluoro substituent and its pronounced electron-withdrawing features at the anomeric center, these derivatives show delayed reactions with acceptor structures [28]. This effect proved to be useful in determination of essential mechanistic aspects, since the intermediate enzyme substrate complex could be crystallized and studied [25, 26]. By further modification via introduction of a 9-benzamide function as in compound **24** this effect was increased. In contrast to the expected outcome the enhanced binding of this aromatic residue was unexpectedly due to its location in the acceptor-binding site rather than in the glycerol sidechain binding domain, as evidenced by X-ray structural studies (Fig. 11) [52].

Recently, Carvalho et al. [53] showed that the aromatic α -sialosides 2-difluoromethyl-4-nitro-phenyl (**89**, Neu5AcFNP) as well as the dansyl-Neu5AcFP derivative (**90**) act as time- and dose-dependent irreversible inactivators of TcTS. In MALDI-TOF/TOF-MS, fragments of enzyme inhibitor structures were detected which proved that inactivation of the enzyme was due to covalent binding between the aromatic groups and the Arg₂₄₅ and Asp₂₄₇ residues (Fig. 12).

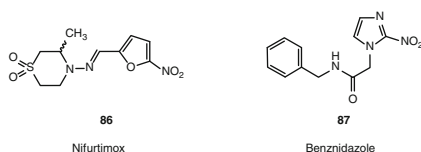


Fig. 10 Chemotherapeutics used in acute phase of Chagas disease

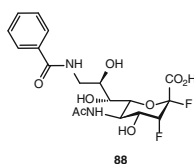


Fig. 11 3-Fluoro sialosyl fluoride derivative with acceptor site binding properties in TcTS [52]

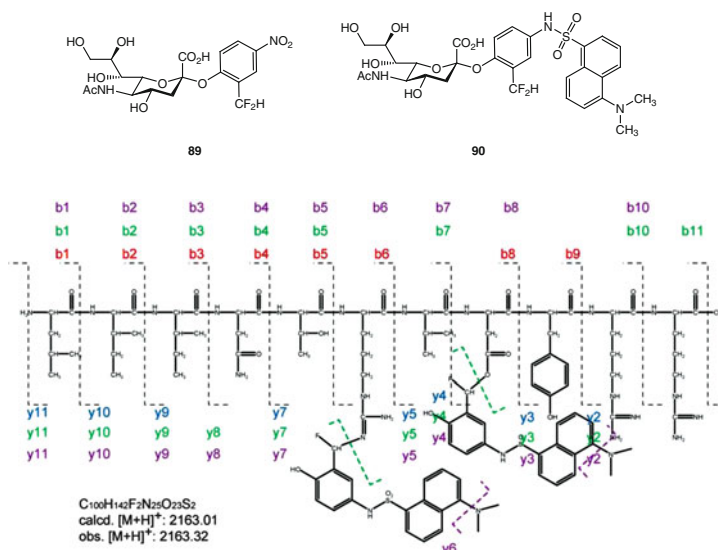


Fig. 12 FNP-Neu5Ac (**89**) and dansylFNP-Neu5Ac (**90**) and covalently linked enzyme-inhibitor structures in MS studies [53]

3.2 Acceptor-Based Inhibitors

Testing of various acceptor substrates with TcTS revealed that stronger binding competing substrates show some inhibitory activity due to their preferred transfer. Thus, even the simple lactitol **91** ($K_M = 0.26$ mM) proved to be a better acceptor substrate than *N*-acetylglucosamine ($K_M = 0.57$ mM) [54].

By employing the omnipresent “click reaction” a substrate library of substituted triazole galactopyranosides was prepared. The most pronounced effect was found for compound **92**, for which some inhibition, 37% reduction of activity at $c = 1$ mM, was detected [55].

Another recent work proposed *N*-PEGylated Gal β 1-6GlcN glycosides (**93**) as potential inhibitors (Fig. 13) [56].

A library of β -thiogalactopyranosides with quite structurally different *S*-substitutions (Fig. 14) showed IC_{50} in the low millimolar range [42].

