Roles of Sialic Acid Diversity in Mammalian Brain Evolution

Permalink
https://escholarship.org/uc/item/5290t4zj

Author
Davies, Leela Rhian Lobo

Publication Date
2013

Peer reviewed|Thesis/dissertation
Roles of Sialic Acid Diversity in Mammalian Brain Evolution

A dissertation submitted in partial satisfaction of the requirements for the degree
Doctor of Philosophy

in

Biomedical Sciences with Specialization in Anthropogeny

by

Leela Rhian Lobo Davies

Committee in charge:

Professor Ajit Varki, Chair
Professor Jeffrey Esko
Professor Pascal Gagneux
Professor Edward Koo
Professor Alysson Muotri

2013
The Dissertation of Leela Rhian Lobo Davies is approved and it is acceptable in quality and form for publication on microfilm and electronically:

Chair

University of California, San Diego
2013
# TABLE OF CONTENTS

Signature Page..............................................................................................................................................iii
Table of Contents........................................................................................................................................iv
List of Figures...............................................................................................................................................v
List of Tables................................................................................................................................................vi
Acknowledgements..........................................................................................................................................viii
Vita...............................................................................................................................................................x
Abstract of the Dissertation............................................................................................................................xiv

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chapter 1</td>
<td>Why is N-Glycolylneuraminic Acid Rare in the Vertebrate Brain?</td>
<td>1</td>
</tr>
<tr>
<td>Chapter 2</td>
<td>The Roles of Polysialic Acid in Brain Development</td>
<td>33</td>
</tr>
<tr>
<td>Chapter 3</td>
<td>Resistance of ( \alpha_{2-8} )-Linked N-Glycolylneuraminic Acid to Enzymatic Cleavage</td>
<td>60</td>
</tr>
<tr>
<td>Chapter 4</td>
<td>The Overexpression of N-Glycolylneuraminic Acid in the Mammalian Brain</td>
<td>102</td>
</tr>
<tr>
<td>Chapter 5</td>
<td>The Effects of a Human-Specific Change in the Polysialyltransferase ST8Sia-II</td>
<td>124</td>
</tr>
<tr>
<td>Chapter 6</td>
<td>Conclusions and Future Directions</td>
<td>158</td>
</tr>
<tr>
<td>Appendix</td>
<td>Changes in Developmental Expression of Polysialic Acid in Sialidase Deficient Mice</td>
<td>170</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

| Figure 1-1. | Structures of the three common sialic acids in vertebrates | 3 |
| Figure 1-2. | Alignment of vertebrate CMAH sequences | 6 |
| Figure 2-1. | Structure of polySia on NCAM | 40 |
| Figure 2-2. | Known domains of ST8Sia-II | 51 |
| Figure 3-1. | Neuroblastoma cells can incorporate Neu5Gc into endogenous PSA-NCAM | 74 |
| Figure 3-2. | Sialidases exhibit no detectable preference for Neu5Ac or Neu5Gc | 76 |
| Figure 3-3. | Synthetic strategy to generate polymers of N-acetyl, N-glycolyl, or N-propionylneuraminic acid | 78 |
| Figure 3-4. | Polymers of Neu5Ac, Neu5Gc, and Neu5Pr used in these studies | 79 |
| Figure 3-5. | Relative rate of hydrolysis of α2-8-linked Sias by sialidases is markedly reduced when Neu5Gc residue is present at terminal positions | 82 |
| Figure 3-6. | Neu5Gc in polySia increases relative susceptibility to acid-catalyzed breakdown | 84 |
| Figure 3-7. | Molecular dynamics modeling demonstrates conformational similarities and differences between hexamers of Neu5Ac, Neu5Gc, and Neu5Pr | 88 |
| Figure 3-8. | Replacement of the glycolyl hydroxyl group with a methyl group restores sialidase activity | 90 |
| Figure 4-1. | The distribution of Neu5Gc in the mammalian brain | 109 |
| Figure 4-2. | Transgenic Cmah driven by CMV promoter causes embryonic abnormalities | 113 |
| Figure 4-3. | The CFE-CMAH construct used for inducible overexpression of Cmah | 115 |
| Figure 4-4. | CFE-CMAH lines are able to express Cmah and Neu5Gc in the brain | 117 |
| Figure 5-1. | Alignment of peptide sequences of ST8Sia-II | 137 |
| Figure 5-2. | Homology modeling of ST8Sia-II to the crystal structure ST3Gal-I suggests N308K is distant from the catalytic site | 138 |
Figure 5-3. Human and chimp ST8SIA2 produces equal amounts of polySia in N2a transfectants..........................140

Figure 5-4. Human ST8Sia-II produces longer polySia chains..........................142

Figure 5-5. N308K confers instability to human ST8Sia-II..........................145

Figure 5-6. PolySia expression does not significantly differ in human and chimpanzee adult frontal cortex..........................147

Figure 5-7. Evidence for increased expression of ST8SIA2 in human development........148

Figure 5-8. Alignment of human, Neandertal, and Denisovan sequences allow dating of N308K to at least 500 kya..........................150

Figure 7-1. Immunoblots for polySia and NCAM in sialidase deficient mice are similar to wild type..........................174

Figure 7-2. Quantification of polySia expression suggests slight increase of polySia in sialidase deficient mice..........................175
LIST OF TABLES

Table 1-1. Distribution of Neu5Gc in vertebrate tissues...........................................8
Table 2-1. Processes and phenotypes associated with polySia......................................45
Table 3-1. States and populations from 0.5 µs molecular dynamics simulations of Sia hexamers.........................................................................................................................87
Table 4-1. Summary of initial efforts to overexpress transgenic Cmah in mouse brain.....112
ACKNOWLEDGEMENTS

I am so grateful for the help and support of many people who have made this dissertation possible. First and foremost, thank you to my thesis advisor, Ajit Varki, whose unending curiosity and joy in his work are inspiring. Thank you for teaching me both fearlessness and focus, and for always challenging me to be a better scientist. Thank you also to Nissi Varki, for your kindness and wisdom from my very first year of medical school.

Thank you to my committee members, Jeff Esko, Pascal Gagneux, Edward Koo, and Alysson Muotri. Your guidance has been invaluable, particularly through changing projects and priorities.

Further thanks are necessary to Pascal Gagneux for his role as director of the Anthropogeny Specialization Track. I have been so lucky to participate in this program, which has broadened the horizons of my graduate studies and ideas for the future. I am very grateful to all of the faculty and students that have been involved with CARTA.

Many thanks are due to the past and present members of the Varki lab. A lab becomes a family, and this is one that generously shares ideas, commiserations, stories, time, spare hands, food, and great friendship. In particular, I would like to thank those I have worked with closely on projects included here. Sandra Diaz, you seem to have an answer to every question, a spare hand for every experiment, and a word of encouragement for every dark moment. Darius Ghaderi, Oliver Pearce, and Yuko Matsui, you have all been outstanding teachers, in very different ways. Thank you all.

Lastly, thank you to my family, who have made me the person I am. To my parents, Keith and Sandra, whose love and support are unfailing. To my brother, Owen, who believes in me more than anyone. And to my partner, John, who has patiently accompanied me through joys and challenges, every step of the way.

The text of Chapter 1 chapter is, in full, a reprint of material originally published in Topics in Current Chemistry by authors Davies LRL and Varki A, 2013, with kind permission of Springer Science+Business Media. The dissertation author was the primary author of this review article.
The text of Chapter 3, in full, is a reprint of the material as it appears in the article of the same title by Davies LRL, Pearce OMT, Tessier MB, Assar S, Smutova V, Pajunen M, Sumida M, Sato C, Kitajima K, Finne J, Gagneux P, Pshezhetsky A, Woods R, and Varki A in the Journal of Biological Chemistry, 2012. The dissertation author was one of two primary authors of this article.
VITA

EDUCATION

2015 Doctor of Medicine (expected), University of California, San Diego, La Jolla, CA
2013 Doctor of Philosophy, Biomedical Sciences with Specialization in Anthropogeny, University of California, San Diego, La Jolla, CA
2005 Bachelor of Science with Honors, Neuroscience, Brown University, Providence, RI

SCIENTIFIC TRAINING

2009-2013 Graduate Researcher, Laboratory of Ajit Varki, UCSD, La Jolla, CA
2006-2007 Research Technician, Laboratory of David Altshuler, Broad Institute of Harvard and MIT, Cambridge, MA
2002-2006 Undergraduate Researcher, Laboratory of James Padbury, Women and Infants Hospital, Providence, RI

LEADERSHIP EXPERIENCE

2013 Planning Committee, NextGen Seminar, San Diego Glycobiology Symposium, San Diego, CA
2011-2012 Teaching Assistant, Anatomy, UCSD School of Medicine, La Jolla, CA
2008-2010 Clinic Manager, UCSD Student-Run Free Clinics, San Diego, CA
HONORS & AWARDS

2011-2013  Predoctoral CARTA Fellowship in Anthropogeny, University of California, San Diego
2012  Poster Award in Glycobiology, American Society for Biochemistry and Molecular Biology
2010-2011  Predoctoral Neuroplasticity of Aging Training Grant, University of California, San Diego
2009-2011  Sponsored member, American Association for the Advancement of Science,
2005  Sigma Xi membership, Brown University
2004  Undergraduate Howard Hughes Research Fellowship, Brown University
2001  National Merit Scholar

PUBLICATIONS

Davies LRL and Varki A. Why is N-glycolylneuraminic acid rare in the vertebrate brain? (Review)
Topics in Current Chemistry, 2013.


PRESENTATIONS
2012  *Unusual chemical and enzymatic stability of polysialic acid containing N-glycolylneuraminic acid.*
American Society for Biochemistry and Molecular Biology Annual Meeting, San Diego, CA

2012  *From sugars to societies: My experience in the anthropogeny specialization track.*
CARTA Annual Student Symposium, San Diego, CA

2011  *Why is the sialic acid Neu5Gc excluded from all mammalian brains?*
UCSD Biomedical Sciences Program Annual Retreat, San Diego, CA

POSTERS

2013  *Potential toxicity of the sialic acid Neu5Gc to the vertebrate brain.*
American Society for Clinical Investigation and Association of American Physicians Joint Meeting, Chicago, IL

2012  *Unusual chemical and enzymatic stability of polysialic acid containing N-glycolylneuraminic acid.*
American Society for Biochemistry and Molecular Biology Annual Meeting, San Diego, CA

2012  *Unexpected resistance of α2-8 linked N-glycolylneuraminic acid to enzymatic cleavage.*
San Diego Glycobiology Symposium, San Diego, CA

2011  *The rarity of N-glycolylneuraminic acid in the vertebrate brain.*
Medical Scientist Training Program Annual Revisit, La Jolla, CA

2006  *A genome-wide scan for variation in the polyadenylation signal and correlation to expression levels.*
American Society of Human Genetics Annual Meeting, New Orleans, LA
ABSTRACT OF THE DISSERTATION

Roles of Sialic Acid Diversity in Mammalian Brain Evolution

by

Leela Rhian Lobo Davies

Doctor of Philosophy in Biomedical Sciences with Specialization in Anthropogeny

University of California, San Diego, 2013

Professor Ajit Varki, Chair

The sialic acids N-acetylneuraminic acid (Neu5Ac) and N-glycolylneuraminic acid (Neu5Gc) are commonly found in the glycolipids and glycoproteins that cover all vertebrate cell surfaces. The brain, which has a higher concentration of sialic acids than any other tissue, exhibits some interesting features of sialic acid biology with evolutionary implications. In the brain of all vertebrate species tested to date, Neu5Ac is present at high quantities, whereas Neu5Gc is very low or even absent. This conserved exclusion suggests a selective pressure against Neu5Gc expression in the vertebrate brain. We have hypothesized that the mechanism for this pressure may be that Neu5Gc presence inhibits the appropriate enzymatic degradation of polysialic acid (polySia), a polymer of alpha2-8 linked sialic acid. PolySia is primarily found in
brain tissue, and has been implicated in a wide array of important developmental processes including migration, neurite outgrowth, plasticity, and repair. In Chapter 3, we test the viability of this hypothesis in vitro. We demonstrate that Neu5Gc present in polySia does in fact exhibit a relative resistance to degradation by vertebrate and bacterial sialidases. To determine whether this process is meaningful in vivo, and to further understand the presumed detrimental effects of Neu5Gc on the vertebrate brain, we have developed transgenic mouse models that overexpress Neu5Gc in brain tissue. In Chapter 4, we describe these ongoing efforts. Studies in the near future will determine whether these mice exhibit problems with polySia degradation, and comprehensively characterize their behavioral and anatomical phenotypes. Lastly, in Chapter 5, we explore the potential role of polySia in human brain evolution. We focus on the effects of a single human-specific mutation in one of the biosynthetic enzymes of polySia, ST8Sia-II. Preliminary work suggests that this mutation decreases the stability of the enzyme while increasing the length of its synthesized polySia product. These effects may have implications for the regulation of polySia, and its associated processes, in the human brain. Together, the work discussed here examines potential roles that sialic acid metabolism may have played in mammalian brain evolution.
CHAPTER 1

Why is N-Glycolylneuraminic Acid Rare in the Vertebrate Brain?
The sialic acids \(N\text{-acetylneuraminic acid (Neu5Ac)}\) and \(N\text{-glycolyneuraminic acid (Neu5Gc)}\) differ by a single oxygen atom and are widely found at the terminal position of glycans on vertebrate cell surfaces. In animals capable of synthesizing Neu5Gc, most tissues and cell types express both sialic acids, in proportions that vary between species. However, it has long been noted that Neu5Gc is consistently expressed at trace to absent levels in the brains of all vertebrates studied to date. Although several reports have claimed to find low levels of Neu5Gc-containing glycans in neural tissue, no study definitively excludes the possibility of contamination with glycans from non-neural cell types. This distribution of a molecule—prominently but variably expressed in extraneural tissues but very low or absent in the brain—is, to our knowledge, unique. The evolutionarily conserved brain-specific suppression of Neu5Gc may indicate that its presence is toxic to this organ; however, no studies to date have directly addressed this very interesting question. Here we provide a historical background to this issue and discuss potential mechanisms causing the suppression of Neu5Gc expression in brain tissue, as well as mechanisms by which Neu5Gc may exert the presumed toxicity. Finally, we discuss future approaches towards understanding the mechanisms and implications of this unusual finding.

The Evolutionary Origins of Sialic Acids

Sialic acids (Sias) commonly occupy the terminal position on the glycoconjugates that cover all vertebrate cell surfaces, and as such are major determinants of the molecular cell surface phenotype \(^1\). The term Sia refers to derivatives of \(N\text{-acetylneuraminic acid (Neu5Ac)}\), \(N\text{-glycolyneuraminic acid (Neu5Gc)}\), and ketodeoxynonulosonic acid (Kdn) (Figure 1-1). In 1963, Leonard Warren used the newly developed thiobarbituric acid test to screen for Sias in a wide variety of animals, algae, plants, and fungi \(^2\), detecting them almost throughout the deuterostome lineage of animals, a group that includes vertebrates and so-called “higher” invertebrates. Sias were found to be present in vertebrates, echinoderms, hemichordates, and cephalochordates, although they were not found in urochordates. Outside the deuterostomes, Warren also noted
Figure 1-1. Structures of the three common sialic acids in vertebrates.

$N$-acetylmuramic acid (Neu5Ac)

$N$-glycolylmuramic acid (Neu5Gc)

Ketodeoxynonulosonic acid (Kdn)
small amounts of Sias within two species of flatworm and the digestive glands of squid and lobster, but considered these likely to be of dietary origin rather than endogenously synthesized. Based on these and later data, Sias were once considered both unique and universal to the deuterostome lineage.

In keeping with this concept, neither Sias nor the genes encoding their biosynthesis can be found in the model protostome C. elegans. However, although early studies failed to identify any Sia-containing structures in invertebrates, more recent analyses identified small amounts of Neu5Ac in certain protostomes, including in gangliosides of squid and octopus, as well as in the slug Arion lusitanicus. Furthermore the Drosophila genome encodes a sialyltransferase that functions in one phase of neural development. Thus it seems likely that Sias originated before the deuterostome:protostome split, became prominent only in deuterostomes, and were mostly discarded in protostomes.

Outside the animal kingdom, there are multiple instances of Neu5Ac production in various species of bacteria. However, the synthetic pathways for Neu5Ac in these species have evolved independently from those of animals, likely as a means of host immune evasion. These bacterial pathways appear to have arisen by convergent evolution, taking advantage of the more ancient pathways for biosynthesis of nonulosonic acids (NulOs), which are likely also the evolutionary precursors of the animal sialic acids.

Neu5Gc is a Major Sialic Acid in Many Mammals

Although Kdn can be prominent in some fish, the two major Sias of most mammalian species are Neu5Ac and Neu5Gc. The CMP-Neu5Ac nucleotide sugar donor molecule originates from UDP-GlcNAc, via ManNAc and free Neu5Ac intermediates. CMP-Neu5Ac can be then be converted to CMP-Neu5Gc by cytidine monophosphate N-acetylneuraminic acid hydroxylase (CMAH). This activity was first identified by Schauer more than 40 years ago, then shown to work specifically on CMP-Neu5Ac, and the enzyme was later cloned and further characterized. Gene knockout studies in mice have indicated that CMAH is likely to be
solely responsible for Neu5Gc synthesis. The polypeptide sequence of CMAH is highly conserved across deuterostomes, but appears to be essentially confined to this lineage (Figure 1-2).

Despite high sequence conservation of vertebrate CMAH, there are some instances of its loss. The human genome is unique among the old world primates in containing a universal deletion spanning exon 6 in the CMAH gene $^{20,21}$. While one group suggested that this causes an N-terminal truncation of the polypeptide $^{20}$, the other showed that an Alu-mediated mutation introduces a premature stop codon and highly truncated polypeptide (see Figure 1-2), consequently preventing endogenous production of Neu5Gc $^{21,22}$. However, very low levels of Neu5Gc can be found in human tissues, likely due to metabolic incorporation from the diet $^{23}$. Meanwhile, a recent study has confirmed the previously suspected absence of Neu5Gc in sauropsids—i.e., birds and reptiles—as well as likely in monotremes $^{24}$. In keeping with this, no sequence strongly homologous to vertebrate CMAH can be found in any of the sauropsid genomes so far sequenced. Although some Neu5Gc was noted in one lizard and one bird—the green basilisk and the budgerigar—this was also suspected to originate from food $^{24}$. Some strains of dogs and cats are deficient in erythrocyte Neu5Gc $^{25,26}$; it is as yet unclear whether this is caused by genomic CMAH mutations. Changes in the CMAH promoter, however, have been identified in some cats that appear to define the feline type B blood group that expresses solely Neu5Ac on erythrocytes $^{27}$.

There are many consequences to the loss of Neu5Gc production in a mammal. A mouse model of the human CMAH mutation, for example, exhibits numerous phenotypes, including delayed wound healing and age-dependent hearing loss $^{19}$, heightened B cell responses $^{18,28}$, and a tendency for decreased insulin production $^{29}$. Like this mouse model, vertebrate lineages that have lost functional CMAH are viable and fertile. However, Neu5Gc loss can cause relative infertility with wild-type animals, because of female anti-Neu5Gc antibodies that attack Neu5Gc-positive sperm or embryos $^{30}$. 
Figure 1-2. Alignment of vertebrate CMAH sequences. Alignment was performed using CLUSTALW.
Rarity of Neu5Gc in the Vertebrate Brain

Despite a universally high concentration of Neu5Ac in the brain, Neu5Gc has long been noted to be rare in the central nervous system (CNS) \(^3,32\). Although is well known in the field that Neu5Gc is rare in the vertebrate brain, we were unable to ascertain when and where this observation was first recorded in writing.

Unlike the highly variable levels of Neu5Gc found in all other tissues, the suppression of brain Neu5Gc expression is remarkably conserved across all vertebrates that have been studied to date (see Table 1-1, reproduced from \(^33\)). The highest published fraction of Neu5Gc in brain tissue is 10% in a sample of adult bovine neocortex, and 5% in calf \(^34\). However, other published studies of bovine brain report much lower fractions of under 2\% \(^32,35,36\). Notably, even species that otherwise have a substantial fraction of their total Sias as Neu5Gc in many non-neural tissues maintain this suppression of Neu5Gc expression within the brain. To our knowledge, no other molecule exhibits such an unusual tissue distribution. The evolutionary conservation of strong CNS suppression seems to indicate that maintaining very low levels of Neu5Gc in the CNS is very important.

Widely Variable Expression of Neu5Gc in Non-Neural Tissues of CMAH-Positive Mammals

The conserved regulation of Neu5Gc expression in a given tissue across species is found only in the brain. As shown in Table 1-1, the Neu5Gc fraction of total sialic acid otherwise varies widely both across tissues and across species. To take the horse as an example, the fraction of total sialic acid that is Neu5Gc is as low as 1-2% in submaxillary mucin \(^32\) but almost 100% on erythrocytes \(^37,38\). This remarkable variability can also be found in interspecies comparisons within a single tissue. For example, in pigs, 90% of sialic acid in submaxillary mucin is Neu5Gc \(^39\), while in cows, it is only 15% \(^32\); in sheep and horses, it is lower still at only 1-3\% \(^32\). Meanwhile, the fraction of Neu5Gc on erythrocytes from these species shows no correlation to that found on the mucins. Thus Neu5Gc expression appears to be extremely dynamic in its regulation and evolving rapidly, in all tissues other than the brain.
Table 1-1. Distribution of Neu5Gc in vertebrate tissues. The Neu5Gc fraction of total Sias in tissues was compared across vertebrates. This Table combines data from the literature with that obtained from samples studied in our laboratory. Neu5Ac and Neu5Gc fractions of samples in our lab were determined by total acid hydrolysis of tissue lysate followed by DMB-HPLC. Conserved suppression of Neu5Gc in the brain is unusual among vertebrate tissues. * indicates data from our laboratory, ** indicates published data confirmed in our laboratory. ++: major fraction; +: minor fraction; –: absent; trace: present at 0.8-3%; nr=not reported.

<table>
<thead>
<tr>
<th>species</th>
<th>serum</th>
<th>RBC</th>
<th>submaxillary gland</th>
<th>liver</th>
<th>kidney</th>
<th>milk</th>
<th>brain</th>
</tr>
</thead>
<tbody>
<tr>
<td>human</td>
<td>-</td>
<td>-</td>
<td>nr</td>
<td>nr</td>
<td>nr</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>chimpanzee</td>
<td>nr</td>
<td>++</td>
<td>nr</td>
<td>+</td>
<td>+</td>
<td>+*</td>
<td>trace**</td>
</tr>
<tr>
<td>macaque</td>
<td>+</td>
<td>+</td>
<td>nr</td>
<td>nr</td>
<td>nr</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>mouse</td>
<td>+</td>
<td>+**</td>
<td>-</td>
<td>++</td>
<td>nr</td>
<td>nr</td>
<td>trace**</td>
</tr>
<tr>
<td>rat</td>
<td>+</td>
<td>+**</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>nr</td>
<td>trace**</td>
</tr>
<tr>
<td>rabbit</td>
<td>trace</td>
<td>+</td>
<td>nr</td>
<td>-</td>
<td>+</td>
<td>nr</td>
<td>-</td>
</tr>
<tr>
<td>pig</td>
<td>nr</td>
<td>++</td>
<td>++</td>
<td>+</td>
<td>+</td>
<td>nr</td>
<td>trace**</td>
</tr>
<tr>
<td>cow</td>
<td>++</td>
<td>++</td>
<td>+</td>
<td>nr</td>
<td>++</td>
<td>trace</td>
<td>trace**</td>
</tr>
<tr>
<td>sheep</td>
<td>+</td>
<td>+</td>
<td>trace</td>
<td>+</td>
<td>+</td>
<td>++</td>
<td>trace</td>
</tr>
<tr>
<td>elephant Afr</td>
<td>nr</td>
<td>nr</td>
<td>nr</td>
<td>++*</td>
<td>nr</td>
<td>+</td>
<td>nr</td>
</tr>
<tr>
<td>elephant Asian</td>
<td>nr</td>
<td>nr</td>
<td>nr</td>
<td>++*</td>
<td>nr</td>
<td>-</td>
<td>nr</td>
</tr>
<tr>
<td>dolphin</td>
<td>nr</td>
<td>nr</td>
<td>nr</td>
<td>++*</td>
<td>++</td>
<td>+*</td>
<td>trace*</td>
</tr>
<tr>
<td>horse</td>
<td>+</td>
<td>++</td>
<td>trace</td>
<td>-</td>
<td>+</td>
<td>nr</td>
<td>trace</td>
</tr>
<tr>
<td>chicken</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>xenopus</td>
<td>nr</td>
<td>nr</td>
<td>nr</td>
<td>nr</td>
<td>nr</td>
<td>nr</td>
<td>-</td>
</tr>
</tbody>
</table>
Our understanding of the regulatory mechanisms establishing the fraction of Neu5Gc in a given cell type is incomplete. One study of the developmental regulation of Neu5Gc in rat tissues found that CMAH enzymatic activity is an imperfect correlate to the level of Neu5Gc, suggesting that other mechanisms may influence the presence of Neu5Gc in sialylated glycans. A more recent study of porcine tissues found that CMAH enzyme activity correlated reasonably well with Neu5Gc levels; however, in lung and heart tissues, a much larger amount of immunoreactive protein was present, suggesting that modification of CMAH may influence its activity. Nevertheless, although other factors may be at play, CMAH expression appears to be the major determinant of Neu5Gc levels.

The factors that determine the optimal ratio of Neu5Gc to Neu5Ac for a given tissue in a given animal are not entirely clear; however, it is easy to imagine the types of evolutionary selection pressures that may have had an influence. Sialic acid-binding lectins, whether pathogenic or endogenous, typically prefer either Neu5Ac or Neu5Gc, and these may affect the balance of Neu5Ac and Neu5Gc presence. Indeed, the Neu5Ac binding preference of ancestral Plasmodium species is postulated to be the impetus for fixation of the CMAH deletion in humans. Siglec-2 (CD22), an endogenous immune modulator found on B cells, preferentially binds Neu5Gc in mice, but will bind either Neu5Ac or Neu5Gc in humans and great apes. Recent work on a mouse model of muscular dystrophy shows that the absence of Neu5Gc is required to produce a severe, human-like phenotype, indicating that the presence of Neu5Gc may be important in normal muscle physiology. Thus both pathogenic and endogenous selective pressures may influence the balance of Neu5Gc and Neu5Ac that is ultimately expressed within a given tissue. The data make it clear, however, that most tissues do not have an intrinsic requirement for a specific proportion of Neu5Gc in total sialic acids, let alone an exclusive preference for Neu5Ac or Neu5Gc.

Sialic Acids in the Vertebrate Brain
In seeking to understand why the vertebrate CNS suppresses the presence of Neu5Gc so consistently, it is important to consider the existing knowledge of CNS sialic acid biology. The brain contains more Sias than any other tissue $^{3,32,48}$. In fact, it was from brain tissue that Ernst Klenk extracted what he then called “neuraminic acid” in 1939, only a few years after the first isolation of sialic acids from salivary gland $^{32}$. Interestingly, Neu5Ac content has been found to be approximately 30% higher in the left hemisphere than in the right hemisphere of a chimpanzee, suggesting that brain Sia concentration may correlate with neurological function or hemispheric dominance $^{49}$. It has also been postulated that sialic acid from dietary sources can increase the Sia concentration in the developing brain, playing an important role in learning and memory $^{50,51}$. Conversely, however, other studies of the incorporation of maternal dietary Neu5Gc into pups in utero have not found uptake of Neu5Gc into brain tissue $^{52}$, suggesting that dietary Neu5Ac and Neu5Gc may differ in their ability to incorporate into brain glycans. It is clear, however, that Sias are extremely important within the brain.

As is the case elsewhere in the body, Neu5Ac in the brain can be found on N- and O-linked glycoproteins and on glycolipids; however, the distribution is somewhat different. It has been estimated that about 65% of brain sialic acid is on gangliosides, 32% on glycoproteins, and only 3% remains unbound $^{53}$. This preponderance of ganglioside-bound, rather than glycoprotein-bound, sialic acid is unique among all tissues $^{50}$. The brain also contains several characteristic sialoglycoconjugates, which we will briefly consider here.

**Polysialic acid**

Polysialic acid (polySia) is an unusual posttranslational modification found on a few mammalian proteins as well as in the jelly coat of certain fish eggs, the voltage-gated sodium channel of eel, and the capsules of certain pathogenic bacteria. The most well studied carrier of polySia in the brain is the neural cell adhesion molecule (NCAM). Brain polySia is composed of long $\alpha 2\text{-}8$-linked polymers of Neu5Ac synthesized by two polysialyltransferases, ST8Sia-II and ST8Sia-IV, within the Golgi apparatus $^{54-56}$. 
There is significant developmental regulation of the expression of polysialylated NCAM (also called embryonic NCAM or PSA-NCAM); in mice, very high levels occur in early postnatal life, mediated primarily by ST8Sia-II, that drop to low adult levels by about 3 weeks of age 57. Polysialylated NCAM plays a wide range of crucial roles throughout development affecting migration of neural progenitors, neurite outgrowth, and formation of appropriate synapses 58. These roles are mediated through two mechanisms. Through the formation of a highly hydrated anionic cloud, polySia attenuates the adhesive properties of NCAM 59 and other adhesion molecules. PolySia also mediates other effects, specifically sensitizing cells to brain derived neurotrophic factor (BDNF) and fibroblast growth factor 2 (FGF-2) through direct interaction 60-63. Later in development, downregulation of polySia has been found to be necessary for myelination of axons 64,65. In adult mammals, polySia is expressed at much lower levels throughout the brain, but can still be found in areas such as the hippocampus that are associated with adult plasticity 66.

Although polySia-NCAM is the most well characterized polysialylated glycoprotein, polySia has been identified on other vertebrate carrier proteins. Within the brain, it has been found on the synaptic cell adhesion molecule (SynCAM) on a specialized subset of glial cells, where it is thought to modulate synaptic formation 67, and has been identified in rat brain on the voltage-gated sodium channel 68. Although one study claimed to find polySia on neural podocalyxin, this was actually an erroneous use of the term to describe a highly sialylated protein 69.

**Oligosialic acid**

A related but less well understood structure is oligosialic acid, *i.e.* two to four residues of α2-8-linked Neu5Ac. Oligosialic acids are enriched in the brains of both embryonic and adult pigs 70. A specific trisialic acid epitope found on mouse brain glycoproteins appears, like polySia, to be developmentally regulated 71. However, the functions of oligosialic acids remain relatively unknown. Notably, the oligosialic acid epitope is recognized by Siglec-11, a protein altered by gene conversion in humans, that is found on tissue macrophages and shows human-specific
expression exclusively in microglia in the brain \textsuperscript{72-74}. Although its function is not well understood, Siglec-11 was recently found to decrease the transcription of inflammatory mediators and thus reduce neurotoxicity induced by bacterial lipopolysaccharide (LPS) \textsuperscript{75}.

**Gangliosides**

Gangliosides are sialic acid-containing glycosphingolipids expressed throughout the body but most highly in brain tissue. In mammals, the simple gangliosides GD3 and GM3 predominate early in development throughout formation of the neural tube and neural stem cell proliferation, but this profile changes by the major period of neurogenesis \textsuperscript{76}. In adult mammals, GM1, GD1a, GD1b, and GT1b together comprise over 85\% of total brain gangliosides in pigs and 95\% in humans \textsuperscript{77}. These were found to be most prevalent in the neuropil surrounding neurons, suggesting an enrichment in neurons, particularly in synaptic membranes, rather than in glial cells \textsuperscript{78}.

Brain gangliosides have been proposed to be important in neurogenesis, neurite outgrowth, synaptogenesis, and synaptic function; however, the mechanisms underlying these roles have not been completely elucidated. Abolishing all ganglioside production by knockout of glucosylceramide synthase causes abrupt embryonic lethality after division of the primitive germ layers \textsuperscript{79}. Neural-specific disruption of glucosylceramide synthase in mice produces live births, but the animals have severe neural abnormalities and die at 2–3 weeks of age, suggesting a critical impairment of appropriate brain maturation \textsuperscript{80}. However, mice lacking only complex gangliosides were reported not to have any major abnormalities in histology or behavior, only a decrease in neural conduction velocity at 10 weeks of age \textsuperscript{81}. Thus it has been hypothesized that gangliosides play a role in neural membrane function, such as signaling, conduction, or stability \textsuperscript{82}. However, another group found that mice lacking complex gangliosides were found by 16 weeks of age to exhibit more severe phenotypes, including progressive axonal degeneration of optic and sciatic nerves, as well as decreased central myelination \textsuperscript{83}. 
The phenotype found in this study was noted to be similar to that caused by a deficiency of another important sialic acid-binding lectin, myelin-associated glycoprotein (MAG, Siglec-4), suggesting a common pathway for these molecules. MAG is found periaxonally on Schwann cells and oligodendrocytes in both peripheral and central nervous systems. Notably, it preferentially binds the epitope Neu5Acα2-3Galβ1-3GalNAc, found on the complex neuronal gangliosides GD1a and GT1. Deficiency of MAG is associated with progressive signs of peripheral demyelination and axonal degeneration, similar to demyelinating peripheral neuropathies found in human patients. Additionally, the binding of MAG to gangliosides has been implicated in inhibition of neuronal regeneration after injury.

Claims for Presence of Neu5Gc in Neural Tissues

A number of studies have claimed to find small amounts of Neu5Gc-containing glycans in neural cells and tissues. It is worth considering these studies in some detail.

Tumor gangliosides

Increased Neu5Gc expression is a feature of human cancers, and it might be expected that neural tumors would express Neu5Gc. An extensive characterization of mouse brain tumor gangliosides by Seyfried and colleagues has indeed repeatedly found Neu5Gc present in these gangliosides. However, Neu5Gc is only found when the tumor cells are injected subcutaneously into the flanks of mice, and not when the tumor cells are cultured independently, suggesting that the observed Neu5Gc is exogenous. Subcutaneous injection of an experimental ependymoma incapable of synthesizing endogenous GM2(Neu5Gc) results in a tumor that contains GM2(Neu5Gc) as well other gangliosides suggestive of the presence of Neu5Gc-expressing macrophages. These Neu5Gc-containing gangliosides are still found when the cells are injected into mice with severe combined immune deficiency (SCID), which have no B or T lymphocytes, further suggesting that the gangliosides present do originate from tissue.
macrophages and/or from metabolic uptake by the tumor cells. To date, no characterization of naturally occurring neural tumors has demonstrated endogenous Neu5Gc synthesis.

**Normal CNS gangliosides**

There have, however, been a number of claims that Neu5Gc exists in normal brain gangliosides of some mammals. In 1970, Yu and Ledeen used gas-liquid chromatography to analyze sialic acids in the brain gangliosides of several species. Although the specific gangliosides were not defined, this study found Neu5Gc at 1-2% of total ganglioside sialic acids in ox, bull, and calf, 0.1-0.2% in pig and sheep, 0.4% in goldfish, but undetectable in rat, rabbit, frog, and chicken. More recently, a study of cetacean brain gangliosides found low percentages (<2%) of Neu5Gc in the total ganglioside content of cerebrum and cerebellum of three toothed whales (killer whale, Dall’s porpoise, and sperm whale). No Neu5Gc was found to be present, however, in the brains of other members of the dolphin family or baleen whale species (minke whale and Bryde’s whale).

Other studies have been able to identify specific gangliosides containing Neu5Gc in the brains of certain species. The first and most common such ganglioside to be identified was GD1a containing both Neu5Ac and Neu5Gc, initially estimated to account for 1% of total ganglioside-bound sialic acids in bovine brain. However, a later estimate put the fraction of GD1a(Neu5Ac/Neu5Gc) much lower, at only 0.1% of total bovine brain ganglioside. Using a two-step DEAE-Sepharose and TLC approach to increase resolution, Iwamori and Nagai also identified Neu5Gc-containing GD1a, as well as GM1, in the brain of cow; these were not found in brains of human or chicken, nor in cat, rabbit, rat, nor dog. Another early study of calf and pig brain tentatively identified two unknown gangliosides as Neu5Gc-containing GD1a as well as GM3. GM3(Neu5Gc) has also been identified in equine brain, where Neu5Gc was found to comprise 18% of total GM3 Sias. Finally, Neu5Gc-containing GT1b has been identified in extracts from bovine brain. There are, therefore, quite a variety of gangliosides identified in these reports.
A major challenge in interpreting all such studies is that they did not separate neural cells, *i.e.*, neurons and glia, from the endothelial cells and blood contents that run throughout the brain. Are the Neu5Gc-containing gangliosides that are being isolated in these experiments truly neural, or do they instead arise from the non-neural vasculature? The only study to directly address this question is an examination of horse brain. The authors allow that the presence of endothelial cells may affect their results; however, they did determine that the lipid composition associated with this ganglioside was 60% 18:0, a feature characteristic of brain gangliosides as compared to aortic endothelial cells, which express a wide range of fatty acids.

A comparison of the types of gangliosides found in neural and endothelial tissues may help to clarify the results of these studies. It can be expected that brain microvasculature, comprising the blood-brain barrier, may have different characteristics than endothelium from other tissues; however, unfortunately, few characterizations of the gangliosides of these cells have been done. Immortalized and cultured human cerebrovascular endothelial cells have been shown to express GM3 (62%), GM2 (18%), GM1 (3%), and GD1a (15%) as the major gangliosides. A similar cell line has GM3 and LM1, with small amounts of GM1, GD1a, GD1b, and GT1b. Cultured microvascular endothelial cells from bovine brain also express GM3 as the major ganglioside component, with approximately 58% of GM3 containing Neu5Gc—although it is important to note that Neu5Gc in these cultured cells could have originated from components of the growth medium. As noted above, the major brain gangliosides are GM1, GD1a, GD1b, and GT1b. Thus, the above identification in horse brain of GM3(Neu5Gc) as a neural ganglioside is puzzling, as GM3 is a very minor component of brain gangliosides but a major fraction of cerebrovascular cells. Unfortunately, the significant overlap in expression of the remaining gangliosides in the two tissues makes it impossible to clarify the published identification of Neu5Gc in GM1, GD1a, and GT1b any further.

The fact that all published studies purporting to find Neu5Gc in brain gangliosides involve a group of closely related species—the even- and odd-toed ungulates and the cetaceans—may support the validity of their findings. Perhaps there is an evolutionary adaptation to allow low
percentages of Neu5Gc within the brain of these animals. However, these species also happen to be mammals with large enough amount of brain tissue to detect very minor ganglioside fractions, and the observation may thus result from a sampling bias.

Overall, the data on Neu5Gc in CNS gangliosides are quite challenging to interpret. While it is impossible to completely rule out endothelial and/or blood contamination in any case, it is also impossible to rule out a low level of Neu5Gc presence in neural cells. Further work will be necessary to fully clarify this issue.

Normal CNS glycoproteins

Although much work has been done to characterize nervous system gangliosides, to our knowledge no study has so far identified Neu5Gc on a brain glycoprotein. It is interesting to note here that polysialyltransferases are able to incorporate a number of unnatural sialic acids into polySia\textsuperscript{102,103}. Recently, our laboratory has demonstrated that cells from a murine neuroblastoma line are similarly able to incorporate Neu5Gc into endogenous polySia\textsuperscript{33}. Although polymers of Neu5Gc are found in the glycoproteins of the eggs of salmonid fish\textsuperscript{104}, Neu5Gc has never been reported in mammalian polysialic acid, neural or otherwise.

PNS glycoconjugates

It is possible that the central nervous system suppression of CMAH and Neu5Gc does not extend to the peripheral nervous system (PNS). Very little work has examined this question. A study of bovine spinal motor neuron gangliosides that reacted to serum antibodies from patients with Guillain-Barré syndrome identified two unknown gangliosides that the authors suggested, although did not show definitively, were GD1a containing one or two Neu5Gc residues\textsuperscript{105}. Additionally, a membrane mixture of the noradrenergic vesicles from bovine sympathetic nerve endings was found to contain close to 50% Neu5Gc\textsuperscript{106}. It is therefore quite possible that Neu5Gc is expressed without consequence in peripheral nerves.
It remains to be seen whether Neu5Gc retains a small presence in or is completely absent from the vertebrate CNS. The repeated finding of Neu5Gc in characterizations of neural gangliosides may indicate a true neural presence. Regardless, the very difficulty of detection and interpretation of these studies makes it clear that any Neu5Gc present is maintained at an extremely low level within the vertebrate brain. Further studies are needed, including in situ staining and/or cell-sorted analyses.

**Possible Mechanisms for the Rarity of Neu5Gc in the Brain**

The published record makes it clear that, for as yet unknown reasons, vertebrate brains have very low levels of Neu5Gc, regardless of the levels in other cell types in the same organism. There are a number of possible mechanisms responsible for this suppression, which has apparently persisted for hundreds of millions of years of vertebrate evolution. The most likely explanation is simply transcriptional repression of CMAH. Indeed, northern blot analysis of Cmah in mouse tissues detected no message in brain. Further, the Allen Brain Atlas, which stores images of in situ hybridization to mouse brain slices, shows no CMAH signal throughout the brain (Allen Mouse Brain Atlas [Internet]. Seattle (WA): Allen Institute for Brain Science. ©2009. Available from: http://mouse.brain-map.org). Microarray analysis of Cmah cDNA in mouse brain gives a low but detectable level, although again endothelial contamination cannot be ruled out in whole brain extracts. Absence of CMAH transcript from brain has also been found by RT-PCR analysis of CMAH in pig tissues. Interestingly, human tissue mRNA microarrays give similar results, indicating that CMAH continues to be transcriptionally regulated long after its pseudogenization in the human lineage.

However, the pathway by which CMAH might be transcriptionally repressed is completely unknown. It may be induced by an extracellular factor, either soluble or membrane-bound. Conversely, it could be cell-intrinsic, a feature of neuronal/gliial differentiation. The latter explanation is supported by the studies of gangliosides of murine neural tumor cells, which remain unable to synthesize Neu5Gc even when grown in culture away from their normal
extracellular environment. Further, this same group found Neu5Gc-containing gangliosides in murine solid tumors of neural origin grown intracerebrally and in brain metastases of subcutaneously grown tumors, demonstrating that although neural cells do not express Neu5Gc, inflammatory cells can maintain Neu5Gc expression even within the cerebral compartment. It is therefore probable that CMAH downregulation is a characteristic of the neural lineage.

Of course, other mechanisms may influence Neu5Gc levels in neural cell types. CMAH may be further regulated at the level of mRNA stability, translational rate, and/or by posttranslational modification. Additionally, it cannot be ruled out that there may be enhanced mechanisms for eliminating Neu5Gc after it is synthesized, either within the neural cell or on the cell surface. The apparent evidence that dietary Neu5Ac, but not Neu5Gc, is incorporated into the developing brain may indicate that such an elimination is in fact taking place. Regardless, we can think of no other molecule that has such an unusual distribution amongst vertebrates: expressed at widely variable levels throughout extra-neural tissues, and yet always at very low levels in the brain. Whatever the mechanism, the neural regulation of Neu5Gc is apparently very tightly controlled.

Is Neu5Gc “Toxic” to the Vertebrate Brain?

The unusual picture of Neu5Gc distribution is suspicious for harmful effects of Neu5Gc on the brain. The degree of suppression is remarkable, with the Neu5Gc fraction being no more than about 2% (see Table 1). This finding implies that Neu5Gc is quite severely detrimental to the brain. To date, no studies have examined whether this presumed toxicity of Neu5Gc does in fact occur.

It is interesting to consider whether toxicity might be cell-intrinsic or extrinsic—that is, whether the toxicity targets a unique property of individual neural cells, or only exerts its effects on the overall organ of the CNS. There is significant evidence to suggest that the latter is the case. Standard cell culture methods for all cell lines, including neurons, frequently include fetal calf serum, a source of Neu5Gc. It has been shown that cells grown in such conditions will take
up free Neu5Gc by macropinocytosis and incorporate it into glycans. This has been shown to be true of human embryonic stem cells as well. Neurons are also routinely grown and differentiated in culture within such Neu5Gc-rich conditions. It is thus unlikely that Neu5Gc is toxic only on an intercellular level within the whole brain.

Despite these indications of a detrimental effect of Neu5Gc to the brain, it would be an unexpected finding. Neu5Gc is expressed to no ill effect throughout a wide range of other tissues, so such an effect would have to specifically target neural tissue. Even more puzzlingly, Neu5Ac, which differs only by a single oxygen atom, is a highly prevalent and critically important molecule in brain glycoconjugates. It will be fascinating to explore this question in future work. A number of mechanisms by which Neu5Gc might exert toxicity in the CNS are considered in the next section.

Possible Mechanisms for the Presumed Toxicity of Neu5Gc in the Brain

Biophysical properties

Neu5Ac and Neu5Gc have some differences in chemical properties. Notably, the presence of the additional hydroxyl group on Neu5Gc may alter its pKa from that of Neu5Ac. This structural change also increases the hydrophilicity of Neu5Gc. These effects are minor when considered on the level of individual molecules. However, a mammalian erythrocyte, for example, contains many millions of surface sialic acid residues and neural cells likely have even more. The overall effect of replacing Neu5Ac with Neu5Gc may thus have major effects on the surface charge and/or hydropathicity of a neural cell as a whole.

Differentiation

It is possible that the toxic effect is one to which neural progenitors or differentiating cells are primarily vulnerable. However, there is no evidence that CMAH and Neu5Gc are expressed in these immature progenitors, nor that their presence is toxic to neural differentiation. In fact, cultured neurons can be differentiated in the laboratory with Neu5Gc-rich fetal calf serum with no
apparent detriment. This explanation also would not account for the continued CNS suppression of CMAH into adulthood. It therefore seems unlikely that the presumed toxicity of Neu5Gc is mediated at the level of differentiation.

**Sialoglycoconjugates**

The presence of Neu5Gc in one of the sialic acid-containing molecules previously discussed may cause a detrimental effect. Although most sialyltransferases exhibit a preference for CMP-Neu5Ac or CMP-Neu5Gc, most will utilize either substrate if available. Does Neu5Gc incorporation into a characteristic brain sialoglycoconjugate cause toxic aberrant function? Possibilities include polysialic acid; one study found polymers of Neu5Gc to be less effective at binding BDNF, although the authors acknowledged these were also shorter than the control Neu5Ac polymers. The conversion of Neu5Ac to Neu5Gc in neuroblastoma-glioma hybrid cells also abrogates MAG binding.

Alternatively, perhaps Neu5Gc inhibits synthetic or degradative enzymes of brain sialoglycoconjugates, mediating toxicity not by affecting the function but by blocking the normal turnover of Neu5Ac-containing glycans. Indeed, a recent paper from our own laboratory demonstrated that Neu5Gc is relatively resistant to degradation by sialidase when present in the α2-8 linkage common in brain glycans, including polySia. This phenomenon may allow for a very small fraction of Neu5Gc to be highly detrimental to the brain, as an entire chain of polySia can be rendered resistant to breakdown by the presence of a terminal Neu5Gc. To date, this is the only hypothesis explaining a potential mechanism for Neu5Gc toxicity.