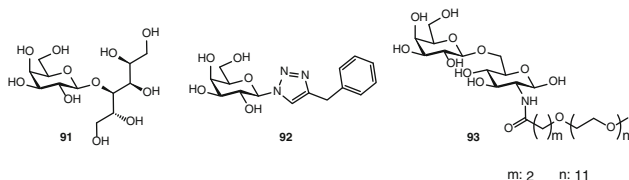


Fig. 13 Lactitol **91** and galactopyranosyl triazole **92** as competitive acceptor substrates and the potential inhibitor structure *N*-PEGylated Gal β 1-6GlcN **93** [54–56]

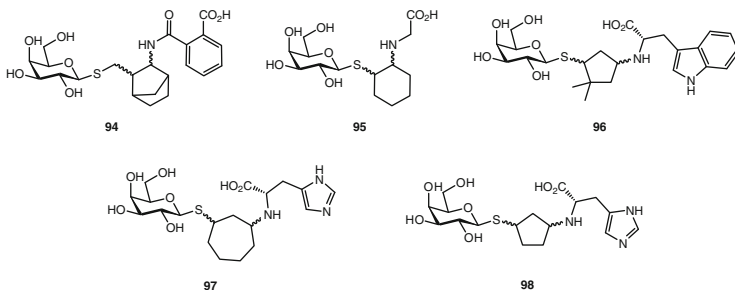


Fig. 14 Some TcTS inhibitors from a β -thiogalactopyranoside library [42]

3.3 Combined Approaches

Due to the transfer mechanism of TcTS analogous to those of sialidases [25] the prototype of sialidase inhibitors Neu5Ac2en **99** [57] was tested. Compound **99** as well as the potent influenza virus neuraminidase inhibitors zanamivir (Relenza, **100**) and the active form of oseltamivir (Tamiflu, **101**) with analogous structures of transition state were also of interest as potential inhibitors of TcTS. However, as determined for **99** the TcTS activity was inhibited with only $K_i = 12.29$ mM (Fig. 15) [33].

Combination of this structural motif with an acceptor analog led to somewhat more potent, yet still considerably weak, inhibitors. Thus, for compound **102** $IC_{50} \sim 1.5$ mM could be determined (Fig. 16) [58, 59].

Recently, based on inspection of the crystal structure of the complex of TcTS with sialyllactose as substrate, a series of novel *C*-sialosides were devised as potential inhibitors. Employing ruthenium-catalyzed cross metatheses, a number of derivatives were synthesized and modified [60]. For these novel acceptor mimetics the affinities were studied by surface plasmon resonance (SPR) using immobilized TcTS [61]. Quite interesting K_D values for compounds **103–105** were measured (Fig. 17). Furthermore, using an NMR-based approach [47], considerable reduction of the TcTS reaction rate was observed in the presence of the most active *C*-sialoside ligand **103** (Fig. 18) [61].

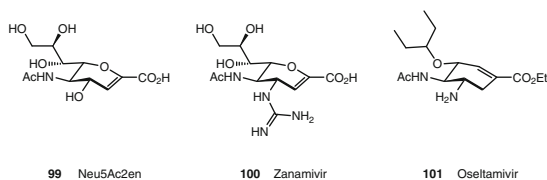


Fig. 15 Sialidase inhibitors tested with TcTS [33, 57]

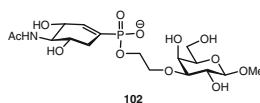


Fig. 16 Transition state analog and acceptor mimetic combined [58, 59]

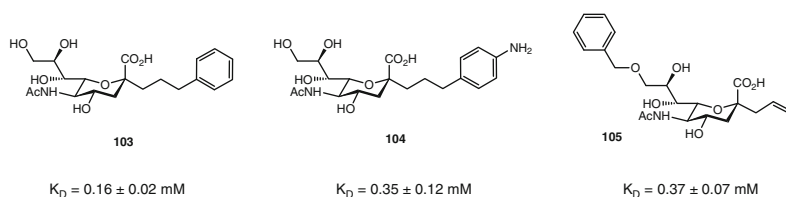


Fig. 17 C-Sialoside TcTS ligands with acceptor mimetic; affinities to immobilized TcTS by SPR [61]