**Alternative role of CMAH**

Lastly, although speculative, it is conceivable that CMAH has an additional role aside from converting Neu5Ac into Neu5Gc, and that it is this alternative function that requires suppression to avoid toxicity. In fact, a recent study of human stem cells suggested that human CMAH, although inactive as a hydroxylase, increases cellular uptake of exogenous Neu5Gc and
decreases Wnt/β-catenin signaling. However, this work depended heavily on an incorrectly reported N-truncated cDNA sequence; the full-length human cDNA actually contains a stop codon upstream of the incorrect start site used. Moreover, the raw data presented do not seem to support the claimed correlations. With the exception of this study, there is no work examining the possibility of an alternative role for CMAH. However, the continued tissue-dependent regulation of CMAH mRNA in humans may indicate that one exists. Perhaps this alternative role is the true mediator of neural toxicity.

The fact that makes few of these explanations entirely satisfactory is that Neu5Gc is not merely low in the brain, but it is almost—perhaps completely—nonexistent. Any model of Neu5Gc toxicity in the vertebrate CNS will have to explain an evolutionarily selective effect requiring that Neu5Gc levels remain below 1-2% of total brain Sias. We believe this feature makes an inhibition of degradation hypothesis the only viable possibility at present. The fact that α2-8 linked Sias are uncommon outside the nervous system and widely distributed in brain glycoconjugates supports this notion.

Evolutionary implications

If Neu5Gc expression is truly toxic to the vertebrate CNS, the expression or lack thereof of CMAH in a given animal may have implications for its neural evolution. The ancestral origins of CMAH are unfortunately not well understood, yet they may be important to our understanding of Neu5Gc in modern vertebrates. N-acetylmuramic acid hydroxylase (namH) has been identified in certain species of Mycobacteria, where it converts UDP-N-acetylmuramic acid to UDP-N-glycolylmuramic acid. Although the homology between namH and murine CMAH is only 12% at the peptide level, namH remains the most likely explanation as to the ancestral source of vertebrate CMAH. Unlike vertebrates, certain early sialic acid expressing invertebrates, such as the echinoderms, actually express a predominance of Neu5Gc, perhaps indicating that the development of CMAH and Neu5Gc was of importance in early lineages.
Within vertebrates, however, species that have since lost Neu5Gc may have gained some neural advantage in the bargain. Admittedly, it is difficult to see what benefit the brains of birds and reptiles have gained from losing Neu5Gc. But the concentration of sialic acid in brain tissue in humans is reported to be extremely high—2-4 fold that of most other mammalian species, and slightly increased over that of chimpanzees. Particularly considering the critical role that polysialic acid plays in neural plasticity, outgrowth, and myelination, it is tempting to speculate that the outright loss of neural Neu5Gc in an already sialic acid-rich brain may have eliminated a residual structural constraint and enabled the evolution of a larger, more complex, and more plastic brain in humans.

**Conclusions and Future Prospects**

The rarity of Neu5Gc in the vertebrate brain is certainly a fascinating observation. The literature to date indicates that this absence is highly conserved, with no animal expressing more than 1-2% Neu5Gc on neural gangliosides. Given the prominent sialylation of important neural glycoconjugates, such as gangliosides and NCAM, the striking absence of CMAH and Neu5Gc from neural cells is highly unusual. Although these findings are suspicious for a detrimental effect of Neu5Gc on the CNS, there is a dearth of studies examining this question. A number of questions will therefore need to be addressed in future work.

1. How conserved is Neu5Gc suppression?

   It is not yet clear whether some species are able to maintain a low percentage of Neu5Gc expression within neural tissue. It will be interesting to see whether some Neu5Gc does persist in the neural tissue of certain species, and to consider what implications this may have for those species. Additionally, no studies to date have observed whether there is similar suppression of Neu5Gc expression in the nervous systems of invertebrate deuterostomes, such as starfish and sea urchin.
2. What is the precise localization of residual Neu5Gc within the CNS?

From existing studies, it cannot be established whether trace Neu5Gc presence is confined to endothelial and blood cells or does extend across the blood brain barrier. If it exists within the CNS itself, cell-specific analyses will be necessary to determine whether all, or only some cell types are able to express it. We are currently studying this issue.

3. What is the developmental regulation of CMAH and Neu5Gc?

No study to date has examined whether Neu5Gc is endogenously expressed in neural progenitor cells or embryonic stem cells. This knowledge may help to clarify the mechanisms and the significance of Neu5Gc absence in mature brain cells.

4. What mechanisms cause the suppression of CMAH expression and Neu5Gc production in the CNS?

This question will be particularly interesting given the unique distribution of CMAH, which exhibits dynamic regulation in non-neural tissues but strict neural suppression. Rigorous examination of the transcriptional regulation of CMAH is a necessary initial step.

5. Does the presence of CMAH and/or Neu5Gc truly cause detrimental effects to the vertebrate brain?

Animal models are needed to address this question, since, as previously discussed, Neu5Gc is not toxic in cell culture. It will further be necessary to study whole brain tissue as well as individual cells. It is worth examining overexpression of CMAH and Neu5Gc separately in this work, as an additional role of CMAH mRNA or protein has not been ruled out.

These questions are critical in determining whether the repeatedly observed “smoke” of CNS rarity of Neu5Gc truly represents clues to a real “fire”. In our ongoing and future work we are beginning to explore this decades-old unexplained observation, using a variety of methods. Our
preliminary studies are promising in this regard. The evidence to date is certainly intriguing, and further investigation may help our understanding of Sias and the CNS.

This chapter is, in full, a reprint of material originally published in Topics in Current Chemistry by authors Davies LRL and Varki A, 2013, with kind permission of Springer Science+Business Media. The dissertation author was the primary author of this review article.

REFERENCES


CHAPTER 2

The Roles of Polysialic Acid in Brain Development
As discussed in Chapter 1, sialic acids have numerous and highly diverse roles in the vertebrate brain. In particular, polysialic acid (polySia), a molecule that is highly expressed throughout vertebrate development, has been implicated in a wide variety of important processes at various stages of development. We have hypothesized that Neu5Gc presence in polySia may impair its degradation by vertebrate sialidases, providing a possible explanation for the conserved exclusion of Neu5Gc from the vertebrate brain. As we explored the diversity of critical roles this glycan plays, we began to wonder whether it might have other evolutionary implications—specifically, within the human brain.

PolySia plays a number of vital roles throughout brain development from embryos to adults. It has been implicated in processes of growth, migration, plasticity, and learning and memory. Notably, some features of the unique human phenotype—a large brain, with high complexity, plasticity, and capacity for learning—may rely on these very processes. Further, a human-specific change in a highly conserved residue of one of polySia's biosynthetic enzymes has been identified, currently with unknown effect. Is polySia differentially regulated in humans as compared to our primate relatives? This question has not been specifically explored to date; however, polySia is an excellent candidate for some of the changes in brain tissue found along the human lineage.

**Unique features of the human brain**

Although a long list of characteristics distinguishes humans from our closest relatives, chimpanzees and other great apes, our brain differences are perhaps the most intriguing. This remarkable organ underlies many of the unique qualities of our species. As the tissue responsible for our intelligence and problem-solving skills, language, social interaction, and our very concept of ourselves and each other, it is perhaps the brain that most deeply defines our sense of what it means to be human.

About six million years have passed since we diverged from the chimpanzee lineage, allowing the biological changes that underlie our phenotype to occur. However, when it comes to
defining these changes and the mechanisms by which they have produced our cognitive skills, the question is surprisingly challenging to answer.

To date, a number of uniquely human features have been identified. Of this lengthy list, we will touch upon three characteristics potentially relevant to polySia biology: size, growth, and maturation.

**Increased size**

Large size is perhaps the most highly discussed feature of the human brain. Although primates themselves have large brains among mammals, among primates humans have an exceptionally large endocranial volume. Endocranial casts of fossil hominid ancestors demonstrate a major expansion around the time of *Homo erectus*, with further increases in size in modern humans \(^1\). Today, the adult human brain is estimated to be about three times as large as expected for an ape of the same size, such as a chimpanzee \(^2\).

Although it is clear that the human brain is unusually large, the specific details of its size have proven surprisingly controversial. It has been demonstrated that the human brain has increased volume of neocortical gray matter, but that the expansion is much more significant in subcortical white matter \(^3\). Although it was initially believed that this expansion was enhanced in the human frontal lobe, more recent studies have not supported this. There is, however, an increase in the gyral white matter in the frontal lobe, suggesting perhaps improved local connectivity \(^4-6\). On the other hand, the temporal lobes, associated with language in humans, do appear to have expanded more than other brain areas \(^7\). The general pattern of expansion seen in white matter throughout the brain, however, suggests that the human cognitive phenotype may depend on connectivity and networks throughout the whole brain, rather than on expansion of any particular region.

**Increased growth rate**
It is unsurprising that a brain of such significant size arises via a striking increase in the rate of growth. The rapid fetal phase of growth of the human brain exceeds that of a chimpanzee, allowing human fetuses to double the brain volume of chimpanzees by only 16 weeks of gestation. Further, humans greatly extend rapid growth past the time period typical for other primates. Chimpanzee brain growth slows around 22 weeks of gestation to a postnatal rate approximately equivalent to that of body growth. Human brain growth, on the other hand, continues extremely rapidly throughout gestation, continuing at a high rate even for several postnatal years.

Lengthened period of maturation

Although all mammals have a period of dependence in early life, humans are remarkable for the protracted period of time it takes us to reach adulthood, and our brains reflect this lengthened period of immaturity. A study of the transcriptional profiles of human and chimpanzee cortex throughout development suggested a “neotenic shift” found only in humans. The authors identified a subset of cortical genes, involved in growth and development, which persisted into adulthood in humans only, suggesting a delayed maturation and more “juvenile” adult state.

Further, myelination of axons in neocortex, seen as a marker of circuit maturation, has been found to be both delayed and protracted in humans when compared to chimpanzees. A study directly measuring staining of sections for 2′,3′-cyclic nucleotide 3′-phosphodiesterase (CNP) and myelin-associated glycoprotein (MAG, Siglec-4), marking early growth and full establishment of myelin, respectively, demonstrated that humans are born with their axons almost completely unmyelinated, then slowly myelinate them throughout childhood and well past adolescence. In stark contrast, chimpanzees start at about one-fifth the adult level of myelination and are essentially finished by sexual maturity.

This short list of anatomical characteristics of the human brain starts to build a picture of a unique organ. Large size, rapid growth, and slow maturation may each contribute to making our
brains complex and cognitively powerful. The molecular mechanisms underlying the features themselves, however, also remain unclear.

**Molecular changes between human and chimpanzee brains**

Understanding how molecular changes translate into anatomical and then functional differences between human and chimpanzee brains is a major challenge. Decades ago, a seminal paper from King and Wilson found over 99% similarity between the human and chimpanzee peptide sequences of a number of proteins. They suggested then that it was likely to be changes at the level of expression, rather than changes in protein coding sequences, that are primarily responsible for the phenotypic differences between the two species.

Accordingly, identifying specific changes underlying the evolution of the human brain has proven challenging. Perhaps the most famous discovery is that of FOXP2—unsurprisingly, a transcription factor. The coding sequence demonstrates a significantly accelerated rate of change along the human lineage, with two changes in the peptide sequences, and a recently identified third intronic change that affects protein expression. In humans, mutations in FOXP2 have been associated with language deficits; replicating the human mutations in mice causes changes in the basal ganglia circuits, including some slight changes in vocalization.

Several other genes show changes along the human lineage that may be relevant to human brain evolution. Further examples include MCPH1 and ASPM1, two genes involved in brain size, though perhaps not associated with changes in cognition. AHI1, a gene involved in normal axonal decussation, has accelerated among primates and especially along the human lineage. (For further discussion, see review)

One notable molecular change specifically related to sialic acid biology is the neural presence of Siglec-11 in humans. Siglec-11 is a member of a family of sialic-acid binding immunoglobulin-like lectins (Siglecs) that is thought to bind to polysialic acid (polySia) epitopes. Siglec-11 is expressed in human, but not chimpanzee microglia, likely due to a gene conversion event with the pseudogene SIGLECP16 that has occurred on the human lineage. The
interaction of Siglec-11 with neuronal polySia has been implicated as an inhibitor of microglia-mediated neurotoxicity \textsuperscript{19}. This suggests that the presence of Siglec-11 may cause different patterns of microglial phagocytosis or pruning in humans. Further work, however, is necessary to fully detail the human-specific effects of microglial Siglec-11.

A host of other studies have taken creative approaches to analyzing the differences in gene expression, rather than protein coding sequence. These studies have made important progress in understanding the genetic basis of the human brain, with analyses ranging from microarray expression data \textsuperscript{20} to more recent studies of regions demonstrating accelerated change on the human lineage \textsuperscript{21}, maps of transcriptional networks \textsuperscript{22}, and patterns of methylation \textsuperscript{23,24}. In all of these cases, the identified differences in the human brain emerge from a complex combination of many individual changes. Nevertheless, this work is helpful in elucidating the network of changes involved in developing the human brain.

The difficulty of tackling these questions, and the work so far completed, have made clear that the story of the human brain is a complex one. It is likely that an array of changes, in conjunction with environmental interactions, are responsible for the unique human phenotype.

**Polysialic acid is an important vertebrate brain glycoconjugate**

A previously ignored candidate for human-chimpanzee brain differences is the critical brain polymeric glycan polysialic acid (polySia). Discussed briefly in the first part of the introduction, polySia is an important brain glycoconjugate that is highly expressed in embryonic development and decreases to much lower levels in postnatal life. It consists of lengthy linear chains of α2-8 linked Neu5Ac, ranging in length from about 8 to over 100 individual residues \textsuperscript{25,26}. Two vertebrate polysialyltransferases, ST8Sia-II (also called STX) and ST8Sia-IV (PST) are responsible for the synthesis of polySia \textsuperscript{27-29} (Figure 2-1).

PolySia is found on a number of neural glycoproteins, foremost among them the neural cell adhesion molecule (NCAM) \textsuperscript{30}. It has also been identified on additional carrier proteins in the nervous system, including the synaptic cell adhesion molecule 1 (SynCAM 1) and the voltage-
gated sodium channel (in rat brain) \(^{31,32}\). The final proteins that have been identified as being polysialylated are the enzymes themselves; this modification increases, but is not necessary for, their ability to polysialylate other proteins \(^{33-35}\).

**Molecular basis for polySia synthesis on NCAM**

Although several carrier proteins are modified by polySia, it is NCAM that has been most extensively characterized. NCAM can be polysialylated by either polysialyltransferase, although the responsibility *in vivo* shifts from primarily ST8Sia-II in development to primarily ST8Sia-IV in the adult brain \(^{36}\). Polysialylation can occur at two N-glycosylation sites, located within the fifth immunoglobulin-like (Ig5) domain \(^{37}\), where it requires initial priming by an α2-3 or 2-6 linked Neu5Ac \(^{38,39}\). The polysialyltransferases require sequences in this domain as well as in the first fibronectin type III repeat (FN1) to recognize the substrate and appropriately act upon it \(^{40}\).

SynCAM 1 has only recently been identified as an additional neural carrier of polySia \(^{31}\). It has three Ig-like domains; a single N-glycosylation site in the Ig1 domain has been identified as the exclusive site of polysialylation, although recognition of both Ig1 and Ig2 domains are necessary for enzyme activity \(^{31,41}\). Unlike NCAM, SynCAM 1 is exclusively polysialylated by ST8Sia-II \(^{41}\).

All animal sialyltransferases share a number of conserved domains, or sialylmotifs, thought to be important for their activity: sialylmotif L (large), S (small), VS (very small), and III. (Figure 2-2). Sialylmotif L and S are generally involved in substrate binding, whereas sialylmotif VS and III are involved in catalytic activity of the enzyme \(^{42}\). The polysialyltransferases, ST8Sia-II and IV, have additional regions that have been implicated in their unique activity. The polysialyltransferase domain (PSTD), enriched for basic amino acids, is proposed to interact with the newly synthesized polySia chain as the enzyme acts processively \(^{43}\). Additionally, a polybasic region (PBR), also enriched for basic residues, is involved in recognition of an acidic patch in the first fibronectin type III repeat of the substrate NCAM protein (Figure 2-2) \(^{44,45}\).
Figure 2-1. Structure of polySia on NCAM. A: PolySia (dark gray) on Ig5 domain of NCAM produces a hydrated, anti-adhesive zone (light gray) around the molecule. Specific regions indicated are immunoglobulin-like domains (Ig1-Ig5) and fibronectin type III repeats (FN1 and FN2). B: Molecular structure of polySia, demonstrating α2-8 linkage of Neu5Ac.
Expression of polySia in the developing brain

PolySia has been identified developmentally in all three germ layers in early embryogenesis. However, it is most highly expressed in the central nervous system. PolySia-NCAM is found at high levels in embryonic life, emerging around E10 in mice and E13 in rats.

In fact, because of the absolute polysialylation of NCAM at this developmental stage, several early studies refer to polySia-NCAM as the embryonic form of NCAM. The expression of polySia-NCAM then rapidly drops postnatally. In mice, for example, polySia in whole brain tissue is high at birth but decreases dramatically within the first weeks of life to reach adult levels by postnatal day 20. Although NCAM expression persists into adulthood, the fraction that remains modified by polySia is quite small.

In humans, the developmental pattern is less well studied. However, one study of prefrontal cortex found that NCAM is most highly polysialylated in fetal tissue, dropping throughout infancy and early childhood to reach low levels by about age five. A study focusing on the hippocampus found a similar decline in polySia-NCAM expression from infancy to early adulthood; this was accompanied by a shift in the distribution of polySia from exclusively within the dentate gyrus to primarily within the hilus. Thus it appears from the available data that human expression of polySia follows a generally similar pattern to that seen in rodents, although with different timing per the very different developmental time scales.

The cell type distribution of polySia during development is remarkably broad, spanning both neuronal and glial populations. It has been identified in interneuron precursors in the subventricular zone and during their migration to the olfactory bulb and cortex. It also appears on the cell bodies of cerebellar granule cells, hippocampal neurons, axons of several major fiber tracts, and oligodendrocyte precursors (for review, see ). This broad expression reflects the diversity of developmental processes involving polySia.

Polysialylated SynCAM 1, only recently characterized, has a much more constrained pattern of expression. It is identified solely on a subset of NG2 cells, a population of glial cells that
serve as precursors of myelinating oligodendrocytes. PolySia-SynCAM 1 also exhibits a temporally specific expression in rodents, vanishing in the first days of postnatal life.

Expression of polySia in the adult brain

Adult expression of polySia is much more limited than that found in the embryo. It can be found most prominently in areas associated with immaturity or plasticity, such as the subventricular zone and the dentate gyrus of the hippocampus. Additionally, it can be found on interneurons of prefrontal cortex, piriform cortex, amygdala, and hippocampus (For review, see 51).

Notably, polySia is also found to be upregulated in cases of injury in adults. Occlusion of the middle cerebral artery in rats causes upregulation of polySia within that area. Further, cortical aspiration lesions induce an increase in polySia expression and total cell number in the subventricular zone. These findings suggest that polySia plays an important role in the response to injury of the adult central nervous system.

Roles of polysialic acid

Since its initial identification 30 years ago a tremendous body of literature has developed implicating polySia in a wide array of developmental processes. Most of this work focuses on polySia present on NCAM; however, recently SynCAM 1 has been included in this examination.

A more complete list of the roles of neural polySia that have been studied to date is presented in Table 1. Here we will highlight several processes in which polySia has been implicated.

Differentiation

PolySia-NCAM is typically associated with precursor or immature cells. Accordingly, treatment of the subventricular zone or subgranular layer of dentate gyrus with endoneuraminidase (endoN), a bacteriophage enzyme that cleaves polySia, is associated with premature differentiation of neuronal progenitors, as measured by cell surface markers of mature
neurons. Similarly, overexpression of ST8Sia-IV within oligodendrocyte precursors prevents their proper differentiation and consequent myelinating activity.

**Neurite outgrowth**

The emergence of neurite outgrowth in progenitors is indicative of differentiation, and indeed this is enhanced with endoN-treatment of the subventricular zone. Further studies have implicated polySia in neurite outgrowth in cell culture. Rat hippocampal neurons treated with RNAi to St8sia2 have shorter neurites; conversely, 3T3 neuroblastoma cells can be induced to have longer neuritic processes by transfection with St8sia2.

**Migration**

The rostral migratory stream (RMS), leading to the olfactory bulb, was one of the first regions identified as having high expression of polySia-NCAM in the adult brain. In the absence of NCAM or polySia, the RMS is significantly reduced, resulting in a smaller olfactory bulb. Enzymatic removal of polySia produces similar effects in other systems. EndoN injection into the hippocampus impairs the migration of neuronal progenitors. In culture, SH-SY5Y neuroblastoma cells also demonstrate impaired migration by scratch assay in the presence of endoN.

**Neurotrophin function**

PolySia has been demonstrated in vitro to interact directly with both BDNF and FGF-2. Interestingly, it has been proposed that polySia sensitizes neurons to the effects of BDNF. In culture of cortical neurons, loss of polySia reduces BDNF signaling. Further, addition of BDNF can rescue deficiency in long-term potentiation caused by loss of polySia. This suggests that polySia and BDNF may act synergistically. Conversely, the interaction of polySia with FGF-2 inhibits its ability to stimulate cell growth.
**Myelination**

In further accord with polySia’s reputation as a marker of neural immaturity, its presence has been associated with inhibition of myelination. Temporally, its downregulation correlates with upregulation of myelin basic protein in human development. In mice, overexpression in oligodendrocytes caused by transgenic *St8sia4* under the proteolipid protein promoter causes reduced myelination and lower levels of myelin basic protein. This relationship appears to be the case even in injury; an *St8sia4* knockout mouse remyelinates more quickly after demyelination induced by cuprizone treatment.

**Synaptic plasticity**

PolySia-NCAM has been demonstrated to play an important role in long term potentiation (LTP) and long term depression (LTD) in the CA1 region of the hippocampus. EndoN treatment, as well as knockout of NCAM or ST8Sia-IV, are all associated with abrogation of normal LTP and LTD in the CA1 area of the hippocampus.

**Learning and memory**

The loss of LTP and LTD associated with the absence of polySia in the hippocampus correlates with deficits in memory. Spatial learning in a Morris water maze is significantly impaired in mice lacking the *Ncam* gene and the *St8sia4* gene, as well as in rats who have had endoN injected into their hippocampus. Formation and consolidation of memories during fear conditioning are also inhibited by the presence of exogenous polySia or polySia-NCAM.

**Potential roles of polySia-SynCAM 1**

Most studies of polySia function have focused on polySia-NCAM. It is currently unclear what role polySia on SynCAM 1 may play in brain development. However, polySia on SynCAM 1, similarly to that on NCAM, reduces the adhesive capacity of these molecules. SynCAM 1 adhesion is known to be involved in synapse formation; notably, its overexpression causes
Table 2-1. Processes and phenotypes associated with polySia. Selected publications are described that identify polySia as being involved in various developmental cellular processes, behavioral phenotypes, and pathologies.

<table>
<thead>
<tr>
<th>Process or phenotype</th>
<th>Finding</th>
<th>Location</th>
<th>Model</th>
<th>Citation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Migration</strong></td>
<td>EndoN treatment of SH-SY5Y neuroblastoma cells impairs migration on scratch assay.</td>
<td>n/a</td>
<td>SH-SY5Y cells</td>
<td>Petridis AK et al., 2004 54</td>
</tr>
<tr>
<td></td>
<td>Loss of polySia is associated with defective rostral migratory stream, small olfactory bulbs. Gain of underlying NCAM is associated with hypoplasia of several tracts.</td>
<td>various</td>
<td>mouse</td>
<td>Weinhold B et al., 2005 74</td>
</tr>
<tr>
<td></td>
<td>EndoN injection into hippocampus impairs migration of neuronal progenitors in subgranular layer of dentate gyrus.</td>
<td>hippocampus</td>
<td>mouse</td>
<td>Burgess A et al., 2008 55</td>
</tr>
<tr>
<td><strong>Differentiation</strong></td>
<td>EndoN injection into subventricular zone causes premature differentiation of explanted cells.</td>
<td>SVZ</td>
<td>mouse</td>
<td>Petridis AK et al., 2004 54</td>
</tr>
<tr>
<td></td>
<td>EndoN injection into hippocampus induces differentiation of progenitors in subgranular layer of dentate gyrus.</td>
<td>hippocampus</td>
<td>mouse</td>
<td>Burgess A et al., 2008 55</td>
</tr>
<tr>
<td><strong>Neurite outgrowth</strong></td>
<td>Transfection with St8sia2 causes increased neurite outgrowth (length).</td>
<td>n/a</td>
<td>3T3 cells</td>
<td>Franceschini I et al., 2001 59</td>
</tr>
<tr>
<td></td>
<td>RNAi against St8sia2 causes shorter neurites.</td>
<td>hippocampal neurons</td>
<td>cultured, rat neurons</td>
<td>Brocco MA and Frasch AC, 2006 57</td>
</tr>
<tr>
<td></td>
<td>Removal of polySia with NEU4 is associated with shorter neurites.</td>
<td>hippocampal neurons</td>
<td>mouse</td>
<td>Takahashi K et al., 2012 76</td>
</tr>
<tr>
<td><strong>Myelination</strong></td>
<td>Downregulation of polySia correlates with upregulation of myelin basic protein.</td>
<td>various</td>
<td>human</td>
<td>Jakovcecski I et al., 2007 55</td>
</tr>
<tr>
<td></td>
<td>Overexpression of St8sia4 under proteolipid protein promoter impairs myelination.</td>
<td>forebrain</td>
<td>mouse</td>
<td>Fewou SN et al., 2007 56</td>
</tr>
<tr>
<td></td>
<td>St8sia4 knockout causes increased oligodendrocyte precursor differentiation and faster remyelination after cuprizone-induced demyelination.</td>
<td>corpus callosum, cortex</td>
<td>mouse</td>
<td>Koutsoudaki PN et al., 2010 66</td>
</tr>
</tbody>
</table>
**Table 2-1, continued.**

<table>
<thead>
<tr>
<th>Neurotrophin function</th>
<th>Loss of polySia reduces BDNF signaling in neurons to BDNF.</th>
<th>PolySia with DP&gt;12 binds BDNF directly.</th>
<th>PolySia binds FGF2 and inhibits its action on growth of 3T3 cells.</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CA1 of hippocampus</td>
<td>cultured cortical neurons</td>
<td>in vitro</td>
<td>n/a</td>
<td>Mouse S et al., 2011</td>
</tr>
<tr>
<td></td>
<td>Muller D et al., 2001</td>
<td>Vutskits L et al., 2001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Muller D et al., 2001</td>
<td>Vutskits L et al., 2001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Synaptic plasticity</td>
<td>Hippocampal LTP and long term depression (LTD) are impaired with loss of polySia.</td>
<td>St8sia4 knockout adult, but not young, mice show reduction in LTP, LTD in CA1 of hippocampus.</td>
<td>Loss of long term potentiation (LTP) in CA1 in NCAM and polysialyltransferase knockout mice is restored by antagonist to GluN2B NMDA receptors.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>hippocampus</td>
<td>hippocampus</td>
<td>hippocampus</td>
<td>mouse</td>
<td>Muller D et al., 1996</td>
</tr>
<tr>
<td></td>
<td>Muller D et al., 1996</td>
<td>Eckhardt et al., 2000</td>
<td>Kochlamazashvili et al., 2010</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Learning and memory</td>
<td>Knockout of Ncam impairs spatial learning in Morris water maze.</td>
<td>EndoN injection into hippocampus impairs spatial learning in Morris water maze.</td>
<td>Injection of exogenous PSA-NCAM or polySia interferes with formation and consolidation of contextual memory.</td>
<td>St8sia4 knockout mice have impaired spatial and reversal learning, but normal fear conditioning.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>whole animal</td>
<td>hippocampus</td>
<td>hippocampus</td>
<td>whole animal</td>
<td>Cremer H et al., 1994</td>
</tr>
<tr>
<td></td>
<td>Becker CG et al., 1996</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Senkov O et al., 2006</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Markram K et al., 2007</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 2-1, continued.

<table>
<thead>
<tr>
<th>Injury and repair</th>
<th>Cortical injury induces polySia-NCAM expression in the subventricular zone.</th>
<th>SVZ</th>
<th>rat</th>
<th>Szele FG and Chesselet MF, 1996</th>
</tr>
</thead>
<tbody>
<tr>
<td>PolySia is upregulated in response to injury caused by occlusion of the middle cerebral artery (MCA).</td>
<td>MCA territory</td>
<td>rat</td>
<td>Hayashi T et al., 2001</td>
<td></td>
</tr>
<tr>
<td>In absence of polySia, appropriate targeting of regenerating peripheral motor neurons is lost.</td>
<td>peripheral nerve</td>
<td>rat</td>
<td>Franz CK et al., 2005</td>
<td></td>
</tr>
<tr>
<td>Lentiviral delivery of St8sia4 can aid recruitment of progenitors from SVZ to cortical lesion.</td>
<td>SVZ</td>
<td>mouse</td>
<td>El Maarouf A et al., 2006</td>
<td></td>
</tr>
<tr>
<td>Lentiviral delivery of St8sia4 to glial cells aids axonal regeneration after transection.</td>
<td>spinal cord</td>
<td>rat</td>
<td>Zhang Y et al., 2007</td>
<td></td>
</tr>
<tr>
<td>Implanted Schwann cells overexpressing polySia have improved migration and resultant axonal growth.</td>
<td>spinal cord</td>
<td>rat</td>
<td>Ghosh M et al., 2012</td>
<td></td>
</tr>
<tr>
<td>Schizophrenia</td>
<td>Less polySia-NCAM is found in the hilar region of hippocampus in schizophrenic patients.</td>
<td>hippocampus</td>
<td>human</td>
<td>Barbeau et al., 1995</td>
</tr>
<tr>
<td>Promoter region of ST8SIA2 is associated with schizophrenia in a Japanese population by GWAS.</td>
<td>n/a</td>
<td>human</td>
<td>Arai M et al., 2006</td>
<td></td>
</tr>
<tr>
<td>Promoter region of ST8SIA2 is associated with schizophrenia in a Han Chinese population by GWAS.</td>
<td>n/a</td>
<td>human</td>
<td>Tao R et al., 2007</td>
<td></td>
</tr>
<tr>
<td>GWAS for bipolar disorder in Italians finds linkage at 15q26, near ST8SIA2.</td>
<td>n/a</td>
<td>human</td>
<td>Vazza G et al., 2007</td>
<td></td>
</tr>
<tr>
<td>Loss of polySia in St8sia2 and St8sia4 knockout mice causes abnormal connectivity “reminiscent of schizophrenia” including callosal hypoplasia.</td>
<td>various</td>
<td>mouse</td>
<td>Hildebrandt H et al., 2009</td>
<td></td>
</tr>
<tr>
<td>Familial mutation in ST8SIA2 associated with schizophrenia gives lower enzymatic activity and shorter resulting polySia chains.</td>
<td>in vitro</td>
<td>human</td>
<td>Isomura R et al., 2011</td>
<td></td>
</tr>
<tr>
<td>Reduced staining for polySia-NCAM is found in dorsolateral prefrontal cortex in schizophrenic patients.</td>
<td>dorsolateral prefrontal cortex</td>
<td>human</td>
<td>Gilabert-Juan K et al., 2012</td>
<td></td>
</tr>
<tr>
<td><strong>Depression</strong></td>
<td>PolySia-NCAM is increased in medial prefrontal and hippocampus with chronic use of the antidepressant imipramine.</td>
<td>prelimbic cortex, hippocampus</td>
<td>rat</td>
<td>Sairanen M et al., 2007&lt;sup&gt;88&lt;/sup&gt;</td>
</tr>
<tr>
<td>----------------</td>
<td>----------------------------------------------------------------------------------------------------------------</td>
<td>-------------------------------</td>
<td>-----</td>
<td>----------------------------------</td>
</tr>
<tr>
<td></td>
<td>EndoN injection in the rat hippocampus increases CA3 dendritic arborization, increases vulnerability to excitotoxicity, exacerbates stress response to chronic immobilization stress.</td>
<td>hippocampus</td>
<td>rat</td>
<td>McCall T et al., 2012&lt;sup&gt;89&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>PolySia-NCAM staining is decreased in basolateral and basomedial amygdala in depressed patients.</td>
<td>amygdala</td>
<td>human</td>
<td>Varea E et al., 2012&lt;sup&gt;90&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>PolySia-NCAM staining is increased in basolateral amygdala in depressed non-suicides.</td>
<td>amygdala</td>
<td>human</td>
<td>Maheu ME et al., 2013&lt;sup&gt;91&lt;/sup&gt;</td>
</tr>
<tr>
<td><strong>Bipolar disorder</strong></td>
<td>GWAS association for bipolar disorder in Italians finds linkage at 15q26, near ST8SIA2.</td>
<td>n/a</td>
<td>human</td>
<td>Vazza G et al., 2007&lt;sup&gt;84&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>ST8SIA2 is associated with bipolar disorder by GWAS in Han Chinese.</td>
<td>n/a</td>
<td>human</td>
<td>Lee MT et al., 2011&lt;sup&gt;92&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>PolySia-NCAM is increased in lateral nucleus of amygdala in bipolar patients.</td>
<td>amygdala</td>
<td>human</td>
<td>Varea E et al., 2012&lt;sup&gt;90&lt;/sup&gt;</td>
</tr>
<tr>
<td><strong>Alcoholism</strong></td>
<td>Low polySia-NCAM, through natural variation or artificial depletion, is associated with alcoholism risk behavior.</td>
<td>prefrontal cortex</td>
<td>mouse</td>
<td>Barker JM et al., 2012&lt;sup&gt;93&lt;/sup&gt;</td>
</tr>
<tr>
<td><strong>Epilepsy</strong></td>
<td>Expression of polySia-NCAM is increased in entorhinal cortex and hippocampus in drug-refractory temporal lobe epilepsy.</td>
<td>hippocampus, entorhinal cortex</td>
<td>human</td>
<td>Mikkonen M et al., 1990&lt;sup&gt;94&lt;/sup&gt;</td>
</tr>
<tr>
<td></td>
<td>ST8Sia-IV deficiency shortens latency to status epilepticus induced by kainate, and also increases mortality.</td>
<td>various</td>
<td>mice</td>
<td>Pekce A et al., 2010&lt;sup&gt;95&lt;/sup&gt;</td>
</tr>
<tr>
<td><strong>Social interaction</strong></td>
<td>St8sia4 knockout causes decreased social interaction; St8sia2 knockout also causes increased aggression.</td>
<td>whole animal</td>
<td>mouse</td>
<td>Calandreau L et al., 2010&lt;sup&gt;96&lt;/sup&gt;</td>
</tr>
</tbody>
</table>
excess excitatory synapse formation. Since NG2 cells do form glutamatergic synapses, it has been hypothesized that polySia specific to these cells is involved in the regulation of those synapses.

**Phenotypes associated with dysregulation of polysialic acid**

Dysregulation of polySia has a number of associated negative phenotypes. These have been characterized quite thoroughly in rodent models. PolySia has also been noted to be associated with a number of human disease states, particularly with psychiatric illnesses (Table 2-1). Mouse models have been helpful in revealing the phenotypes associated with loss of ST8Sia-II and IV. Comparison of the neural anatomy of polysialyltransferase knockout mice with that of NCAM knockout mice has revealed defects two categories of defects: those caused by loss of independent polySia function, and those caused by gain of function of the underlying NCAM. While the former category is present in mice lacking either polysialyltransferases or NCAM, the latter category is present only in mice lacking polySia, and the phenotype is rescued with the additional knockout of NCAM. Phenotypes attributed to loss of function of polySia include defective rostral migratory stream with small olfactory bulbs, and delamination of hippocampal mossy fibers. Those attributed to gain of function of NCAM include absence or hypoplasia of several tracts, including the anterior commissure, corpus callosum, and internal capsule, as well as hydrocephalus. Further studies have demonstrated that the amount of polySia-free NCAM correlates with the severity of these phenotypes. ST8Sia-II and ST8Sia IV are, to some extent, able to compensate for each other in individual knockouts; however, a lack of both polysialyltransferases is lethal, causing failure to thrive and death within the first few weeks of life.

In contrast, the overexpression of polySia has not been thoroughly explored in animal models. A single study overexpressed St8sia4 under the proteolipid protein promoter in mouse oligodendrocytes, increasing and extending the expression of developmental polySia. The
identified result was defective oligodendrocyte differentiation and impaired myelination. Phenotypes resulting from the overexpression of polySia in other cell types have not been examined.

In humans, polySia has been associated with disease states—specifically, with psychiatric diseases. Most evidence to date supports an association with schizophrenia. Three genome-wide association studies, in Japanese, Han Chinese, and Italian populations, have associated polymorphisms in the promoter region of **ST8SIA2** with schizophrenia. Additionally, analysis of schizophrenic patients’ brain tissue has shown decreased presence of polySia in the hilar region of the hippocampus as well as in prefrontal cortex. Recently, a mutation in the coding sequence of ST8Sia-II found in a family with heritable schizophrenia was demonstrated to reduce the activity of the enzyme, resulting in shorter polySia chains with less capacity to bind neurotrophins.

A case has also been made for the association of polySia with major depression, although this evidence is not as clear. Alterations in the levels of polySia-NCAM in the amygdala have been identified in depressed patients; however, the direction of the reported change is inconsistent. In animal models of depression, chronic use of the antidepressant imipramine causes an increase in polySia-NCAM in prefrontal cortex and hippocampus. Correspondingly, removal of polySia by endoN injection into the hippocampus exacerbates the stress response to chronic immobilization stress, a common animal model for depression. In fact, a case has been made that adult neurogenesis in the hippocampus, and the associated roles of polySia-NCAM, are important factors in the pathogenesis of depression.

Although fewer studies have examined other human diseases, there is limited evidence that polySia is increased in bipolar disorder, and decreased in alcoholism (Table 1).

It has been speculated that the pathogenesis of psychiatric diseases such as schizophrenia and autism involves derangements of human-specific brain qualities. Indeed, such diseases are not known to be present in any other animal. It is unclear
Figure 2-2. Known domains of ST8Sia-II. TM: Transmembrane region. SML, SMS, SM3, SMVS: conserved sialylmotifs L, S, III, and VS, respectively. PBR: polybasic region, thought to interact with acidic patch on NCAM. PSTD: polysialyltransferase domain, thought to interact with growing polySia chain. N308K: human-specific change identified at amino acid residue 308, within sialylmotif S.
what such a diagnosis in a non-human animal would even entail, given that the clinical deficits identified in schizophrenia and other diseases are in affect, social interaction, and other qualities that are greatly enhanced in human beings. The involvement of polySia in such diseases thus makes an even stronger case for its role in human evolution.

Humans have a unique mutation in ST8Sia-II

A former publication from our laboratory made use of the available genomic sequence data to find human-specific changes in conserved residues of proteins related to sialic acid metabolism. One of the changes identified was in the polysialyltransferase ST8Sia-II. The single base change occurs at position 308, a highly conserved residue in sialylmotif S. An asparagine is present in a range of species, including primates, rodents, and extending as far as Xenopus. In humans, a single mutation has replaced this asparagine with a basic lysine residue (Figure 2-2).

Analysis of the predicted secondary structure revealed that this coding change was likely associated with a structural change. The presence of the lysine in the human sequence is predicted to obliterate a short ancestral coil, extending the length of an adjacent helix. However, the effect of this change on the activity of the enzyme, and consequently on the human brain, was not pursued at the time.

PolySia as a candidate for human brain uniqueness

PolySia is an interesting candidate for some of the unique features of the human brain. The diverse roles of polySia throughout development potentially allow for similarly diverse results of changes in polySia metabolism. Many of these roles may affect factors such as growth, plasticity, and maturation—factors that, as discussed, are highly relevant to human brain evolution.

The observation of a uniquely human change in the polysialyltransferases ST8Sia-II is particularly of great interest: a human-specific coding change that is predicted to alter the protein
secondary structure of an enzyme involved in essential neural developmental processes. If this change does affect the activity of the enzyme, the predicted effects may be quite significant.

Regardless of the particular outcome of this single change in human ST8Sia-II, however, the importance of polySia within the vertebrate brain make it an excellent candidate for further study within the context of human evolution. It may be a previously overlooked factor impacting the development of the remarkable human brain.

REFERENCES


66. Koutoudaki, P. N. et al. Remyelination after cuprizone induced demyelination is accelerated in mice deficient in the polysialic acid synthesizing enzyme St8siaIV. *Neuroscience* 171, 235-244 (2010).


CHAPTER 3

Resistance of α2-8-Linked N-Glycolyneuraminic Acid to Enzymatic Cleavage
ABSTRACT

The sialic acid N-acetylneuraminic acid (Neu5Ac) and its hydroxylated derivative N-glycolyneuraminic acid (Neu5Gc) differ by one oxygen atom. CMP-Neu5Gc is synthesized from CMP-Neu5Ac, with Neu5Gc representing a highly variable fraction of total Sias in various tissues, and amongst different species. The exception may be the brain, where Neu5Ac is abundant and Neu5Gc is reported to be rare. Here we confirm this unusual pattern and its evolutionary conservation in additional samples from various species, concluding that brain Neu5Gc expression has been maintained at extremely low levels over hundreds of millions of years of vertebrate evolution. Most explanations for this pattern do not require maintaining neural Neu5Gc at such low levels. We hypothesized that resistance of α2-8-linked Neu5Gc to vertebrate sialidases is the detrimental effect requiring relative absence of Neu5Gc from brain. This linkage is prominent in polysialic acid (polySia), a molecule with critical roles in vertebrate neural development. We show that Neu5Gc is incorporated into neural polySia, and does not cause *in vitro* toxicity. Synthetic polymers of Neu5Ac and Neu5Gc showed that mammalian and bacterial sialidases are much less able to hydrolyze α2-8-linked Neu5Gc at the non-reducing terminus. Notably, this difference was not seen with acid-catalyzed hydrolysis of polySias. Molecular dynamics modeling indicates that differences in the 3-dimensional conformation of terminal saccharides may partly explain reduced enzymatic activity. In keeping with this, polymers of N-propionylneuraminic acid are sensitive to sialidases. Resistance of Neu5Gc-containing polySia to sialidases provides a potential explanation for the rarity of Neu5Gc in the vertebrate brain.

INTRODUCTION

The sialic acid N-acetylneuraminic acid (Neu5Ac) and its derivative N-glycolyneuraminic acid (Neu5Gc) differ only by a single oxygen atom, are widely synthesized throughout most animals of the deuterostome lineage, and are commonly positioned as the terminal residues of the glycoconjugates that cover all cell surfaces. The conversion of CMP-Neu5Ac to CMP-Neu5Gc is catalyzed by a single highly conserved enzyme, cytidine monophosphate N-
acetylneuraminic acid hydroxylase (CMAH)\(^3\)-\(^5\). Decades of studies indicate that the Neu5Gc fraction of total sialic acids (Sias) is highly variable amongst tissues and between species\(^6\)-\(^33\). The notable exception appears to be the vertebrate brain. Although this tissue contains more total Sias than any other\(^34\), Neu5Gc is reported to be present at only trace levels, if at all. Studies of mouse and pig tissues have accordingly found \(Cmah\) mRNA to be undetectable in brain\(^35\),\(^36\). As we confirm and expand in this study, no other tissue displays this unusual suppression of Neu5Gc expression across vertebrate taxa.

The evolutionarily conserved suppression of Neu5Gc expression in the vertebrate brain suggests that a brain-specific detrimental effect requires its consistent down-regulation. We have considered a number of possible mechanisms to explain this observation. There is no evidence to suggest that CMAH has any function beyond the conversion of CMP-Neu5Ac to CMP-Neu5Gc. The only known homologue of CMAH is \(N\)-acetylmuramic acid hydroxylase (namH), which also converts an \(N\)-acetyl group to an \(N\)-glycolyl group on muramic acid in the actinomycete bacteria\(^37\). Also, \(Cmah\)^{\(-/-\)} mice do not manifest any gross defect in brain development\(^38\). It is therefore reasonable to assume that CMAH is responsible primarily for the synthesis of Neu5Gc, and that any detrimental effect results from the presence of Neu5Gc and not from a secondary effect of CMAH.

Neu5Ac is widely prevalent on a variety of sialoglycoconjugates (SGCs) in vertebrate brain tissue. We considered the possibility that Neu5Gc may interfere with the synthesis, recognition, or function of one or more SGCs. However, although sialyltransferases or Sia-binding proteins may have a relative preference for Neu5Ac or Neu5Gc, they will generally accept either Sia\(^39\). Even in the cases where there is selectivity, biological function is not likely to require such a remarkably low fraction of Neu5Gc to be maintained. Similarly, while neural lectins such as the myelin-associated glycoprotein (MAG) do prefer Neu5Ac to Neu5Gc\(^40\), a minority of Neu5Gc should not interfere with the recognition of majority Neu5Ac. Rather, based on the remarkably low amounts of Neu5Gc that are maintained in vertebrate brains, we presume that even a very small amount of Neu5Gc must cause toxicity. As a first step towards testing this hypothesis we
propose that the presence of Neu5Gc in SGCs in the vertebrate brain critically impairs their degradation. This model allows for even a very small fraction of Neu5Gc to exert substantial toxicity through widespread resistance of underlying glycans to breakdown.

Of course, any proposed mechanism underlying Neu5Gc toxicity must selectively impact brain tissue. As a first step towards testing our hypothesis that Neu5Gc impairs degradation of brain SGCs, we propose a candidate glycan for this mechanism, polysialic acid (polySia, sometimes called PSA). PolySia is a homopolymer of α2-8-linked Neu5Ac, most prevalent on the neural cell adhesion molecule (NCAM) ⁴¹,⁴², and also found in the capsules of certain neuroinvasive bacteria ⁴³. Polysialylated NCAM is highly expressed through embryonic and early postnatal development and plays critical roles in neurite outgrowth and plasticity ⁴⁴. Its absence in mice causes severe developmental phenotypes ⁴⁵,⁴⁶, and its dysregulation has been implicated in neural disorders such as schizophrenia ⁴⁶-⁵⁰.

The cellular mechanisms responsible for polySia breakdown are not well understood. The molecule has been shown to undergo an intra-molecular self-cleavage into short oligomers over many hours at mildly acidic pH ⁵¹, under conditions that would be found in the lysosome. However, the resulting oligomers are resistant to further intra-molecular self-cleavage, and these would presumably require further digestion by the prevalent lysosomal mammalian sialidase NEU1. Notably, NEU1 has also been implicated in desialylation of cell surface SGCs ⁵²,⁵³, another process which could be critical in appropriate regulation of polySia.