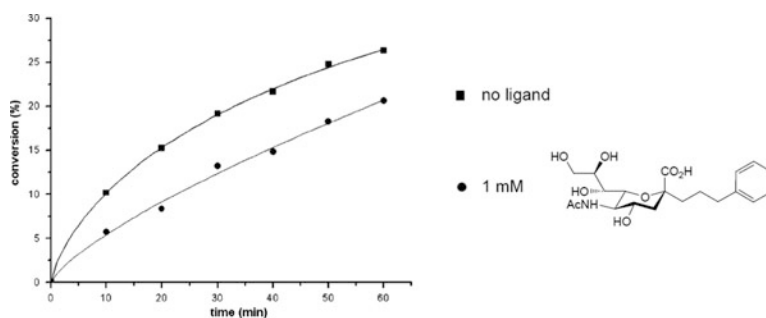
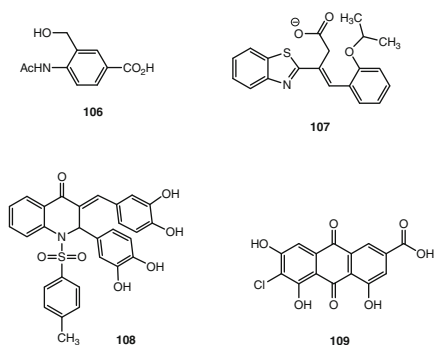


Fig. 18 Influence on TcTS reaction rates by C-sialoside ligand **103** [61]

3.4 Further Structures

Some time ago *N*-acyl anilines were tested and found to inhibit the sialidase from *Vibrio cholerae*; however, they did not show any effect with TcTS [62]. There is some similarity in structural features between Neu5Ac and compound **106** (Fig. 19), which showed $K_i = 0.33$ mM for TcTS, measured by hydrolysis of MU-Neu5Ac [63].

Fig. 19 Non-carbohydrate structures as potential TcTS inhibitors [63–66]



An approach based on molecular modeling and screening resulted in a library of structures that resemble Neu5Ac only because they contain a carboxylate function [64]. For compound **107** $IC_{50} = 0.15$ mM was determined (assay: hydrolysis of MU-Neu5Ac).

Recently Withers et al. tested chalcones carrying sulfonamide groups as potential inhibitors of TcTS [65]. Impressive IC_{50} data could be measured – e.g., $IC_{50} = 0.6$ μ M for compound **108**; however, with these structural elements it could not be excluded that covalent modifications of the enzyme may occur due to the Michael acceptor properties of these derivatives. The authors did address this question, and a covalent enzyme modification was excluded since the compounds did not show a time-dependent inhibitor activity [65].

Furthermore, some flavones and anthraquinones were studied. A quite significant inhibition activity comparable to that of **108** was measured with the chlorinated derivative **109**, that is $IC_{50} = 0.58$ μ M (MU-Neu5Ac hydrolysis assay) [66]. Again, principally arguments similar to the case of chalcones may be cited to influence these data (Fig. 19).

It should be noted that in all the above cases quite different assays were employed. These lead to different data such as K_i , K_M , and IC_{50} , which in turn may represent different effects and are consequently difficult to compare.

References

1. Morel CM, Lazdins J (2003) *Nat Rev Microbiol* 1:14
2. Chagas C (1909) *Mem Inst Oswaldo Cruz* 1:159
3. Marshall E (1995) *Science* 267:811
4. Teixeira ARL, Nitz N, Guimaro MC, Gomes C, Santos-Buch CA (2006) *Postgrad Med J* 82:788
5. Schauer R, Reuter G, Muehlpfordt H, Andrade AFB, Pereira MEA (1983) *Hoppe-Seyler's Z Physiol Chem* 364:1053
6. Barrett MP, Burchmore RJS, Stich A, Lazzari JO, Frascch AC, Cazzulo JJ, Krishna S (2003) *Lancet* 362:1469
7. Alves MJM, Colli W (2007) *IUBMB Life* 59:274