It is well established that bacterial and viral sialidases have a relative preference for Neu5Ac over Neu5Gc in α2-3 and α2-6 linkages ⁵⁴-⁵⁶. Relatively little is known about the four mammalian sialidases NEU1-4 in this regard ⁵⁷. Only NEU2 has been crystallized to date ⁵⁸, although the other related sialidases have been homology modeled to this structure ⁵⁹. Unlike bacterial sialidases, human NEU2 has been demonstrated to have very similar activity on Neu5Ac and Neu5Gc in α2-3 or α2-6 linkage to an underlying galactose residue ⁶⁰. To our knowledge, there are no reports examining the cleavage of of α2-8-linked Neu5Gc by vertebrate exosialidases.
We have taken a biochemical approach to address the question of whether NEU1 and other vertebrate sialidases are able to digest α2-8-linked Neu5Gc. We compare here the chemical and enzymatic breakdown of polymers of α2-8-linked Neu5Ac and Neu5Gc under various conditions, showing that polymers containing Neu5Gc exhibit an unusual resistance to breakdown by sialidases, including NEU1. Using molecular dynamics modeling, we propose a structural mechanism by which Neu5Gc in an α2-8 linkage is less likely to be found in the conformation optimal for enzymatic cleavage. The proposed mechanism is supported by replacement of the glycolyl moiety with a propionyl group, which, while occupying a similar spatial arrangement to the glycolyl residue, negates the stability. Given the critical importance of developmental regulation of polySia in the central nervous system and the likely need for its rapid turnover under specific circumstances, this finding provides the first steps towards an explanation for the evolutionary conservation of the suppression of Neu5Gc synthesis in the vertebrate brain.

EXPERIMENTAL PROCEDURES

DMB-HPLC analysis of sialic acids

Quantification of Sia content and type on acid-hydrolyzed samples of vertebrate tissues, polymers, and disaccharides was done using previously described methods of DMB derivatization at 50°C for 2.5 h followed by HPLC on a Phenomenex C18 column using an isocratic elution in 85% water, 7% methanol, and 8% acetonitrile. Other details are in a companion paper.

For samples in which heated DMB derivatization would cause unwanted breakdown of polySia in solution (partial enzymatic and acid-catalyzed breakdown reactions of polymers and disaccharides), an extended derivatization reaction was used, for 48 h at 4°C, that allowed labeling of all reducing ends in solution without the destruction of polymers. To quantify monomer following such reactions, the derivatization reaction was not quenched and samples were run on a Phenomenex C18 column as described above. To observe polySias, reactions were quenched with the addition of one-fifth volume 1 M NaOH and analyzed on a Dionex DNAPac PA-100 column along a 2–35% gradient of 1M NaNO₃ in water.
Studies of vertebrate tissue Neu5Gc content

Brain samples from mouse, rat, pig, and cow were obtained from Pel-Freez Biologicals, Rogers, AR; chimpanzee brain samples from the Yerkes Primate Center, Atlanta, GA; dolphin liver and brain samples were obtained from the Southwest Fisheries Science Center, La Jolla, CA. Dolphin milk samples were obtained from the US Navy marine mammal program, Point Loma. African and Asian elephant liver samples were obtained from the Zoological Society of San Diego. All samples were rinsed in PBS to remove blood, and 1 g of sample was homogenized using a polytron in 2M acetic acid directly or in methanol chloroform for separate analysis of glycolipids and glycoproteins. Monomeric sialic acids were then released from bound glycans by treatment with 2M acetic acid for 3 h at 80˚C, and quantified by DMB-HPLC.

Cell culture and flow cytometry

SH-SY5Y cells (ATCC) were cultured in DMEM/F12 (Gibco) with 10% human serum (Valley Biomedical Inc.) and supplemented with 2 mM Neu5Ac or Neu5Gc (Nacalai) for 1 week. Cells were lifted with 2 mM EDTA in PBS. Neu5Gc incorporation in the 2mM Neu5Gc supplemented cells was confirmed as 76.5% of total Sias by complete acid hydrolysis of cell pellets in 0.1M trifluoroacetic acid (TFA) for 4 h at 80˚C followed by DMB-HPLC. Neu5Gc was not detected in 2mM Neu5Ac supplemented cells. One aliquot of suspended cells was treated with endoNF (a gift from Rita Gerardy-Schahn; 40 ng/µL) 64 in PBS for 45 min on ice. Cells were stained with 12E3 (10 µg/mL) followed by anti-mouse IgM secondary antibody. Concurrently, a separate aliquot was stained with inactive endoNA-GFP (10 µg/mL) 65. Flow cytometry was conducted on a BD FACSCalibur. Cells were gated on a population of cells positive for polySia and analyzed with FlowJo software (TreeStar).

Preparation of sialic acid dimers
A mixed polymer of Neu5Ac and Neu5Gc (1 mg) was hydrolyzed with 50 mM sodium acetate buffer (pH 4.8) at 50°C for 20 h. The hydrolysate was subjected to DEAE-Sephadex A-25 anion-exchange chromatography (1 x 10 cm, pre-equilibrated in 10 mM Tris-HCl, pH 8.0). After washing the column with 3 volumes of 10 mM Tris-HCl (pH 8.0), oligoSias were eluted with a linear gradient of NaCl (0-0.4M NaCl in 10 mM Tris-HCl, pH 8.0). The diSia fraction was collected, desalted with Sephadex G-25 chromatography (1 x 10 cm, water), and lyophilized. Dried diSia was then dissolved in 200 µL of water. Neu5Acα2-8Neu5Ac, Neu5Acα2-8Neu5Gc, Neu5Gcα2-8Neu5Ac, and Neu5Gcα2-8Neu5Gc were separated and purified by preparative thin-layer chromatography (TLC) as described previously. Each fraction was subjected to the Sephadex G-25 chromatography and lyophilized. The identity of each disaccharide was confirmed by hydrolysis in 0.1M HCl for 1 h at 80°C followed by DMB-HPLC.

Synthesis of sialic acid polymers

De-N-acetylation of polySia: polySia was de-N-acetylated as previously described. N-acetyl-polySia (Lipoxen, 500 mg, 1.6 mmol) was dissolved in a solution of 2 M sodium hydroxide (20 mL) with sodium borohydride (50 mg, 1.3 mmol). The solution was heated at reflux. To achieve over 90 % deacetylation, the polymer was refluxed for 50 h, during which time a white precipitation was seen. The deacetylated polymer intermediate was dialyzed (dialysis tubing molecular weight cut off of 3000 Daltons) with a solution of ammonium carbonate (0.01 M) for 4 h, then water overnight. The solution was dried under vacuum to yield a white powder. The polymer could now be reacetylated, glycolated, or propionylated as described below.

Ac100: Deacetylated polymer (98 mg, 0.04 mmol) was slurried in water (5 mL) with sodium bicarbonate (492 mg, 6 mmol). Acetic anhydride (612 µL, 6 mmol) was added and the solution stirred vigorously for 30 min. The solution was filtered, then dialyzed as described above. The N-acetylation procedure was repeated on the isolated material. DMB-HPLC analysis of the acid hydrolyzed polymer revealed one single peak, which correlated (qualitatively and quantitatively) with a Neu5Ac standard. In addition, ion-exchange polymer profiles of Ac100 were
comparable to the polySia starting material. H\textsuperscript{1}NMR (D\textsubscript{2}O, 500 MHz): Although the CH region of the polymer was broad, due to the polymeric nature of the molecule; characteristic peaks were seen in the spectra, which was in good agreement with a standard of the polySia starting material (ppm). δ = 1.5 (m, 1H, CHH), 1.8 (s, 3H, OCH\textsubscript{3}), 2.6 (m, 1H, CHH) 3.4-4.2 (CH region, 7H).

Gc100: Deacetylated polymer (70 mg, 0.03 mmol) was slurried in water (2 mL) with sodium bicarbonate (176 mg, 2 mmol). Acetoxyacetyl chloride (118 µL, 1.5 mmol) was added and the reaction mixture was stirred vigorously for 30 min. an additional portion of sodium bicarbonate then acetoxyacetyl chloride was added, and the reaction mixture stirred for an additional 30 min. The solution was filtered then dialysed as described above. The product was dissolved in 0.1 M sodium hydroxide solution (5 ml) and heated at 50°C for 30 min. The solution was then dialyzed as described above and then dried under vacuum to yield a white powder. The white powder was analyzed for composition by DMB-HPLC. To insure no free amine residues remained, the product was put through the acetoxyacetylation procedure described above again, then composition analyzed again by DMB-HPLC. If no difference was seen in composition of Sia content then we interpreted this as no free amines remaining. If a difference was seen the procedure was repeated until the composition remained constant. H\textsuperscript{1}NMR (D\textsubscript{2}O, 500 MHz): Characteristic peaks were seen (ppm). δ = 1.5 (m, 1H, CHH), 1.8 (s, 0.1H, OCH\textsubscript{3}), 2.6 (m, 1H, CHH) 3.4-4.2 (CH region, 7H), 4.0 (s, 1.9H, COCH\textsubscript{2}OH).

Mixed polymers, Gc40 and Gc60. Mixed polymers were made and analyzed using a similar procedure as described above for Gc100, using a mixture of acetic anhydride and acetoxyacetyl chloride. Gc40 was generated using a 1:1 premix of acetoxyacetyl chloride and acetic anhydride. Gc60 was generated using a 5:3 premix of acetoxyacetyl chloride and acetic anhydride respectively.

Pr70: polymers containing monomeric residues of N-propionylneuraminic acid (Neu5Pr) were generated in a similar way as described under the mixed polymer method. Pr70 was generated using a 4:1 premix of propionyl chloride and acetic anhydride respectively. The
composition was confirmed as described above, namely checking the composition after running the polymer through additional acetylation reactions.

EndoNF digests of polymers. 200 µg of polymers were subjected to digest with 1 µL EndoNF \(^{64}\) in 100 mM sodium phosphate buffer pH 7.4 for 45 minutes at 37˚C. The reaction was terminated by the addition of an equal volume of ice-cold EtOH, samples were dried under vacuum, and resuspended in water. Digestion to oligomers of length 3-7 Sias was confirmed by extended DMB labeling of reducing Sias followed by HPLC on a DNAPac anion exchange column \(^{63,68}\). To determine concentration of polymers, endoNF-digested samples were then hydrolyzed in 0.1 M TFA at 80˚C for 4 h, and quantified using DMB-HPLC.

Preparation of inactive EndoNA-GFP probe. The pQE31-based construct for inactive EndoNA-eGFP \(^{65}\) was transformed to the M15 [pREP4] (Qiagen) expression strain and expressed as a histidine-tagged fusion protein. The cells were grown in SOB medium (2% tryptone, 0.5% yeast extract, 8.5 mM NaCl, 10 mM MgCl\(_2\), 10 mM MgSO\(_4\)) to OD\(_{600nm}\) ~0.4 and expression induced with 0.1 mM Isopropyl \(\beta\)-D-1-thiogalactopyranoside (IPTG) for 3 h at room temperature. After centrifugation at 3000 x g for 20 min, the pellet was stored at –80˚C. The cells were thawed on ice and resuspended in 50 mM sodium phosphate buffer pH 7.4, 300 mM NaCl, 20 mM imidazole. After adding complete EDTA-free protease inhibitor cocktail (Roche) and lysozyme to a final concentration of 5 mg/ml, the cells were incubated for 30 min on ice and lysed by sonication. Cell wall debris was separated by centrifugation twice at 10,000 x g at 4˚C for 20 min. Ni-NTA resin (Qiagen) equilibrated with the above phosphate buffer was added to the supernatant and stirred at 4˚C for 1 h. The protein–resin complex was washed 5 times as a batch with 8 times (v/v) excess of the buffer and packed into a column. The packed column was washed with the buffer until OD\(_{280}\) ≤0.01. Bound protein was eluted by adding imidazole to 100 mM concentration. The buffer was changed using Amicon Ultra filter devices (Millipore) to 50 mM sodium phosphate pH 7.4, 300 mM NaCl.

Preparation of sialidases
Murine NEU1: NEU1 sialidase was purified from mouse kidney tissue by the affinity chromatography on a concanavalin A-Sepharose column followed by fast protein liquid chromatography gel-filtration on Superose 6 column, as previously described. 69

Rat NEU1: COS-7 cells (2 × 10^6 cells) were transfected with rat NEU1 (rNEU1) plasmid or mock plasmid using the Genejuice transfection reagent (Novagen) and incubated for 48 h. After washing, transfected cells were collected, sonically disrupted in PBS containing protease inhibitors (1 mM PMSF, 1 µg/ml aprotinin, 1 µg/ml of leupeptin, 1 µg/ml pepstein, and 1 mM EDTA) and centrifuged at 600 × g at 4°C for 10 min to remove debris and nucleus. The supernatant was ultracentrifuged at 100,000 × g at 4°C for 1 h. The precipitate was washed with cold PBS containing protease inhibitors and ultracentrifuged. The precipitate, representing the membrane fraction, was dissolved in PBS containing protease inhibitors, and then used as the rNEU1 enzyme fraction.

Murine NEU2 and NEU4. COS-7 cells cultured in 10 cm dishes in Eagles’s minimal essential medium supplemented with 10% (v/v) fetal calf serum (Wisent) and 5% DMSO were transfected with pCTAP-Neu2, and pCTAP-Neu4 plasmids using Lipofectamine LTX (Invitrogen) as described in the manufacturer’s protocol. 48 h post-transfection, cells were washed with PBS and harvested by scraping. Cell pellets from 10 dishes were resuspended in 2 mL lysis buffer from InterPlay TAP Purification Kit (Stratagene) supplemented with 0.1% NP-40 and Sigma protease and phosphatase inhibitor cocktail (P8340, 10 µL per mL of cell suspension). The homogenates were sonicated for 5 sec to solubilize proteins. The suspension was then centrifuged at 13,000 × g for 30 min. The supernatant was first passed through a 0.4 mL of avidin-agarose resin (Sigma A9207), then affinity purification of TAP-tagged Neu2 and Neu4 was performed using streptavidin resin (Stratagene) according to the manufacturer's protocol. Purified enzymes were stabilized in 20% glycerol and stored at −20°C until use.

AUS: Arthrobacter ureafaciens sialidase (AUS) was purchased from EY Laboratories. Enzyme digests of 4-methylumbelliferyl sialic acids. 4-methylumbelliferyl (4MU) Neu5Ac was purchased from Nacalai (Japan). 4MU-Neu5Gc and 4MU-KDN were gifted from Dr. Kimio
Furuhata (Kitasato University). 25 nmol of 4MU-Sia was digested in 90 µL of 50 mM sodium acetate, pH 4.75 (NEU1) or PBS, pH 6.5 (AUS). 20 µL aliquots were quenched at each time point by the addition of 180 µL stop buffer (0.1 M glycine, 25% EtOH, pH 10.7). Fluorescence was read in a 96-well plate reader at Ex 365 nm, Em 450 nm.

Rat NEU1 digests were conducted separately at 37˚C in 50 mM sodium acetate buffer (pH 4.5), 100 µM 4MU-Sia, 0.1 % BSA, 0.1 % Triton X-100 and enzyme fractions (10 µL) in a final volume of 0.1 mL. The reaction was terminated by addition of 1.0 mL of 0.25 M glycine-NaOH (pH 10.4), and the amount of released 4-MU was fluorometrically determined with FluoroMax-3 (Ex 365 nm, Em 448 nm). The rat NEU1 activity was calculated based on the activity for membrane fraction of mock-transfected COS7 cells and natural degradation of the substrate (negative control).

**Enzyme digests of polySia**

1 nmol (Sia equivalent) of endoNF-digested polymers Ac100, Gc40, Gc60, Gc100, and Pr70 were further digested at 37˚C in 50 mM sodium acetate pH 4.75 (NEU1, NEU4), pH 5.5 (NEU2), or PBS pH 6.5 (AUS). For all samples except Pr70, 250 pmol of Neu5Pr were added to each tube as an internal standard. All samples were run in parallel with a no-enzyme and enzyme-only control. Total acid hydrolysis in 0.1 M TFA at 80˚C for 4 h was performed in parallel. At each time-point, aliquots were removed and flash-frozen in liquid nitrogen. All samples were diluted 5-fold in water and derivatized for 48 h at 4˚C using an extended DMB protocol that allows labeling of monomer but does not cause further breakdown of polySia. Analysis was then performed by HPLC on a Phenomenex C18 column, with fluorescence detection.

Total monomeric Neu5Ac, Neu5Gc, and Neu5Pr were quantified by integration of spectra compared to standards for each sample. Each run was normalized according to known concentration of Neu5Pr (this was not done for the Pr70 sample). No-enzyme and enzyme-only control values were subtracted from the corresponding timepoint for each sample. Values at t=0 were also subtracted from values for all time points. Total enzymatic breakdown of Neu5Ac,
Neu5Gc (Gc40, Gc60, Gc100) and Neu5Pr (Pr70 only) was determined as a fraction of total monomer calculated from TFA-hydrolyzed samples. To independently measure Neu5Ac and Neu5Gc breakdown within the Gc40 polymer, released monomer was calculated as a fraction of the total Neu5Ac or Neu5Gc content of the polymer.

Digests of 1 nmol of disaccharides of Neu5Ac and Neu5Gc were conducted and analyzed as described above. Reactions were conducted in triplicate. In parallel, disaccharides were hydrolyzed in 0.1 M HCl for 1 h at 80°C to confirm total possible monomer release.

**Analysis of polymer breakdown under acidic conditions**

*Analysis of monosaccharide release by DMB-HPLC.* Ac100 or Gc100 polymers were dissolved in HCl-KCl, citrate-phosphate, or phosphate based buffers at pH of 1, 2, 4.5, and 7.5 at a concentration of 20 µg/mL (total volume=500 µL). Samples were heated at 80°C, and aliquots (40 µL) taken at hourly intervals. The pH of samples was corrected to pH 1.0 on ice using TFA (13 M) prior to DMB derivatization. Samples were derivatized using the extended DMB reaction followed by HPLC analysis of monomer release.

*Analysis of glycosidic linkage hydrolysis at pH 4.5 and 37 °C using a reducing sugar assay.* Ac100 or Gc100 polymers (0.4 mg) were dissolved in an acetate buffer (pH 4.5, 200 µL) and incubated at 37°C. 20 µL samples were taken at regular time intervals over a 200 h time period. Before analysis samples were stored at –20°C. In parallel, complete hydrolysis was achieved by taking a sample of the polymer in acetate buffer and heating at 80°C with 0.1 M H₂SO₄ for 1 h. Samples were analyzed using a reducing end assay variant of the Mopper-Grindler method. Standards of monomeric Neu5Ac or Neu5Gc (1-20 nmol) were used to quantify the hydrolysis of the polymer linkage. A fractional value was obtained by comparing the molar breakdown to a completely hydrolyzed polymer sample.

Molecular dynamics modeling. Hexamers of α2-8-linked Neu5Ac, Neu5Gc and Neu5Pr were developed using the GLYCAM06 structure library and parameter sets. Charges for Neu5Ac and Neu5Gc were obtained from this parameter set. Charges for the Neu5Pr residue
were calculated from a quantum mechanically optimized Neu5Pr with a reducing O-methyl aglycone (HF/6-31++g**). This model was used to obtain the restrained electrostatic potential (RESP) at HF/6-31++g**. Ensemble averaged charges (EAC) for Neu5Pr were obtained from 100 equally-spaced snapshots from a 10 ns molecular dynamics (MD) simulation using HF/6-31++g** and a RESP weight of 0.0174, consistent with the formalism used in GLYCAM0673. Simulations were performed using the *pmemd* module of AMBER 1175. All simulations followed a standard protocol of minimization (5000 steps conjugate gradient, 5000 steps steepest descent), heating (50 ps from 5 to 300 K), equilibration (300K temperature throughout 100 ps for EAC development and 10 ns for the hexasaccharides) and production (300K throughout 10 ns for EAC and 0.5 µs for the hexasaccharides). The MD simulations were performed using an nPT setup at 1 atm, employing a Berendsen-type76 thermostat and barostat, with temperature and pressure coupling constants of 10 ps and 0.1 ps respectively. Sodium ions were employed to neutralize the system charge and a cubic box of TIP3P waters77 were placed with a minimum of 12 Å between the box edge and solute. Direct non-bonded interactions were truncated at 10 Å while the PME method78 was used to treat long-range electrostatics beyond this cutoff. Scaling factors for 1,4-interactions were not employed and high-frequency motions involving hydrogen atoms were restrained using the SHAKE algorithm79. Initial conformations for all three homopolymers were symmetrized from the dominant conformer identified from a previous study by Yongye et al. 2008 for di- and tri-saccharides of similar systems80. Conformations were extracted and analyzed every 1 ps throughout the MD simulation. Glycosidic torsion angles $\varphi$ (C1–C2–O8′–C8′), $\psi$ (C2–O8′–C8′–H8), $\omega_8$ (H8–C8–C7–H7), and $\omega_7$ (H7–C7–C6–H6) between the non-reducing terminal residues were used to check for convergence and identify the geometries and relative populations. NMR J-couplings for the entire hexasaccharide were consistent with experimental di- and tri-saccharide values presented in Yongye *et al.*80.

**RESULTS**

Relative absence of Neu5Gc is a conserved feature of vertebrate brain.
Prior studies of sialic acids in the vertebrate brain have either failed to find Neu5Gc or reported it to be present only at very low levels\textsuperscript{10,26,81-88}. To confirm and extend these findings, we collected samples of brain tissue from chimpanzee, mouse, rat, cow, pig, and dolphin. Elephant liver and dolphin liver and milk were also analyzed. High performance liquid chromatography (HPLC) was used to characterize the Neu5Ac and Neu5Gc fractions of total Sias in tissues of these mammalian taxa. The combined data, including those from prior publications, are shown in Table 1-1. These data support previous observations that while Neu5Gc is widely and variably distributed in other tissues in various species, it is relatively low or absent from all vertebrate brains. This is corroborated by RT-QPCR data of mouse tissues from our lab (data not shown), as well as published \textit{Cmah} gene expression data from pig that demonstrate highly variable expression in most tissues, with the brain as the only tissue with markedly lowered levels\textsuperscript{35}. It is clear from these results that Neu5Gc and CMAH are otherwise highly variable in their expression, both between species and amongst tissues within an individual species. In striking contrast, Neu5Gc is expressed at very low levels in the brain of all species studied. This suppression of expression is highly conserved across vertebrates, including the frog.

\textbf{Neu5Gc is efficiently incorporated into polysialic acids on neural cells}

Polysialic acid is synthesized in the mammalian brain by two polysialyltransferases, ST8SialII (STX) and ST8SialIV (PST)\textsuperscript{89-92}. Many sialyltransferases studied to date have some preference, but not outright specificity for CMP-Neu5Ac or CMP-Neu5Gc\textsuperscript{39}. With regard to ST8SialII and IV, the situation is less clear. While they have been found to be variably active on a number of unnatural sialic acids\textsuperscript{93-96}, their activity on CMP-Neu5Gc has not been reported.

We have previously shown that cells in culture will take up free Neu5Gc from the culture medium by macropinocytosis, and further, that they will incorporate Neu5Gc into endogenous SGCs\textsuperscript{97}. We exploited this finding to determine whether neural cells are capable of incorporating Neu5Gc into polySia. SH-SY5Y cells, a human neuroblastoma line that expresses both ST8SialII and ST8SialIV and synthesizes PSA-NCAM\textsuperscript{98}, were grown for 1 week in medium supplemented
Figure 3-1. Neuroblastoma cells can incorporate Neu5Gc into endogenous PSA-NCAM.
The human neuroblastoma cell line SH-SY5Y was cultured in the presence of 2 mM Neu5Ac or Neu5Gc for 1 week. An antibody to polymers of Neu5Ac (12E3, left panel) showed reduced staining in the Neu5Gc-supplemented cells. Cells were also stained with a fluorescent polySia-binding probe that does not demonstrate Neu5Ac or Neu5Gc specificity (inactive endoNA-GFP, right panel). Here, staining was comparable between Neu5Ac- and Neu5Gc-supplemented cells. EndoN treatment indicates negative control cells treated with endoneuraminidase from K1F.
with 2 mM Neu5Ac or Neu5Gc. This extended timeline was chosen to allow adequate turnover of the long surface polymers. After 1 week, surface polySia was measured using flow cytometry. Cells were stained with 12E3, an antibody against \( \alpha_2-8 \)-linked Neu5Ac \(^{99,100}\). 12E3 showed reduced binding in cells supplemented with 2mM Neu5Gc (Figure 3-1), indicating either reduced synthesis of polySia or impaired binding of the antibody due to Neu5Gc incorporation.

To distinguish between these two scenarios, cells were further stained with inactive endoNA-GFP \(^{65}\). This molecule is a modified version of an endoneuraminidase from the bacteriophage PK1A. This endoneuraminidase from phage K1F has homology to the endoneuraminidases from PK-series bacteriophages \(^{101}\), and has been shown to cleave polymers of both Neu5Ac and Neu5Gc \(^{66,102}\). In this inactivated form, endoNA-GFP binds but does not cleave polySia. Unlike 12E3, inactive endoNA-GFP demonstrated equal binding in both Neu5Ac- and Neu5Gc-supplemented cells.

Taken together, these data indicate that Neu5Gc can be incorporated into endogenous polySia in SH-SY5Y cells. Thus, the presence of Neu5Gc does not appear to impair the synthesis of polySia. If Neu5Gc were present in the vertebrate brain, therefore, it is likely that it would be incorporated into PSA-NCAM on cell surfaces \textit{in vivo}.

\textbf{No detectable difference in relative rates of release of \( \alpha \)-linked Neu5Ac and Neu5Gc by NEU1}

Many bacterial sialidases have been shown to exhibit a relative preference for \( \alpha_2-3 \) or \( 2-6 \)-linked Neu5Ac over Neu5Gc \(^{54,55}\). To determine the intrinsic preference of sialidases for recognizing Neu5Ac or Neu5Gc independent of linkage, we assessed their ability to digest substrates of 4-methylumbelliferyl (4MU) Neu5Ac, Neu5Gc, and the related nonulosonic acid, deaminoneuraminic acid (KDN). Digests were performed either with the well-characterized \textit{Arthrobacter ureafaciens} sialidase (AUS) or with preparations of murine NEU1 from liver and kidney. Both enzymes demonstrated very similar activity on substrates containing Neu5Ac and
Figure 3-2. Sialidases exhibit no detectable preference for Neu5Ac or Neu5Gc. 1 nmol of substrates of Neu5Ac, Neu5Gc, and KDN α-linked to the fluorescent molecule 4-methylumbelliferone were digested with a murine sialidase (NEU1, A) and a bacterial sialidase (Arthrobacter ureafaciens sialidase, AUS, B). Enzymatic breakdown at each time-point was calculated by subtracting fluorescence in the absence of enzyme to fluorescence in enzyme digest samples.
Neu5Gc (Figure 3-2). Neither sialidase demonstrated appreciable activity on 4MU-KDN. A similar preference was seen with a preparation of NEU1 from rat liver (data not shown).

**Synthesis and characterization of α2-8-linked sialic acid polymers with varying ratios of different N-acyl groups**

Since natural sources of α2-8-linked sialic acid polymers containing Neu5Gc are rare, we adapted current synthetic methods \(^{67,103}\) to generate polymers containing varying levels of Neu5Gc and Neu5Ac (Figure 3-3). Briefly, the N-acetyl groups of polySia were removed under strong basic conditions to reveal free amines. The molecule could then be re-N-acetylated or N-glycolylated, (or a mixture of both obtained) by reacting with acetic anhydride and/or acetoxyacetyl chloride in an aqueous solution of sodium bicarbonate. The final composition was determined using NMR spectroscopy and HPLC analysis. Figure 3-4 shows the polymers used in this study: 100% Neu5Ac (Ac100), 40% Neu5Gc/60% Neu5Ac (Gc40), 60% Neu5Gc/40% Neu5Ac (Gc60), 100% Neu5Gc (Gc100), and 70% Neu5Pr/30% Neu5Ac (Pr70).

HPLC analysis on the DNAPac anion exchange column showed that the process of synthesis causes some breakdown of long polymers, which presumably occurred during the base-catalyzed de-acetylation. To insure consistency in length between synthesized polymers and control Ac100 polymers we employed two methods. Firstly, Ac100 was made in the same way as the Neu5Gc containing polymers, *i.e.* from the de-acetylated polymer backbone (Figure 3-3). Secondly, we treated all polymers with the enzymatically active phage endoneuraminidase endoNF. This cleaves polySia into oligomers of 3-7 residues (EndoNF has previously been shown to digest polymers of Neu5Gc in salmonid eggs) \(^{102}\). We confirmed that endoNF digested all polymers of Neu5Ac and Neu5Gc by anion exchange HPLC (size range, 4-7 residues, data not shown). We next tested the effect of varying levels of Neu5Gc within the polymer backbone on sialidase activity.
Figure 3-3. Synthetic strategy to generate polymers of N-acetyl, N-glycolyl, or N-propionylneuraminic acid (Neu5Ac, Neu5Gc, and Neu5Pr, respectively). A: From commercial colominic acid (bacterial N-acetyl polySia, Lipoxen), the de-acetylated form was accessed by refluxing in an aqueous solution of sodium hydroxide (2M) with sodium borohydride. B: The control polymer of Neu5Ac (Ac100) was made by reacetylation of the deacetylated polymer with acetic anhydride and an aqueous solution of sodium bicarbonate. C: Mixed polymers of Neu5Ac and Neu5Gc (Gc60, Gc40) were generated using a mixture of acetic anhydride and acetoxyacetyl chloride in an aqueous solution of sodium bicarbonate. D: A polymer of Neu5Gc (Gc100) was generated using acetoxyacetyl chloride under the same mildly basic conditions described above. E: The Neu5Pr polymer was synthesized using propionyl chloride in the same way as described above (also see Figure 8).
Figure 3-4. Polymers of Neu5Ac, Neu5Gc, and Neu5Pr used in these studies. The upper panel indicates a trimer of (from left to right) Neu5Ac, Neu5Gc, and Neu5Pr to depict the structure of the individual Sias. Below, the 5 polymer substrates used in these experiments and their respective percentages of each Sia, determined by DMB-HPLC and $^1$H-NMR.

<table>
<thead>
<tr>
<th>ENTRY</th>
<th>Polymer Name</th>
<th>Neu5Ac (%)</th>
<th>Neu5Gc (%)</th>
<th>Neu5Pr (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Ac100</td>
<td>100</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>Gc100</td>
<td>&lt;10</td>
<td>&gt;90</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>Gc60</td>
<td>40</td>
<td>60</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>Gc40</td>
<td>61</td>
<td>39</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>Prop70</td>
<td>31</td>
<td>0</td>
<td>69</td>
</tr>
</tbody>
</table>
Major difference in relative rates of cleavage of terminal sialic acids in α2-8-linked polymers of Neu5Ac and Neu5Gc by various sialidases

To determine the sensitivity of Neu5Ac and Neu5Gc to sialidase digestion in an α2-8 linkage, all Neu5Ac- and Neu5Gc-containing polymers (prepared as described in the previous section) were subjected to degradation by AUS or NEU1. Both enzymes hydrolyze the glycosidic linkage of the terminal residue, releasing a monomeric sialic acid residue. The released Neu5Ac and Neu5Gc monomer was detected by DMB-HPLC.

Although polymers of Neu5Ac are quite sensitive to breakdown by either AUS or NEU1, the polymers containing Neu5Gc exhibit significant resistance to both sialidases (Figure 3-5A). This result was consistent across multiple experimental replicates. The same observation was made with a preparation of NEU1 from rat (data not shown). In mixed polymers, increasing the percentage composition of Neu5Gc within the polymer decreased sialidase sensitivity (Figure 3-5B). Notably, even though the total Sia release from Gc40 is similar to that of Ac100, further analysis indicates that Neu5Ac, and not Neu5Gc, is the dominant product (Figure 3-5C). This suggests that while the enzymes are still able to hydrolyze terminal Neu5Ac residues within the polymers, the presence of a Neu5Gc residue inhibits the hydrolysis rate.

To further explore this idea, and determine the position of the inhibitory Neu5Gc moiety, we digested disaccharides of α2-8-linked Neu5Ac and Neu5Gc. These were Neu5Ac-Neu5Ac, Neu5Gc-Neu5Gc, Neu5Ac-Neu5Gc, and Neu5Gc-Neu5Ac (nomenclature indicates non-reducing Sia followed by reducing Sia). Both AUS and NEU1 were able to digest disaccharides with Neu5Ac at the non-reducing terminus; however, Neu5Gc showed significant resistance to breakdown (Figure 3-5D). The underlying residue appeared to have no effect on the ability of the sialidases to break the glycosidic bond. AUS showed a slight preference for dissacharides with a terminal Neu5Ac and an underlying Neu5Gc, an effect that NEU1 did not exhibit. Taken together, these data indicate that the terminal Neu5Gc residue confers resistance of a polymer to sialidase digestion. Since we see no breakdown of the Neu5Gc-terminated disaccharides and the released
Figure 3-5. Relative rate of hydrolysis of α2-8-linked Sias by sialidases is markedly reduced when Neu5Gc residue is present at terminal positions. Polymer substrates (Figure 4) and disaccharides of Neu5Ac and Neu5Gc were digested with NEU1 (upper panel) and AUS (lower panel). Enzymatic breakdown was measured using DMB-HPLC. A: Upper panel, NEU1 demonstrates much greater activity on Ac100 than on Gc100. Lower panel, total breakdown of mixed Neu5Ac and Neu5Gc polymers by NEU1 at 80 min shows increasing resistance with increasing Neu5Gc fraction. Fractions were calculated relative to total amount of monomer released by complete hydrolysis in 0.1 M TFA. Data shown are from a representative experiment. B: AUS activity on Neu5Ac and Neu5Gc polymers as in part A. C: Fractional release of monomeric Neu5Ac and Neu5Gc from the Gc40 polymer by NEU1 and AUS indicates preferential release of Neu5Ac from this polymer. Fractions are expressed relative to the total amount of the respective monomer (Neu5Ac or Neu5Gc) present in the sample. D: 1 nmol of α2-8-linked disaccharides of Neu5Ac and Neu5Gc (Neu5Ac-Neu5Ac, Neu5Ac-Neu5Gc, Neu5Gc-Neu5Ac, Neu5Gc-Neu5Gc, nomenclature indicates nonreducing followed by reducing sugar) were digested by AUS and NEU1, and released monomer detected by DMB-HPLC. Background subtraction of PBS or acetate buffer controls as well as enzyme-only controls were applied to the values obtained. Neu5Gc at the nonreducing terminus conferred significant resistance to digestion by both AUS and NEU1. (n = 3. Error bars represent standard deviation. ***: p<0.0005, **: p<0.005, analyzed by t-test relative to corresponding Neu5Ac-Neu5Ac sample.) E and F: As in A and B, polymers Ac100 and Gc100 were subjected to digestion by mammalian sialidases NEU2 (E) and NEU4 (F)
monomer from the mixed polymers is predominantly Neu5Ac, this indicates that relatively low percentages of Neu5Gc within a polySia chain would dramatically inhibit hydrolysis.

The other known vertebrate sialidases NEU2, NEU3 and NEU4 are far less prevalent than NEU1 in the brain. NEU3 has a strong preference for gangliosides rather than glycoproteins. We considered the possibility that the NEU2 or NEU4 may also be affected by the presence of Neu5Gc in polySia. Indeed, a recent study has implicated NEU4 as being involved in polySia degradation. Although this work demonstrated the ability of NEU4 to break down polySia, it did not show exclusivity of this enzyme in breaking down polySia in vivo. As NEU1 is present at much higher levels in brain tissue, it is likely that both enzymes are involved in polySia degradation.

To explore the impact of Neu5Gc on these other sialidases the Ac100 and Gc100 substrates were digested with purified preparations of NEU2 and NEU4. These sialidases also exhibit a strong preference for Neu5Ac over Neu5Gc. Notably, the difference in activity on the two polymers was greatest in NEU4. Further studies are needed to determine whether NEU1 or NEU4 is most impacted by the presence of Neu5Gc on polySia in vivo.

**α2-8-linked polymers of Neu5Gc are more sensitive to acid hydrolysis than α2-8-linked Neu5Ac**

Our previous work demonstrated that polymers of α2-8-linked Neu5Ac undergo an intramolecular self-cleavage (hydrolysis of the glycosidic linkage) under mildly acidic conditions, similar to conditions found in the lysosome. To ask if polymers containing Neu5Gc were also more resistant to acidic degradation, we looked at the breakdown of Ac100 and Gc100 under various acidic pH conditions. Using DMB-HPLC analysis we were able to compare monomer release between Ac100 and Gc100 when heated at 80°C and pH 1, 2, 4.5, and 7.5 (Figure 3-6A).
Figure 3-6. Neu5Gc in polySia increases relative susceptibility to acid-catalyzed breakdown. A: To determine differences in acid-catalyzed cleavage, Ac100 and Gc100 were incubated at 80°C at pH of 1, 2, or 4.5. Monomer release was measured by DMB-HPLC. At pH of 2 and 4.5 Gc100 appears to break down faster. B: To determine breakdown of polymer chains to short oligomers, at lysosomal pH of 4.5 and physiological temperature of 37°C, the fraction of linkages hydrolyzed was analyzed using the Moppler-Grindler reducing end assay. Fractions were calculated relative to complete hydrolysis of polymers in H$_2$SO$_4$. In keeping with A, Gc100 appeared to break down faster, with all linkages hydrolyzed by the end point.
While there was no appreciable difference in breakdown at pH 1, at higher pH’s of 2 and 4.5 it appeared that Gc100 was more sensitive to hydrolysis. At pH 7.5, no appreciable breakdown was seen in either polymer (data not shown). We also looked at the breakdown of Ac100 and Gc100 under conditions similar to that in the lysosome (37°C, pH 4.5), this time analyzing the glycosidic linkages hydrolyzed over time using a reducing sugar assay (see methods). As before, we demonstrate that Gc100 was more sensitive to acid hydrolysis (Figure 3-6B). We found that by 125 hours, 100% of the glycosidic linkages in the Gc100 polymer had hydrolyzed, whereas approximately 50% still remained in the Ac100 polymer. This may indicate that polymers of Neu5Ac are more likely to form protective lactone rings at acidic pH 107.

Regardless, taken together, these data suggest that spontaneous acid hydrolysis of polySia at lysosomal pH is not inhibited by the presence of Neu5Gc residues. Indeed, in striking contrast to enzymatic sensitivity, Gc100 polymers in fact appear to be more sensitive to acid catalyzed hydrolysis. However, the time course of polySia breakdown by acid alone is extremely slow, over a period of several days. It is unlikely, therefore, that adequate breakdown of either polymer could occur in the absence of sialidase. Further, this mechanism is not available at the neutral pH of the plasma membrane, and so it cannot impact surface turnover of polySia.

**Molecular modeling shows differences in the 3-dimensional shape of α2-8-linked Neu5Gc, Neu5Ac and Neu5Pr**

To investigate further the mechanism by which Neu5Gc residues inhibit sialidase activity, we modeled α2-8-linked hexamers of Neu5Ac (Ac100) and Neu5Gc (Gc100). While sialidase is active on fragments as small as disaccharides, the use of a hexamer model was selected to introduce the 3-dimensional structural effects of the polymer chain on the enzymatically cleaved non-reducing terminal residue pair. Many studies have indicated that polyNeu5Ac has a helical conformation in solution; although the number of residues per turn appears to vary under different conditions, a recent crystallization of an endoneuraminidase estimated this number to be around
However, studies of 3-dimensional conformation have thus far been limited to di- and trisaccharides and have never been performed on polyNeu5Gc.

Hexamers of Neu5Ac and Neu5Gc were analyzed based on the torsion angles (φ, ψ, ω₆, and ω₇) around the bonds of the final glycosidic linkage (Figure 3-7A). With the exception of the ω₇ rotamer, which did not transition away from -60°, all other torsions frequently transitioned between states, suggesting a converged simulation by 0.5 μs (data not shown). Examining the glycosidic linkage between the terminal residues revealed the apparent absence of one conformer in Neu5Gc hexamers, indicating a constriction within the ω₈ conformational space for Neu5Gc relative to Neu5Ac (Figure 3-7B).

Grouping the conformation states based on the torsion angles (φ, ψ, ω₆, ω₇) produced seven states between the hexamer models (Table 3-1 and Figure 3-7C). The most populated state, centered near (-60°,0°,80°,-60°) (entry a, Table 3-1), was common to both of the hexamers and was observed for about 40% of each simulation, indicating that the overall structure of the polymers is largely similar. However, one of the states centered near (-75°,15°,-60°,-68°) (entry b, Table 3-1) was identified as absent in Neu5Gc; this state represented 24% of the simulation time for Neu5Ac. Another rotamer, centered at (-90°,160°,-170°,-50°) (entry g, Table 3-1) was observed to be exclusive to Neu5Gc, but only for 6% of the simulation.

The accessibility of the b rotamer in hexameric Neu5Ac may represent a partial explanation for its increased susceptibility to enzymatic breakdown. This conformer may be accommodated by the catalytic site better than some of the other states, reducing the entropic penalty to the recognition of Neu5Ac. Alternatively, the b state may allow access to a transition state required for enzymatic activity.

The differences observed between Neu5Gc and Neu5Ac could be the result of increased steric repulsion from a larger substituent or may be an effect of the polar hydroxyl on the 3-dimensional structure. To test this, we additionally modeled a hexamer of N-propionylneuraminic acid (Neu5Pr); the propionyl group is of a similar size to the glycolyl but does not contain the polar group (Figure 3-3B and Table 3-1). Again, the same major conformation (a) was observed.
Table 3-1. States and populations from 0.5 µs molecular dynamics simulations of Sia hexamers. Bond torsion angles across the glycosidic linkage at the non-reducing terminus of hexamers were calculated for Neu5Ac, Neu5Gc, and Neu5Pr as indicated in Figure 3-7B. Seven total populations were grouped and state centers calculated, indicating average bond angles for Neu5Ac, Neu5Gc, and Neu5Pr within each state. Percentages of Neu5Ac, Neu5Gc, and Neu5Pr hexamers are indicated for each conformer.

<table>
<thead>
<tr>
<th>State ID</th>
<th>State centers (\varphi,\psi,\omega_6,\omega_7) (°)</th>
<th>Neu5Ac Popn (%)</th>
<th>Neu5Gc Popn (%)</th>
<th>Neu5Pr Popn (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>-62,1,81,-64 -60,5,77,-66 -64,1,80,-63</td>
<td>42</td>
<td>43</td>
<td>37</td>
</tr>
<tr>
<td>b</td>
<td>-76,16,-58,-60 -77,16,-68,-75</td>
<td>24</td>
<td>0</td>
<td>9</td>
</tr>
<tr>
<td>c</td>
<td>-76,-53,74,-64 -92,-54,67,-66 -90,-56,71,-67</td>
<td>8</td>
<td>24</td>
<td>12</td>
</tr>
<tr>
<td>d</td>
<td>-50,63,-178,-60 -48,56,173,-65 -44,57,-179,-54</td>
<td>7</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>e</td>
<td>-164,-15,77,-57 -166,-16,76,-59 -164,-18,80,-57</td>
<td>7</td>
<td>10</td>
<td>9</td>
</tr>
<tr>
<td>f</td>
<td>-52,49,81,-61 -54,46,78,-63 -53,48,80,-61</td>
<td>6</td>
<td>8</td>
<td>6</td>
</tr>
<tr>
<td>g</td>
<td>-85,159,-169,-49 -89,163,-170,-53</td>
<td>0</td>
<td>6</td>
<td>13</td>
</tr>
</tbody>
</table>
Figure 3-7. Molecular dynamics modeling demonstrates conformational similarities and differences between hexamers of Neu5Ac, Neu5Gc, and Neu5Pr. Hexamers were modeled and analyzed for conformational differences. A: Location of ψ, φ, ω7, and ω8 bond torsion angles across a glycosidic linkage. The arrow indicates the position of the terminal glycosidic linkage of the hexamer whose data is shown in Table 2. B: Plots for the non-reducing terminal glycosidic linkage torsion angles (˚) for ω7 and ω8, demonstrating the absence of one state (b, Table 2) in Neu5Gc hexamers. C: The 3-dimensional structures of the two terminal disaccharides in the most populated rotamers (a) and those with significant differences between Neu5Ac, Neu5Pr and Neu5Gc (b, c, and g) shown in Table 2. Each of the sialic acid models is shown overlaid. D: Molecular model indicating a conformation close to the likely transition state for lactonization. The molecular modeling predicts this conformation (and therefore the transition state of lactonization) is not easily accessible by the polyNeu5Gc molecules.
for about 40% of the simulation. The b state absent in Neu5Gc was observed in Neu5Pr, as in Neu5Ac, but for a shorter percentage of the simulation (9%), suggesting that the hydroxyl in Neu5Gc may in fact inhibit access to this conformation. On the other hand, the minor conformer g, centered at (-90˚, 160˚, -50˚, -170˚) and found at 6% in Neu5Gc, was also observed in Neu5Pr for 13% of the simulation. Although this could indicate a role for the length of the N-substituent in affecting the conformation, it could also suggest that convergence had not yet been achieved after 0.5 µs.

It is important to note that catalytic domain properties of the enzyme itself will also play a critical role in sialidase recognition. Homology modeling of NEU1 performed by Magesh et al. revealed a hydrophobic pocket where the N-acetyl, N-glycolyl or N-propionyl group is expected to reside. However, their homology model for NEU4 shows a more polar pocket, which may account for slight differences in the enzymatic activities of NEU4 and NEU1. Of course, the hydrophobicity of this pocket may not be the only factor at play as the hydroxyl group in the glycolyl moiety may also play a role in perturbing the orientation of the enzymatic machinery. The overall reduction in activity on Neu5Gc by sialidases may therefore be a result of multiple features impacting the enzyme-substrate interaction.

We further considered these models in the context of the enhanced sensitivity of Neu5Gc polymers to acid hydrolysis. It has been previously shown that at low pH polymers of Sia will form lactones that stabilize the molecule to acid breakdown. Modeling of a polyNeu5Ac lactone showed that \( \varphi \) and \( \psi \) angles of \(-45^\circ\), \(-50^\circ\) and \(56^\circ\), \(63^\circ\) respectively would be required within the transition state of lactonization (Figure 3-7D). Interestingly, molecular modeling data predicted that this transition state would fall close to conformation b and d (Table 3-1). As b cannot be achieved in polymers of Neu5Gc, the transition state for lactonization is likely not easily accessible. Therefore, a propensity towards protective lactonization may explain the relative stability of polyNeu5Ac to acid-catalyzed self-cleavage.
Figure 3-8. Replacement of the glycolyl hydroxyl group with a methyl group restores sialidase activity. To determine whether the addition of a glycolyl group in Neu5Gc confers resistance by steric effects, polymers of 70% Neu5Pr (Pr70) were digested with NEU1 (A) and AUS (B) as described in Figure 5. For comparison, we have included enzymatic activity on Ac100 (black line) and Gc100 (grey line). In both cases Pr70 shows enhanced sensitivity to sialidase hydrolysis.
Replacement of N-glycolyl groups with N-propionyl groups restores susceptibility to sialidase cleavage

The addition of an oxygen atom in Neu5Gc from Neu5Ac may conceivably have an array of effects, including steric hindrance of enzymatic binding, increased solvation of the polymer, or altered interaction with the sialidase binding pocket. To address these possibilities, polymers containing Neu5Pr and Neu5Ac residues were synthesized as previously discussed (Figure 3-3E and Figure 3-4 entry 5). We used a polymer with an average composition of 70% Neu5Pr and 30% Neu5Ac (Pr70). This is similar in composition ratios to the mixed polymer Gc60, which we previously showed had an inhibitory effect on sialidase activity (Figure 3-5, A & B). In striking contrast to Gc60, Pr70 polymers show susceptibility to breakdown by the sialidases NEU1 and AUS (Figure 3-8). In both cases it appears that the presence of N-propionyl actually increased sensitivity to sialidase hydrolysis. This indicates that the resistance of Neu5Gc to NEU1 and AUS is not mediated by steric effects. Further, as the b rotamer was found in equal fractions in the modeling of both the Neu5Ac and Neu5Pr hexamers (Table 3-1), this may indicate that conformational restriction is not a critical determinant of sialidase activity. This finding lends support to the idea that sialidase activity on α2-8-linked polymers may depend strongly on the effects of substrate binding and solvation, as the aliphatic pocket present at the binding site of NEU1 and AUS favors the longer propionyl group of Neu5Pr over the shorter acetyl group of Neu5Ac.