8. Schenkman S, Jiang MS, Hart GW, Nussenzweig V (1991) *Cell* 65:1117
9. Rubin-de-Celis SSC, Uemura H, Yoshida N, Schenkman S (2006) *Cell Microbiol* 8:1888
10. Leguizamon MS, Mocetti E, Rivello HG, Argibay P, Campetella O (1999) *J Infect Dis* 180:1398
11. Erdmann H, Steeg C, Koch-Nolte F, Fleischer B, Jacobs T (2009) *Cell Microbiol* 11:1600
12. Pereira ME, Zhang K, Gong Y, Herrera EM, Ming M (1996) *Infect Immun* 64:3884
13. Amaya MF, Buschiazzo A, Nguyen T, Alzari PM (2003) *J Mol Biol* 325:773
14. Buschiazzo A, Campetella O, Frasch ACC (1997) *Glycobiology* 7:1167
15. Montagna G, Cremona ML, Paris G, Amaya MF, Buschiazzo A, Alzari PM, Frasch ACC (2002) *Eur J Biochem* 269:2941
16. Koliwer-Brandl H, Gbem TT, Waespy M, Reichert O, Mandel P, Drebitz E, Dietz F, Kelm S (2011) *BMC Biochem* 12:39
17. Engstler M, Reuter G, Schauer R (1993) *Mol Biochem Parasitol* 61:1
18. Nagamune K, Acosta-Serrano A, Uemura H, Brun R, Kunz-Renggli C, Maeda Y, Ferguson MAJ, Kinoshita T (2004) *J Exp Med* 199:1445
19. Aguero F, Campo V, Cremona L, Jager A, Di Noia JM, Overath P, Sanchez DO, Frasch AC (2002) *Infect Immun* 70:7140
20. Pereira ME (1983) *Science* 219:1444
21. Previato JO, Andrade AFB, Pessolani MCV, Mendonca-Previato L (1985) *Mol Biochem Parasitol* 16:85
22. Uemura H, Schenkman S, Nussenzweig V, Eichinger D (1992) *EMBO J* 11:3837
23. Scudder P, Doom JP, Chuenkova M, Manger ID, Pereira MEA (1993) *J Biol Chem* 268:9886
24. Schenkman S, Chaves LB, Pontes de Carvalho LC, Eichinger D (1994) *J Biol Chem* 269:7970
25. Buschiazzo A, Amaya MF, Cremona ML, Frasch AC, Alzari PM (2002) *Mol Cell* 10:757
26. Amaya MF, Watts AG, Damager I, Wehenkel A, Nguyen T, Buschiazzo A, Paris G, Frasch AC, Withers SG, Alzari PM (2004) *Structure* 12:775
27. Cremona ML, Sanchez DO, Frasch ACC, Campetella O (1995) *Gene* 160:123
28. Watts AG, Damager I, Amaya ML, Buschiazzo A, Alzari P, Frasch AC, Withers SG (2003) *J Am Chem Soc* 125:7532
29. Damager I, Buchini S, Amaya MF, Buschiazzo A, Alzari P, Frasch AC, Watts A, Withers SG (2008) *Biochemistry* 47:3507
30. Todeschini AR, Mendonça-Previato L, Previato JO, Varki A, van Halbeek H (2000) *Glycobiology* 10:213
31. Ribeiro M, Pereira-Chioccola VL, Eichinger D, Rodrigues MM, Schenkman S (1997) *Glycobiology* 7:1237
32. Buschiazzo A, Tavares GA, Campetella O, Spinelli S, Cremona ML, Paris G, Amaya MF, Frasch ACC, Alzari PM (2000) *EMBO J* 19:16
33. Paris G, Ratier L, Amaya MF, Nguyen T, Alzari PM, Frasch ACC (2005) *J Mol Biol* 345:923
34. Lee S-G, Shin D-H, Kim B-G (2002) *Enzyme Microb Technol* 31:742
35. Neubacher B, Schmidt D, Ziegelmüller P, Thiem J (2005) *Org Biomol Chem* 3:1551
36. Neubacher B (2005) Chemoenzymatic syntheses of complex oligosaccharides employing trans-sialidase from *Trypanosoma cruzi*. PhD thesis, University of Hamburg
37. Ito Y, Paulson JC (1993) *J Am Chem Soc* 115:7862
38. Schmidt D, Sauerbrei B, Thiem J (2000) *J Org Chem* 65:8518
39. Kroeger L, Thiem J (2005) *J Carbohydr Chem* 24:717
40. Scheppokat AM, Gerber A, Schroven A, Meinke S, Kopitzki S, Beketow E, Thimm J, Thiem J (2010) *Eur J Cell Biol* 89:39
41. Kroeger L, Scudlo A, Thiem J (2006) *Adv Synth Catal* 348:1217
42. Harrison JA, Kartha KPR, Fournier EJJ, Lowary TL, Malet C, Nilsson UJ, Hindsgaul O, Schenkman S, Naismith JH, Field RA (2011) *Org Biomol Chem* 9:1653
43. Campo VL, Carvalho I, Da Silva CHTP, Schenkman S, Hill L, Nepogodiev SA, Field RA (2010) *Chem Sci* 1:507
44. Neubacher B, Scheid S, Kelm S, Frasch AC, Meyer B, Thiem J (2006) *Chembiochem* 7:896