DISCUSSION

It is well established that sialidases exhibit relative preferences for the type of sialic acid or linkage that they are able to digest. However, most studies to date have focused on α2-3 or 2-6-linked Sias, and the differences found in vertebrate sialidases have been small. Here we demonstrate a remarkable intrinsic stability of terminal α2-8-linked Neu5Gc residues to enzymatic digestion. This stability extends beyond simple enzymatic selectivity for the terminal Sia, to a molecular stability in the presence of a terminal glycolyl moiety in combination with an α2-8-
linkage. Molecular dynamics modeling suggests that conformational differences of $\alpha_{2-8}$ Neu5Gc, as well as a potentially decreased interaction of the glycolyl moiety with a sialidase binding pocket, may contribute to this finding. The significant inhibition of enzymatic activity by Neu5Gc-terminated disaccharides and the relatively low release of Neu5Gc in mixed polymers indicate that even a low percentage of Neu5Gc in polySia would dramatically inhibit its enzymatic hydrolysis by exosialidases. Since there are no endosialidases in vertebrates, this finding is expected to affect tissues expressing $\alpha_{2-8}$ polySia. Although further work is clearly needed to demonstrate a clear link, this finding provides a potential mechanism by which even small amounts of Neu5Gc may exert a detrimental effect in the vertebrate brain. We speculate that the resistance of $\alpha_{2-8}$ linked Neu5Gc to sialidase breakdown may therefore underlie the relative absence of Neu5Gc from all vertebrate brains studied to date.

PolySia is a strong candidate for this neural-specific effect, being highly and widely expressed in neural development and playing critical roles in growth and plasticity. It is expressed during development in some extraneural tissues, but in much lower amounts, and so would not be expected to show major toxicity in other, Neu5Gc-containing tissues. Further, the $\alpha_{2-8}$ linkage that affords this resistance is enriched in brain, not only in polySia and on gangliosides, but also on short di- and oligosialosyl SGC epitopes that are much rarer in other tissues. The large number of Sias present in each polymer must be broken down one by one by mammalian sialidases, which as exoglycosidases can only cleave a terminal sialic acid. This allows a very low fraction of Neu5Gc to render an entire chain relatively resistant to breakdown, due to the presence of a single terminal Neu5Gc residue.

We propose a model in which Neu5Gc can be incorporated into polySia but then renders the molecule relatively indigestible by vertebrate sialidases. The inability to rapidly desialylate NCAM may then result in widespread inappropriate polySia distribution in the brain. We show here that Neu5Gc can be incorporated into endogenous polySia by polysialyltransferases. Such polymers containing Neu5Gc are relatively resistant to breakdown by NEU1, NEU2, and NEU4. Previous work has indicated that polymers of Neu5Ac do undergo cleavage by both
extramolecular and intramolecular protons at mild acid pH \(^{51}\); however, it is a slow process that occurs over days via short oligomer intermediates, without much breakdown to monomer. Although we find that polymers of Neu5Gc are actually relatively more susceptible to acid-catalyzed breakdown, it still occurs over a very long time course. In embryonic development, when PSA-NCAM is highly enriched in brain, lysosomal acid hydrolysis alone may be insufficient to degrade the large quantities of polySia present. Further, this process would additionally be ineffective for the rapid desialylation of NCAM that may be necessary at the cell surface.

The enzymatic resistance of Neu5Gc-containing polySia may of course have effects in tissues outside the vertebrate brain. While polySia is relatively rare in vertebrate tissues outside the developing brain, it is found in the capsule of certain bacteria such as *E. coli* K1 and *N. meningitidis* \(^{43}\). However, no bacterium has yet been shown to be capable of Neu5Gc synthesis, and thus these capsules contain only Neu5Ac. There is a single report of Neu5Gc present in \(\alpha2-8\)-linked polySia is in the eggs of salmonid fish \(^{113}\). While we are not aware of any organism that digests these glycoproteins, we can hypothesize that eggs that contain Neu5Gc may be more resistant to attack by sialidase-producing pathogens in the wild.

Further studies of Neu5Gc and polySia *in vivo* will of course be necessary to determine whether the observed enzymatic resistance is, in fact, sufficient to exert the proposed detrimental effect on vertebrate brain development. Regardless, this finding is an unusual biochemical consequence of exchanging Neu5Ac for Neu5Gc in a biological system, showing how a single oxygen atom can have dramatic consequences for the biology of a tissue-specific polysaccharide.


**REFERENCES**


CHAPTER 4

The Overexpression of Neu5Gc in the Mammalian Brain
ABSTRACT

The sialic acids (Sias) N-acetylneuraminic acid (Neu5Ac) and N-glycolylneuraminic acid (Neu5Gc) are sugars commonly found at the terminal ends of glycosylated structures in vertebrates. Neu5Gc makes up a highly variable proportion of total sialic acid in most tissues when compared across species. The sole exception to this is the vertebrate brain, which, despite containing the highest concentration of Neu5Ac of any vertebrate tissue, also contains strikingly low levels of Neu5Gc (<3% in all vertebrate species tested to date). The expression of Neu5Gc’s biosynthetic enzyme, CMP-Neu5Ac hydroxylase (CMAH), is correspondingly very low in the brain. To our knowledge, no other molecule demonstrates this distribution, being highly expressed throughout most tissues but exhibiting evolutionarily conserved suppression of expression in the brain alone. This unusual distribution suggests that Neu5Gc presence may exert a detrimental effect in brain. We hypothesized that the presence of Neu5Gc may interfere with the degradation of brain glycoconjugates, and recently proposed a candidate mechanism by which this effect might occur. The important brain glycan polysialic acid (polySia, commonly found on the neural cell adhesion molecule) is degraded less effectively by vertebrate sialidase in the presence of Neu5Gc. In the current study, we further characterize the distribution of Neu5Gc-containing glycans within the brains of mice and several other mammalian species, and explore the effects of its overexpression. Several murine models of Neu5Gc overexpression have been developed in order to characterize the effects on the brain. Initial efforts using CMV- and NSE-driven transgenes resulted in high embryonic lethality, preventing the establishment of a transgenic line. The use of the lox-Cre system to create inducible, tissue-specific expression of Cmah has allowed us to circumvent this problem. Surprisingly, however, lethality was not observed with the inducible transgene. The resulting viable mice express high levels of brain Neu5Gc, with diffuse expression throughout the parenchyma of the brain. Further work is underway to characterize the phenotypes of these mice. These models will provide an important foundation for further understanding the unusual evolutionarily universal exclusion of Neu5Gc from the vertebrate brain.
INTRODUCTION

Most vertebrate tissues can contain either of the two common vertebrate sialic acids: \(N\)-acetylneuraminic acid (Neu5Ac) and \(N\)-glycolyneuraminic acid (Neu5Gc). The fractions of each vary widely both among tissues and between species. The sole exception to this pattern is found in the vertebrate brain, where only trace amounts of Neu5Gc are present in all species tested to date. (For further detail, see Table 2-1.)

The dynamic variability seen in non-neural tissues is the expected pattern for Neu5Ac and Neu5Gc expression. The two molecules are extremely similar, differing by only a single oxygen atom. In fact, many proteins related to sialic acid metabolism are able to utilize either sialic acid \(^1\). There is, therefore, no \textit{a priori} reason to assume a requirement for one or the other sialic acid in a given tissue.

There are some cases in which the small difference between Neu5Ac and Neu5Gc becomes relevant within a given tissue. For example, specific pathogens can cause a tissue to preferentially express one type of sialic acid—indeed, the binding of ancestral \textit{Plasmodium} strains to Neu5Gc has been proposed to have caused the complete loss of functional \textit{Cmah} in the human lineage \(^2\)-\(^4\). Although they are less well understood, there are also cases in which the type of sialic acid expressed has intrinsic effects on a tissue. For example, the absence of Neu5Gc from muscle tissue is required for the profound phenotypes found in Duchenne and limb-girdle muscular dystrophies \(^5\),\(^6\). However, even these examples have not produced sufficient selective pressure to cause the phenomenon seen in brain, where a single tissue expresses solely Neu5Ac or Neu5Gc across all vertebrates.

For decades, it has been noted with curiosity that Neu5Gc is either absent or present in extremely low quantities in brain tissue. This correlates with very low levels of the \textit{CMAH} transcript in species including mouse, pig, and human, suggesting that regulation at the level of mRNA is responsible for the resulting low expression of Neu5Gc \(^7\)-\(^9\). The suppression of Neu5Gc expression, an evolutionarily conserved feature of the brain alone, suggests that its presence may have a detrimental effect in this tissue.
To date, minimal work has addressed this question. Our laboratory has recently published a finding that Neu5Gc in α2-8 linkage demonstrates remarkable intrinsic resistance to breakdown by vertebrate and bacterial sialidases. This finding is particularly notable considering that α2-8 linkages are enriched in the brain on a number of sialoglycoconjugates. One of these, polysialic acid (polySia), plays a number of important roles in the brain throughout life, including migration and neurite outgrowth during development, and plasticity and repair in adult life (for further discussion, see Chapter 2, or 12). Recent work has implicated rapid turnover of polySia on microglia to be an important response to LPS stimulation (unpublished, laboratory of Ken Kitajima), a process that is likely to be impeded if a terminal Neu5Gc confers resistance to the underlying chain. We therefore hypothesized that the absence of Neu5Gc from the vertebrate brain may be necessary to allow this rapid turnover to occur.

In order to properly address this hypothesis and to develop new possibilities, it is necessary to determine what effect, if any, the presence of Neu5Gc may have on brain tissue in vivo. To address this question, our lab has made several efforts to overexpress Neu5Gc in the murine brain and to characterize the resulting phenotypes. Initial transgenic efforts supported prenatal lethality, and we therefore pursued a mouse line with Cre-inducible expression of CMAH and Neu5Gc. Such mice express high percentages of neural Neu5Gc; however, surprisingly, they are viable and grossly normal, unlike earlier transgenics. Efforts are ongoing to characterize the phenotypes of these transgenic animals. This model will be critical in furthering our understanding of the reasons underlying the conserved exclusion of Neu5Gc from the vertebrate brain.

**EXPERIMENTAL PROCEDURES**

**Creation of mouse lines**

Linearized insertions of CMV-\textit{Cmah} and NSE-\textit{Cmah} were injected into the male pronucleus of fertilized mouse embryos. The resulting embryos were implanted into pseudopregnant females. After birth, mice were genotyped by PCR of genomic DNA taken from tail or toe clipping.
Mice expressing the inducible CFE-CMAH construct were also created via pronuclear injection. T-CFE-CMAH was created through site-specific transgenesis using the TARGATT technology of Applied StemCell (Menlo Park, CA). Mice were genotyped for the CFE-CMAH insertion by using a fluorescent headlamp to observe GFP expression. Cre genotype was determined by PCR.

Synapsin-Cre mice came from the laboratory of Jamey Marth, Sanford-Burnham Medical Research Institute. Nestin-Cre mice were from the laboratory of Joseph Gleeson, University of California, San Diego. Ella-Cre mice, originally from Jackson labs, were from the laboratory of Jeffrey Esko, University of California, San Diego. NSE-Cre mice were purchased from Jackson Labs.

All lines were maintained in C57BL/6 background, with the exception of T-CFE-CMAH, which was initially created in a FVB background and maintained as a hybrid with C57BL/6.

**Analysis of CMV-Cmah transgenic embryos**

After pronuclear injection of the CMV-Cmah construct, three pregnant dams were sacrificed and embryos collected at embryonic day 13.5. The embryos were manually scored as normal or abnormal, with abnormal categories including small size, absence of eye development, or decomposition of embryos. Embryonic tissue was genotyped for presence or absence of the transgene by PCR.

**Mouse tissue collection**

Mice were euthanized by isoflurane inhalation, the thoracic cavity was opened, and they were perfused through the left ventricle with 30 mL of ice-cold PBS. Tissues for analysis were removed intact. If brains were to be collected for staining of tissue sections, mice were further perfused with 4% paraformaldehyde, and the brain immersed in 30% sucrose at 4°C until the tissue sank.
**RT-qPCR**

RNA was extracted from mouse tissues by homogenization of whole tissue in TRIzol (Invitrogen) and extraction according to the manufacturer’s instructions. The first-strand product of reverse transcriptase was synthesized using the AffinityScript qPCR cDNA synthesis kit (Agilent) using random primers. The resulting cDNA was subjected to quantitative PCR in 25 µl total reaction volume using SYBRGreen master mix (Qiagen) on a BioRad CFX Connect real-time PCR detection system. All samples were run in triplicate.

Primers used for Cmah were AGACATTCCATTTATGGCG (forward) and CACCTCCTGCGAAATCACTCA (reverse). Control primers against the housekeeping gene Gapdh were AGGTGGTGTGAACCGGATTG (forward) and TTAGACCATGAGTTGAGGTCA (reverse).

**Tissue immunohistochemistry**

Brain tissue from mouse, rat, rabbit, cow, and pig was collected from animals used for other studies. Tissue was rapidly frozen in O.C.T. compound (Tissue-Tek). 10 µm sections were fixed in 2% paraformaldehyde. Three-step staining was performed using a polyclonal antibody to Neu5Gc (Sialix) or with a corresponding IgY isotype control (Sialix), and developed using a horseradish-peroxidase tertiary reagent, followed by counterstain with hematoxylin. All staining was performed in a humid chamber.

**DMB-HPLC analysis for total sialic acid content**

Whole murine tissues were homogenized by sonication in lysis buffer (20 mM Tris-HCl, pH 8.0, 150 mM NaCl, 5 mM EDTA, 2% Triton-X100, with 1X protease inhibitors added just before use). Aliquots of lysate were hydrolyzed in 2M acetic acid for 3 hours at 80˚C and then derivatized with the addition of an equal volume of DMB reagent (7 mM DMB, 1.4 M acetic acid, 0.75 M β-mercaptoethanol, 18 mM sodium hydrosulfite) and incubation in the dark at 50˚C for 2.5 h. Low molecular weight material was collected by filtration through a 10 kDa centrifugal filter unit.
(Millipore). The derivatized samples were subjected to HPLC on a Hitachi LaChrom Elite, using a Phenomenex C18 column. The method used was isocratic in 85% water, 7% methanol, and 8% acetonitrile, with excitation at 373 nm and emission at 448 nm.

To measure free versus bound sialic acid, a control aliquot of lysate in 2M acetic acid was stored at 4°C for 3 h. Both acid-hydrolyzed and control aliquots were then subjected to a modified extended DMB derivatization reaction. Originally developed for labeling of oligo- and polySia with minimal cleavage within the chains, this method is also useful for labeling free Sia while minimizing hydrolysis and labeling of bound Sias. Briefly, an equal volume of 2X extended DMB reagent (14 mM DMB, 0.5M Na₂S₂O₄, 1 M 2-mercaptoethanol, 40 mM trifluoroacetic acid) was added to each sample, and they were derivatized in the dark at 4°C for 48 h. The original method describes the addition of NaOH to eliminate lactones; this step was eliminated because it was found to destroy the detectable labeled monomer. Samples were filtered and analyzed by HPLC as described above.

To separate glycolipid- and glycoprotein-bound Sias, lipids were extracted by adding 20 volumes of 1:1 chloroform:methanol to the sample. Tubes were mixed thoroughly, centrifuged at 2000 rpm for 5-10 min, and the supernatant collected as the glycolipid fraction. The remaining sample was saved as the glycoprotein fraction. Both fraction were dried under a stream of nitrogen gas, resuspended in water, and then subjected to acid hydrolysis and DMB-HPLC as described above.

RESULTS

The distribution of Neu5Gc in the mammalian brain

Although Neu5Gc is known to be expressed at very low levels in the vertebrate brain, the precise localization of the trace amounts of Neu5Gc present has not been explored. To examine the specific localization of Neu5Gc within the vertebrate brain, frozen sections of whole brain tissue from mouse, rat, rabbit, cow, and pig were stained with a polyclonal antibody recognizing Neu5Gc glycans. As shown in Figure 4-1A, this staining has a punctate quality and appears to be
Figure 4-1. The distribution of Neu5Gc in the mammalian brain. **A:** Frozen cortical sections from various mammalian species were stained with a polyclonal antibody to Neu5Gc-containing glycans (brown), and counterstained with hematoxylin (purple). In all cases, the distribution is punctate and does not extend throughout most of the brain parenchyma. **B:** Immunofluorescence of mouse brain sections stained for Neu5Gc along with the cell-specific markers CD31 (endothelia), F480 (microglia), and NeuN (neurons). Neu5Gc colocalizes completely with CD31, but not at all with F480 or NeuN. This work was done by former graduate student Pam Tangvoranuntakul and Nissi Varki. **C:** DMB-HPLC is used to quantify total sialic acid content after acid hydrolysis of mouse whole brain lysate. Neu5Gc is present at approximately 1% of total sialic acids. **D:** RT-qPCR of Cmah in wild-type mouse tissues demonstrates approximately a 3000-fold difference in expression between brain and liver. Error bars represent standard deviation of triplicate samples.
confined to certain cells rather than expressed throughout the brain parenchyma. Interestingly, although the pattern in mouse and pig appears to be confined to endothelia, other Neu5Gc-positive cell types appear in other species. Despite these slight variations, all species tested demonstrate very low Neu5Gc expression in the brain parenchyma.

To further characterize the distribution of Neu5Gc in murine brain, sections of brain tissue were stained with an antibody for Neu5Gc in conjunction with other cell surface markers: CD31 (endothelia), F480 (microglia), and NeuN (neuronal) (Figure 4-1B). Notably, Neu5Gc expression appears to correlate completely with expression of CD31, suggesting that—at least in the mouse—the trace amounts of Neu5Gc remaining in the vertebrate brain are expressed solely in endothelial cells and do not extend across the blood brain barrier.

Lysate of whole brain tissue of wild-type mice was further subjected to total acid hydrolysis in 2M acetic acid, which hydrolyzes glycosidic linkages, followed by labeling of alpha-keto acids with the fluorescent molecule 1,2-diamino-4,5-methylenedioxybenzene (DMB). This method allows quantitative measurement of Neu5Ac and Neu5Gc by HPLC. Neu5Gc is quantified as approximately 0.6 pmol per µg protein in brain, compared to approximately 65 pmol of Neu5Ac per µg protein. Neu5Gc thus comprises 1% of total sialic acids in brain tissue (Figure 4-1C). This corresponds to published data indicating very low expression of Neu5Gc in the murine brain (Table 1-1).

Lastly, we sought to determine whether this observed low expression of Neu5Gc is caused by mRNA-level repression of its synthetic enzyme Cmah. We compared the level of Cmah mRNA in wild type mouse brain to that in liver by RT-qPCR. We found approximately a 3000-fold lower level of Cmah in the wild-type mouse brain than in the liver of the same mouse (Figure 4-1D). This indicates that Cmah is in fact regulated at the mRNA level, either by transcriptional regulation or by degradation of the transcript itself. It is important to note, however, that this finding does not rule out additional mechanisms by which Neu5Gc is regulated.

Transgenic mouse models suggest Cmah overexpression is toxic to mouse embryos
To explore the phenotypes resulting from high expression of Neu5Gc in the vertebrate nervous system, we pursued the transgenic overexpression of Cmah (and, therefore, Neu5Gc) in a mouse model. Initially, a cytomegalovirus (CMV) promoter was used to drive Cmah, forming a construct that should produce high expression of Neu5Gc throughout the whole body of a transgenic animal. Approximately 1000 copies of the linearized construct were injected into the male pronucleus of fertilized mouse eggs, which were then implanted into pseudopregnant dams. This procedure is expected to produce about 15-20% transgene-positive offspring from all injected embryos. Surprisingly, only four transgenic founders were born out of 105 injected embryos (3.8%). Of these, none were found to express Cmah or Neu5Gc in the brain when tested by northern blot and immunohistochemistry, respectively (Table 4-1). This finding was highly suggestive of prenatal lethality, with only pups that were able to silence the inserted transgene surviving until birth.

Three pregnant dams carrying litters of CMV-Cmah injected embryos were sacrificed at embryonic day 13.5 in order to determine why transgenic founders were not surviving until birth. Embryos at this stage were scored as being grossly normal or abnormal, i.e. having characteristics such as small size, absence of eyes, or degeneration of the embryo. They were also all genotyped for genomic presence of the transgene. Among the embryos that genotyped as wild type, only one out of nine was scored as abnormal. However, among eight embryos genotyped as containing the CMV-Cmah transgene, a remarkable seven were found to be abnormal (Figure 4-2) (OR=28.0, 95% CI 2.1 – 379.3). It is important to consider that this experiment was done only once and its small numbers make interpretation difficult. Additionally, there may have been embryos lost earlier in development that were not included in this analysis. However, the result obtained even with this small sample is strikingly suggestive of abnormalities in embryonic development resulting from the overexpression of Cmah.

CMV-driven expression is systemic, not neural-specific, and so another attempt at generating a transgenic line was made. This study instead used the promoter of neuron-specific enolase (NSE, also referred to as ENO2) to drive Cmah. This promoter is primarily neuronal, but
Table 4-1. Summary of initial efforts to overexpress transgenic Cmah in mouse brain. CMV: cytomegalovirus, NSE: neuron-specific enolase (ENO2), TG: transgene. These efforts were undertaken by former lab members Hiromu Takematsu and Hsun-Hua Chou.

<table>
<thead>
<tr>
<th>promoter</th>
<th>Cmah cDNA</th>
<th># live births</th>
<th># TG + founders</th>
<th>% TG + founders</th>
<th>Brain Cmah mRNA?</th>
<th>Brain Neu5Gc?</th>
</tr>
</thead>
<tbody>
<tr>
<td>CMV</td>
<td>wild type</td>
<td>105</td>
<td>4</td>
<td>3.8</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>NSE</td>
<td>wild type</td>
<td>72</td>
<td>2</td>
<td>2.8</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>NSE</td>
<td>mutant (frameshift)</td>
<td>57</td>
<td>9</td>
<td>15.8</td>
<td>yes (5)</td>
<td>no</td>
</tr>
</tbody>
</table>
Figure 4-2. Transgenic Cmah driven by CMV promoter causes embryonic abnormalities. **A:** In work done by former postdoctoral fellow Hiromu Takematsu, embryos were collected at E13.5 and observed for normal or abnormal development. Normal development is shown at left. Observed abnormal phenotypes included decomposing embryos (center) and embryos without proper eye development (right). **B:** The percentage of pups of each genotype (wild type or transgenic) that demonstrated a normal or abnormal phenotype at E13.5. Transgenic embryos were much more likely to be abnormal (OR = 28.0).
is thought to be less than optimally specific to this cell type. Again, the number of founder animals was extremely low, only two out of 72 live births, with no Cmah mRNA or Neu5Gc detected in brain tissue. A frameshift mutant of the same NSE-Cmah construct was also tested, in which a single base-pair insertion rendered the resulting protein non-functional. This mutant construct produced a much higher number of transgenic founder mice (nine out of 57, 15.8%), and transcribed mRNA was in fact detected in the brains of positive mice (Table 4-1). This indicates that the mRNA itself is neither unstable nor toxic to the developing embryo. Again, these findings are highly suggestive of lethality caused by overexpression of Cmah, this time more specifically in the mouse brain.