45. Cornforth JW, Firth ME, Gottschalk A (1958) *Biochem J* 68:57
46. Nakamura M, Furuhashi K, Ogura H (1988) *Chem Pharm Bull* 36:4807
47. Schrovén A, Meinke S, Ziegelmeier P, Thiem J (2007) *Chem Eur J* 13:9012
48. Schrovén A (2007) Synthesis and investigations on modified neuraminic acid donors for regio- and stereoselective sialylation with trans sialidase (*Trypanosoma cruzi*). PhD thesis, University of Hamburg
49. Lauria-Pires L, Braga MS, Vexenat AC, Nitz N, Simoes-Barbosa A, Tinoco DL, Teixeira ARL (2000) *Am J Trop Med Hyg* 63:111
50. Schofield CJ, Jannin J, Salvatella R (2006) *Trends Parasitol* 22:583
51. Neres J, Bryce RA, Douglas KT (2008) *Drug Discov Today* 13:110
52. Buchini S, Buschiazzo A, Withers SG (2008) *Angew Chem Int Ed* 47:2700
53. Carvalho ST, Sola-Penna M, Oliveira IA, Pita S, Goncalves AS, Neves BC, Sousa FR, Freire-de-Lima L, Kurogochi M, Hinou H, Nishimura S-i, Mendonca-Previato L, Previato JO, Todeschini AR (2010) *Glycobiology* 20:1034
54. Agusti R, Paris G, Ratier L, Frasch ACC, de Lederkremer RM (2004) *Glycobiology* 14:659
55. Carvalho I, Andrade P, Campo VL, Guedes PMM, Sesti-Costa R, Silva JS, Schenkman S, Dedola S, Hill L, Rejzek M, Nepogodiev SA, Field RA (2010) *Bioorg Med Chem* 18:2412
56. Giorgi ME, Ratier L, Agusti R, Frasch ACC, Lederkremer RM (2010) *Glycoconjugate J* 27:549
57. Meindl P, Bodo G, Palese P, Schulman J, Tuppy H (1974) *Virology* 58:457
58. Streicher H, Busse H (2006) *Bioorg Med Chem* 14:1047
59. Busse H, Hakoda M, Stanley M, Streicher H (2007) *J Carbohydr Chem* 26:159
60. Meinke S, Thiem J (2008) *Carbohydr Res* 343:1824
61. Meinke S, Schrovén A, Thiem J (2011) *Org Biomol Chem* 9:4487
62. Engstler M, Ferrero-Garcia MA, Parodi AJ, Schauer R, Storz-Eckerlin T, Vasella A, Witzig C, Zhu X (1994) *Helv Chim Acta* 77:1166
63. Neres J, Bonnet P, Edwards PN, Kotian PL, Buschiazzo A, Alzari PM, Bryce RA, Douglas KT (2007) *Bioorg Med Chem* 15:2106
64. Neres J, Brewer ML, Ratier L, Botti H, Buschiazzo A, Edwards PN, Mortenson PN, Charlton MH, Alzari PM, Frasch AC, Bryce RA, Douglas KT (2009) *Bioorg Med Chem Lett* 19:589
65. Kim JH, Ryu HW, Shim JH, Park KH, Withers SG (2009) *Chembiochem* 10:2475
66. Arioka S, Sakagami M, Uematsu R, Yamaguchi H, Togame H, Takemoto H, Hinou H, Nishimura S-I (2010) *Bioorg Med Chem* 18:1633

Index

A

O-Acetylated sialic acid, 75, 77
N-Acetyl-D-neuraminic acid, 2, 6, 32, 76
Acetylesterase, 10
Adeno-associated viruses (AAVs), 17
Adenoviridae, 17
Alphacoronaviruses, 11
Anti-adhesion, 137
Antibody-dependent cellular cytotoxicity (ADCC), 202
Anti-polySia antibodies, 87
Arrays, 125
Arthrobacter ureafaciens, 35
Asialoorosomuroid, 209
Asparagine-linked glycan (*N*-glycan), 201
Avian influenza, 6
Avulavirus, 13
Azaelectrocyclization, 201, 203

B

Bacteria, 29
Bacteriophages, 29, 31
Benzimidazole, 246
Betacoronaviruses, 9
Binding kinetics, 151
Bio-affinity, 105
Biomarkers, 105
Blood–brain barrier, 33
Bovine enterovirus 261, 14

C

Caliciviruses, 14
Carbohydrate epitopes, 151

Caudovirales, 31
CD22, 217
Cell adhesion, 29
Cell surface engineering, 201
Chagas disease, 246
Chaperones, 49
Chemoenzymatic synthesis, 231
p-Chlorobenzamide, 181
Clostridium perfringens, 40
CMP-Neu5Ac hydroxylase (CMAH), 78
Complement-dependent cytotoxicity (CDC), 202
Coronaviruses, 9

D

Deaminoneuraminic acid, 32, 75, 78, 126
Dendrimers, 210
Densovirus, 17
Diisopropyl fluorophosphate, 10
Disialic acid, 75, 79, 82
DMB derivatization, 75
Donor and acceptor substrates/inhibitors, 231
DOTA (tetraazacyclodecane tetraacetic acid), 203
Drug discovery, 151
DTPA (diethylenetriamine pentaacetic acid), 203

E

Electrophorus electricus, 58
Endosialidases, 29, 31, 33, 35
active site, 42
catalysis, 45
F, crystal structure, 39
processivity, 52

Endosialidases (*cont.*)

specificity, 35

Epidemic keratoconjunctivitis (EKC), 17

Erythropoietin (EPO), 202

Escherichia coli K1, 32

Esterase, 8

Exosialidases, 46

F

Fluorescence, 201

labeling, living cells, 219

3-Fluoro sialosyl fluorides, 245

Fragment-based drug discovery (FBDD), 190

Functional glycomics, 105

G

Galactopyranosyl triazole, 246

Gammacoronaviruses, 12

Gangliosides, 1, 136, 151

Glycans, 8, 34, 59, 86, 105, 119, 154, 165, 201

asparagine-linked *N*-glycan, 201*N*-glycans, 14, 76, 91, 115, 127*O*-glycans, 8, 76, 91, 108, 115, 127

profiling, 105

Glycobioenvironment, 210, 224

Glycoclusters, 201, 210

polylysine-based dendrimer-type, 210

N-Glycolylneuraminic acid (Neu5Gc),

75, 78

Glycomics, 105

Glycomimetics, 151

Glycoproteins, 1, 75, 201

H

Hemagglutinating encephalomyelitis virus (HEF), 10

Hemagglutination inhibitors, 2

Hemagglutinin, 3, 127, 139–141

Hemagglutinin esterase, 93

fusion (HEF) protein, 7

Hemagglutinin-neuraminidase

(HN) glycoprotein, 13

Hemicentrotus pulcherrimus, sperm lysate, 87

H1N1, 5

H5N1, 4

Huisgen 1,3-cycloaddition, 212, 218, 221

Human parainfluenza viruses (HPIV), 13

I

Immunochemical analysis, 75

Immunodetection, 86

Immunoglobulin (IgG), 202

Immunohistochemistry, 75, 90

Infectious bronchitis virus (IBV), 9

Infectious salmon anemia virus (ISAV), 2, 9

Influenza, 2, 3, 139

C, 7

hemagglutinin esterase (InfCHE), 93

neuraminidases, 60

Inhibitor design, 231

Insulin, 202, 209

Isavirus, 2, 9

J

Junctional adhesion molecule 1 (JAM-1), 15

K

Ketodeoxy nonulosonic acid (KDN), 162, 240

Ketononose, 32

L

Labeling, 201

Lead optimization, 151

Lectins, 4, 106, 110, 151, 187, 202

glycomics, 106

microarray, 105

Lymphocytes, 201

Lymphotropic papovavirus (LPV), 17

Lysine, 201

azaelectrocyclization-based labeling, 203

M*Maackia amurensis*, isolectins, 91*Mannheimia (Pasteurella) haemolytica* A2, 33

Mengovirus, 14

Merkel cell polyomavirus (MCPyV), 16

Methyl sialyllactoside, 235

4-Methylumbelliferyl α -sialoside, 235

Modifications, 231

Moraxella nonliquefaciens, 32

Mucins, 1, 11, 92, 108, 137

Mumps virus, 13

Myelin-associated glycoproteins (MAG), 151

Myoviridae, 31, 33

N

- Nanoparticles, 125
- Nanotechnology, 142
- Neisseria meningitidis*, 32, 36, 79, 139
- Neural cell adhesion molecule (NCAM), 30, 56, 79
- Neuraminic acid (5-amino-3,5-dideoxy-D-glycero-D-galacto-non-2-ulosonic acid), 32
- Neuraminidase, 1, 2, 6
- Neuropilin-2 (NRP-2), 59
- Newcastle disease virus (NDV), 13
- Nifurtimox, 246
- 5-Nitroindole, 190
- Norovirus, 14

O

- Oligosaccharide, 3, 84, 110, 115
 - mimics, 151, 156
- Oligosialic acid, 75, 79
 - binding sites, 48
- Opsonophagocytosis, 33
- Orosomucoid, 209
- Orthomyxoviridae, 2
- Oseltamivir (Tamiflu), 246

P

- Panstrongylus megistus*, 246
- Paramyxoviridae, 13
- Parvoviridae, 17
- Picornaviridae, 14
- Podoviridae, 31, 33
- Polyomaviridae, 16
- Polysialic acid, 29, 32, 75
 - detection, 61, 82
- Polysialoglycoprotein (PSGP), 79
- Porcine aminopeptidase N, 11
- Porcine respiratory coronavirus (PRCoV), 11
- Positron emission tomography (PET), 201
- Preparative use, 231

R

- Receptor bindings, 1
- Receptor-destroying enzyme, 2
- Reoviridae, 15
- Respirovirus, 13
- Rhodnius prolixus*, 246
- Rubulavirus, 13

S

- Sambucus nigra* agglutinin (SNA), 91
- Sambucus sieboldiana* lectin (SSA), 86
- Selectins, 128, 138
- Sendai virus, 13
- Sialate-O-acetyl-esterases, 1, 77
- Sialate-O-acetyl-transferases, 77
- Sialic acid, 125
 - receptors, 1
 - O-sulfated, 75, 78
- Sialidase, 247
- Sialoadhesin (Siglec-1), 190, 193
- Sialoglycoconjugates, 75, 80
- Sialo-N-glycans, 203
- Sialosides, 125, 151
 - arrays, 140
 - synthesis, 128
- Sialyllactose, 235, 246
- Sialyl Lewis structures, tumor research, 138
 - neoglycoproteins, 14
- Sialyloligosaccharides, 13
- Sialyltransferases, 125
- Siglecs (sialic acid-binding immunoglobulin-type lectins), 141, 151
 - Siglec-1 (sialoadhesin), 190, 193
 - Siglec-2 (CD22), 217
 - Siglec-4 (MAG) antagonists, 151
- Siphoviridae, 32, 33
- Somatostatin, 206
- Sphingobacterium multivorum*, 86
- Surface plasmon resonance (SPR) based biosensor (Biacore), 181

T

- TAMRA, 219
- Tetrasaccharide mimic, 172
- Theiler's murine encephalomyelitis virus (TMEV), 14
- Thiogalactopyranosides, 245
- Thogotovirus, 2
- Toroviruses, 12
- Transmissible gastroenteritis virus (TGEV), 9, 11
- Trans-sialidases, trypanosomal, 231, 234
- Triatoma infestans*, 246
- Triazole galactopyranosides, 245
- Trifluoromethylumbelliferyl, 46
- Trypanosoma brucei brucei*, 233
- Trypanosoma cruzi*, 231, 232, 244
- Trypanosomiasis, 232

Tumors, invasion, 137
targeting, 201

V

Vibrio cholerae, 2, 40, 236, 247

Viruses, 1

Y

Yersinia enterocolitica adhesin (yadA), 61

Z

Zanamivir, 246