**CFE-CMAH construct allows inducible expression of Cmah in vertebrate brain**

Since constitutive expression of transgenic Cmah proved lethal to developing embryos, we developed an inducible construct that makes use of the lox-Cre system to express Cmah in the presence of a Cre recombinase (Figure 4-3A). This construct contains a floxed GFP that is constitutively expressed, allowing easy genotyping of the mice by headlamp; the presence of a Cre recombinase causes excision of GFP and expression of Cmah.

Prior to injection, we confirmed that the construct was functional in HEK 293T cells, a human embryonic kidney line that does not express endogenous Neu5Gc. Cells were transfected with a plasmid containing CFE-CMAH in the presence or absence of an additional plasmid containing a Cre recombinase, and analyzed by flow cytometry for the presence of GFP and surface Neu5Gc using a polyclonal antibody for Neu5Gc-containing glycans. As expected, surface Neu5Gc was found on these cells only in the presence of both CFE-CMAH and Cre (Figure 4-3B).

Several founder lines were created using the CFE-CMAH construct. The linearized insertion of CFE-CMAH was injected into the male pronucleus of fertilized mouse eggs, resulting in three initial lines (CFE-CMAH-A, -B, and -C). Additionally, a fourth line, T-CFE-CMAH was produced with the company Applied StemCell. This commercial method uses an integrase to
Figure 4-3. The CFE-CMAH construct used for inducible overexpression of Cmah. A: A cartoon of the CFE-CMAH insertion. The floxed GFP is normally expressed under the CAG promoter; in the presence of Cre recombinase, it is excised, allowing expression of Cmah. Note that the HS4 insulators, which are intended to reduce the chances that the insertion will be silenced, are not present in the T-CFE-CMAH line. B: To confirm the functionality of the CFE-CMAH insertion, 293T cells were doubly transfected with pCFE-CMAH and pCAG-NCre, and expression of Neu5Gc and GFP were analyzed by flow cytometry. As expected, cells transfected with pCFE-CMAH express GFP, but cells positive for Neu5Gc are found only in the presence of both constructs.
target single-copy insertion to a specific locus, in this case \( H11^{14} \). Unlike the other founders, TCFE-CMAH lacks the flanking HS4 insulator sequences.

**CFE-CMAH lines are able to express high levels of \( Cmah \) and Neu5Gc in the brain**

Initially, heterozygous CFE-CMAH founder lines were crossed with mice heterozygous for Synapsin-Cre, which is expressed in neurons at E13.5. If embryos of all genotypes survive until birth, these crosses are expected to produce transgenic expression of \( Cmah \) in 25% of offspring, with the remainder expressing only CFE-CMAH, only Synapsin-Cre, or neither transgene. The ratios of genotypes in the offspring indicated Mendelian inheritance, indicating that profound embryonic lethality was in fact not present in these animals (data not shown). All surviving mice appeared grossly identical, regardless of genotype, and were fertile.

To confirm \( Cmah \) and Neu5Gc expression in brain tissue, mice positive for both Cre and CFE-CMAH were sacrificed at approximately 8 weeks of age, and brain tissue was analyzed. Quantitative RT-PCR analysis for \( Cmah \) from CFE-CMAH-A/Synapsin-Cre double positive mice confirmed very high expression in brain tissue, as predicted. In fact, this expression was much higher even than that seen in wild type liver, a tissue that naturally expresses approximately 80% Neu5Gc in mouse (Figure 4-4A). Quantitative DMB-HPLC analysis of total brain sialic acids revealed that, although there was substantial variability in the expression of Neu5Gc across founders, overall the Neu5Gc fraction was quite high, ranging from about 25% to as high as 90% (Figure 4-4B).

To verify that this high neural Neu5Gc was in fact successfully incorporated into brain glycoconjugates, we collected the brain of one CFE-CMAH-A/Synapsin-Cre individual and examined the fraction of Neu5Gc in free sialic acids versus in glycoprotein- or glycolipid-bound sialic acids. Two aliquots of brain homogenate were suspended in 2 M acetic acid. One aliquot was hydrolyzed at 80°C for 3 h to cleave the glycosidic linkages of bound Sias, to give the total Sia concentration. The second was stored at 4°C for this time, leaving the linkages generally intact, thus yielding only free Sia to analysis. The two samples were then DMB-derivatized using
Figure 4-4. CFE-CMAH lines are able to express Cmah and Neu5Gc in the brain. A: Cmah expression was analyzed by RT-qPCR in the brain of a CFE-CMAH-A x Synapsin-Cre mouse, as well as in brain and liver of wild type mouse. The transgenic expression of Cmah is extremely high. Error bars represent standard deviation of triplicate samples. B: Neu5Gc percentage of total brain Sias was determined by acid hydrolysis and DMB-HPLC analysis of mice from each of the four CFE-CMAH lines crossed with Synapsin-Cre. The resulting percentage ranges from 26.2% (CFE-CMAH-A) to 91.4% (CFE-CMAH-C). C: Whole brain lysate from CFE-CMAH-A x Synapsin-Cre mouse was subjected to analysis for total, free, bound, glycoprotein- and glycolipid-bound Sia. Free Sia content was determined by DMB-derivatizing non-acid hydrolyzed material, and bound Sia was calculated by subtracting the free Sia from the total Sia. Glycoprotein- and glycolipid-bound Sias were separated by using methanol:chloroform extraction to remove the lipids from the initial lysate. D: CFE-CMAH-C mice were crossed with Synapsin-Cre, Nestin-Cre, or NSE-Cre lines, and total Sias quantified by DMB-HPLC. All crosses produced high Neu5Gc expression in brain tissue. D: 10 µm frozen sections from the brains of CFE-CMAH-C x Nestin-Cre offspring were stained with a polyclonal IgY antibody for Neu5Gc-containing glycans. Unlike the punctate endothelial pattern of expression seen in wild type mice, the double-transgenic offspring demonstrate diffuse expression of Neu5Gc throughout the parenchyma of the brain, with lighter staining found in the vessels. Control IgY isotype control is presented for comparison. E: As in D, frozen brain sections were stained with an antibody to Neu5Gc-containing glycans, and additionally with one to CD31, a marker for endothelial cells, to demonstrate normal distribution of blood vessels in the transgenic mouse brain.
a modified protocol that limits further acid-catalyzed hydrolysis during the derivatization itself. HPLC analysis revealed similar total bound Sia compared to wild type littermates (data not shown). Additionally, the Neu5Gc fraction is almost identical in both free and bound Sia fractions (Figure 4-4C). This confirms that the overexpressed Neu5Gc does not remain sequestered within the cell, but these transgenic mice are able to utilize it in sialoglycoconjugates.

The homogenate was also subjected to methanol:chloroform extraction to separate glycoproteins from glycolipids, and the resulting fractions analyzed again by DMB-HPLC. These studies demonstrated that Neu5Gc fraction of total sialic acids is in fact very similar across glycolipids and glycoproteins, as well as in free and bound states (Figure 4-4C). This further indicates that the neural cells of transgenic mice do not discriminate in their utilization of the forms of sialic acid present, but effectively use Neu5Gc and Neu5Ac similarly in their sialoglycoconjugates. Importantly, taken together with the finding that the Neu5Gc fraction is similar across free and bound forms of sialic acid, this finding demonstrates that the presence of Neu5Gc does not cause any gross impairment of glycan synthesis in the brain.

We then used immunohistochemical analysis to better understand the distribution of induced Neu5Gc expression throughout the brain. Frozen brain tissue sections were stained with a polyclonal IgY antibody to Neu5Gc-containing glycans. In all CFE-CMAH/Cre individuals examined, Neu5Gc was expressed diffusely throughout the parenchyma of the brain (Figure 4-4D). Remarkably, the bright endothelial staining for Neu5Gc that characterizes wild-type mouse brains was not found in the transgenic animals, despite normal distribution of CD31 (Figure 4-4E). This may indicate that endogenous CMAH and Neu5Gc are suppressed in the vessels in the presence of high expression in adjacent cells. However, it is not clear that this immunohistochemical method is quantitative, and further analysis is underway to determine if this apparent suppression of endogenous Neu5Gc synthesis is in fact taking place.

Overexpression of CFE-CMAH with other Cre lines causes no gross abnormalities
The previous results from CMV-Cmah and NSE-Cmah transgenic mice showed strong evidence for embryonic lethality, leading us to expect that the CFE-CMAH lines would reproduce this embryonic lethality. However, we were surprised to discover that the Synapsin-Cre x CFE-CMAH mice demonstrated no lethality or gross abnormalities. We initially considered two possible hypotheses for this discrepancy. The first was that Synapsin-Cre, expressed at E13.5 in neurons, is turned on too late to induce a toxic effect. The second was that the previously observed lethality was in fact an extraneural effect in an NSE-expressing tissue.

We therefore proceeded to test a number of additional lines of Cre mice crossed with CFE-CMAH-C, the founder that produced the highest Neu5Gc expression when crossed with Synapsin-Cre. One was Nestin-Cre, which is expressed in neurons and glia around E9.5. We further tested NSE-Cre (also called ENO2-Cre), which is expressed at around the same time point but has some nonspecific expression in other tissues. Importantly, the NSE promoter was one of those used to drive Cmah in the initial transgenic models that demonstrated lethality. Although high levels of neural Neu5Gc can be produced by each of these transgenic Cre lines (Figure 4-5), there is surprisingly no deviation from Mendelian inheritance, perinatal lethality, or grossly abnormal phenotype yet observed. To account for lethality caused by Neu5Gc expression in non-neural tissues, we are also testing EIIa-Cre, which is expressed in the oocyte and should produce whole-body expression of Cmah in offspring.

Further efforts to characterize the anatomical and behavioral phenotypes of these animals are ongoing, in order to better understand the effects of high levels of neural Neu5Gc.

**DISCUSSION**

Here we describe the transgenic overexpression of Neu5Gc in the murine brain. The highly conserved absence of Neu5Gc from the vertebrate brain suggests an evolutionary requirement for its exclusion. Several initial experiments supported the hypothesis that overexpression of Neu5Gc in the brain would be detrimental to the organism; attempts to create mouse lines with CMV- and NSE-driven overexpression of Cmah produced no surviving founders.
with detectable transcription or functionality of the insertion. Further, the CMV-driven transgenic mice were demonstrated to have severe embryonic abnormalities in utero. The following experiments in Cre-inducible CFE-CMAH and T-CFE-CMAH lines have complicated these early findings. Mice from these lines can be induced to express Neu5Gc as a major fraction of total brain sialic acids. However, to date no lethality or grossly abnormal phenotypes have been observed. It is clear from the work completed to date that neural Neu5Gc expression is not, in and of itself, lethal to the embryonic mammalian brain.

This finding raises two significant questions that are currently being pursued. First, what was the cause of the lethality and abnormal phenotypes initially observed in the early transgenic lines? And second, if neural Neu5Gc presence is itself not lethal, what is the reason for its conserved absence from the brain? The successful development of this model for Neu5Gc overexpression aids us in answering these questions.

The reason for the initial lethality, and its subsequent failure to replicate in the inducible CFE-CMAH lines, is a puzzling question. We have confirmed that the sequence of Cmah in the CFE-CMAH constructs is identical to that used in the CMV- and NSE-driven transgenes. The successful survival of the CMV frameshift mutant lines, made in parallel, indicates that the observed lethality is unlikely to result from any earlier technical error. We therefore believe that the discrepancy is caused by the different methods used to generate the original and inducible lines. Early work on pronuclear injection of transgenes suggested that there is an immediate early burst of transcription of the transgenic construct prior to its integration in the genome. It is possible that the lethality of Neu5Gc overexpression actually occurs at this early point, at the single-cell stage. The EIIa-Cre crosses may support this if they reproduce the lethality. However, these have been reported to cause mosaicism of adult animals, suggesting that even expression of EIIa-Cre does not consistently occur at the single-cell stage. We are currently determining whether Neu5Gc presence is in fact toxic to a fertilized egg. Replication of the initial transgenic work, and comparison of the resulting embryos with those from the CFE-CMAH lines, may be necessary in order to fully address this puzzle.
The reason for the conserved absence of Neu5Gc from the vertebrate brain is a question we are actively pursuing using the CFE-CMAH lines now available. Since some of these mice are able to reach fractions of around 90% Neu5Gc in the brain, we can assume that Neu5Gc itself is not grossly toxic to neurons. Further, since expression throughout neural development does not cause lethality, we can conclude that there is no critical window during which Neu5Gc is specifically toxic to the nervous system. We are currently subjecting these mice to a series of anatomical and behavioral phenotypic analyses in order to determine whether there are any structural or cognitive abnormalities that can be observed. We are also considering gene microarray or RNA-seq analysis of brain tissue to determine whether any compensatory changes in gene expression occur in the presence of excess Neu5Gc.

In addition to these general analyses, we are specifically investigating phenotypes related to polySia and to myelin, since we expect them to be altered in the CFE-CMAH mice. As discussed in the previous chapter, alpha2-8-linked Neu5Gc demonstrates an unusual resistance to degradation by sialidase. We therefore predict terminal alpha2-8-linked Neu5Gc to be enriched in polySia and other alpha2-8-linked brain glycoconjugates, possibly as an age-related phenotype. PolySia dysregulation may also manifest as aberrant migration or connectivity within the brain. Additionally, it has previously been established that myelin-associated glycoprotein (MAG, Siglec-4) is unable to bind appropriately to its sialic acid-containing ligand in complex gangliosides when Neu5Gc, instead of Neu5Ac, is present. We expect, therefore, that aged CFE-CMAH mice may exhibit a degenerative phenotype of myelination similar to that observed in Mag⁻/⁻ mice. These specific molecular processes are being addressed in parallel with the broad phenotype analyses.

It is important to consider that, even if abnormal phenotypes are identified in these mice, they do not necessarily underlie the conserved exclusion of Neu5Gc from the vertebrate brain. It is therefore interesting to consider separately whether the overexpression of Neu5Gc is evolutionarily disadvantageous. We have not yet observed mortality or gross morbidity in these mice. However, a decrease in reproductive fitness is all that is necessary for selective pressure to act. To use an example from sialic acid biology, it has been demonstrated that a minor decrease
in fertility in matings of female $Cmah^{-/-}$ mice with male $Cmah^{+/+}$ mice can drive the deletion allele to fixation. This may be the very effect that drove the $Cmah$ deletion to fixation in humans $^{18}$. If neural Neu5Gc overexpression affects the fertility of an individual, then, this may be sufficient to require the exclusion of Neu5Gc from the brain. This could occur through numerous possible mechanisms, ranging from very specific abnormalities, for example, in pituitary neurons that secrete reproductive hormones, to very broad impairments, such as in social interaction or motor coordination. Although CFE-CMAH/Cre transgenic mice are fertile, we have not yet examined whether there is any resulting detriment to their fertility. Future studies will examine this question.

Even taking this multi-pronged approach to characterizing the CFE-CMAH mice, it is very challenging to properly tackle the question of why Neu5Gc is excluded from the brain. It is possible that the abnormalities caused by the presence of Neu5Gc require a trigger, such as stress or an infectious process, that is not replicated in laboratory conditions. Finally, we cannot yet rule out that this unusual finding is simply a side effect. For example, perhaps expression of $Cmah$ happens to be co-regulated with another gene on which the selective pressure acts. Or perhaps, even, the conservation of its suppression is coincidental—though, given the extent of the conservation only in brain tissue and over at least 300 million years, we find this unlikely.

Of course, until we are able to concretely identify the effects of the overexpression of neural Neu5Gc, these ideas remain speculative. Regardless, the development of the CFE-CMAH mouse lines will enable us to further investigate a decades-old mystery: the remarkable exclusion of Neu5Gc from the vertebrate brain.

REFERENCES


CHAPTER 5

The Effects of a Human-Specific Change in the Polysialyltransferase ST8Sia-II
ABSTRACT

Our laboratory previously identified a human-specific mutation in the polysialyltransferase ST8Sia-II, N308K, predicted to alter the protein’s secondary structure. This change occurs in sialylmotif S, one of the functional motifs found in all sialyltransferases, and is located at a residue that is otherwise highly conserved from chimpanzees to *Xenopus*. ST8Sia-II is one of two vertebrate enzymes responsible for the synthesis of polysialic acid (polySia), a polymer of α2-8 linked *N*-acetylneuraminic acid (Neu5Ac) primarily found on the vertebrate neural cell adhesion molecule (NCAM). PolySia plays numerous and diverse roles in neural development. It has been implicated in neuronal migration, neurite outgrowth, synaptic plasticity, and its dysregulation has been associated with several neuropsychiatric diseases including schizophrenia, bipolar disorder, and major depression, disorders that appear to be unique to humans. The human-specific change in ST8Sia-II is therefore an interesting finding to explore with regards to its effects on the evolution of the human brain. We have begun to characterize the effects of N308K on the activity and structure of ST8Sia-II. Although the overall enzymatic activity appears grossly unchanged, evidence so far suggests that the human ST8Sia-II produces longer chains of polySia, perhaps by interacting more effectively with a growing polySia chain through an expansion of the base-enriched polysialyltransferase domain. Further, N308K appears to destabilize the enzyme, potentially allowing for more rapid downregulation of polySia in development. Ancient genomic sequences of Neandertal and Denisovan individuals now available allow dating of N308K to at least 500,000 years ago, a time scale at which our ability to detect positive selection is very limited. We are currently exploring the consequences of this mutation in *in vivo* samples. ST8Sia-II and its product polySia form a fascinating, previously unexplored avenue in our understanding of the evolution of the human brain.

INTRODUCTION

The human brain is a remarkable organ. When compared with those of our closest relatives, the chimpanzees, our brains are striking for their large size, rapid rate of growth, and
delayed maturation, among other features. The combined result of these changes is a tissue that provides us significant cognitive complexity, including, among many other features, the universal use of language, advanced planning and problem solving, and an awareness of self and others within social networks. These features must result from specific biological changes along our lineage. Although many candidates have been proposed, identifying these presumed biological underpinnings has proved challenging.

Several years ago, our group identified a human-specific coding change in a highly conserved residue in polysialyltransferase ST8Sia-II. ST8Sia-II is highly expressed in early brain development, and is one of two vertebrate enzymes responsible for the synthesis of the important glycan polysialic acid (polySia). Phylogenetic analysis revealed a human-specific change in sialylmotif S of ST8Sia-II. Whereas the ancestral sequence contains a highly conserved asparagine, extending as far as Xenopus, the human sequence contains a lysine at position 308. (For known domains of ST8Sia-II, see Figure 2-2). This change, N308K, is predicted to alter the secondary structure from the ancestral coil structure to an alpha helix. The effect of N308K on protein structure and function has not been further characterized. However, the numerous and diverse roles played by polySia throughout brain development suggest that N308K is worth exploring further for its potential effects on brain evolution.

PolySia is an α2-8 linked polymer of the sialic acid N-acetylneuraminic acid (Neu5Ac) that is found as a posttranslational modification on a handful of vertebrate glycoproteins. Two polysialyltransferase enzymes are responsible for its synthesis: ST8Sia-II (also sometimes called STX) and its cousin ST8Sia-IV (PST). Knockout studies of these two enzymes in mice have indicated that ST8Sia-II is of primary importance in embryonic development, whereas ST8SiaIV is responsible for the trace levels of PSA-NCAM that remain in adult life. By far the most prevalently polySia-modified protein is the neural cell adhesion molecule (NCAM). Addition of polySia forms this protein’s “embryonic form”, commonly denoted PSA-NCAM—however, due to the more common use of the abbreviation PSA to refer to prostate-specific antigen, we will here use the abbreviation polySia-NCAM.
PolySia-NCAM is expressed at extremely high levels through embryonic and early postnatal development in mammals; however, levels throughout most of the brain drop to low levels as adulthood is reached. A similar pattern of expression appears to be present in humans, with adult levels of polySia reached by about five years of age. In the brain, two additional carrier proteins for polySia have been identified, although these are much less prevalent and less well-characterized: a second cell adhesion molecule, SynCAM, as well as the voltage-gated sodium channel in rat brain. PolySia-NCAM plays numerous and diverse roles in the developing nervous system. A vast body of publications have implicated polySia in developmental processes, including neurite outgrowth, neuronal migration, and synaptic plasticity, among others. Its dysregulation has also been implicated in a variety of human-specific neuropsychiatric disease, including schizophrenia, depression, and bipolar disorder. (For review and discussion, see Chapter 2 and Table 2-1). The molecular mechanisms of its involvement in these processes and disease states are not fully understood. PolySia is known to form an anti-adhesive cushion around NCAM, measurably increasing the distance between adjacent cells and preventing inappropriate gain of function of NCAM. However, it also has NCAM-independent roles, for example interacting directly with BDNF and FGF-2. It has additionally been identified as a potential ligand of Siglec-11, a sialic-acid binding lectin that has human-specific expression in brain microglia.

The expression and the time course of polySia are not well studied in humans, and no comparison has been made with other hominids. The importance and diversity of its roles, coupled with a known human-specific change in its major developmental synthetic enzyme, make it an interesting candidate that may be involved in some of the human-specific features of the brain. The fact that N308K occurs in a highly conserved residue in an important motif, where it is predicted to alter secondary structure, strengthens the likelihood that the change is functionally relevant to the enzyme.

In order to explore the significance of N308K in human evolution, we are characterizing the effects of this change on the protein’s structure and activity. Since human and chimpanzee
sequence of ST8Sia-II differ at only this single site, reverting the human coding sequence to the ancestral allele to create a “chimpanized” (chimp) enzyme allows us to study the effect of this single change in isolation, while also providing a direct comparison of the actual human and chimpanzee proteins.

We have compared the human and chimp ST8Sia-II enzymes in cell culture as well as in vitro. We find that, while the activity of the two enzymes is relatively similar in cell culture, human ST8Sia-II appears to produce longer chains of polySia. The resulting product binds much more strongly to both BDNF and FGF-2. In vitro assays additionally reveal an unusual increased instability of the human enzyme resulting from N308K that may be relevant to allowing rapid switches in polySia expression during development.

To better understand these changes, we used homology modeling of ST8Sia-II to the crystal structure of the related porcine ST3Gal-I. Major alterations in protein structure do not appear to be present. N308K may expand the basic surface patch produced by the polysialyltransferase domain, a region that interacts with the growing polySia chain during synthesis, corresponding to the results demonstrating increased length of polySia synthesized by human ST8Sia-II.

Phylogenetic analysis of ST8Sia-II indicates that N308K is fixed in humans, and at least 500,000 years old, as it is found to be present in both Neandertal and Denisovan archaic sequences. This age allows N308K to be involved in major developments in human brain structure, including a massive expansion in size that occurred with Homo erectus. There is some indication of positive selection at the site, but a single base-pair change of this age has very limited signatures of positive selection remaining in the genome.

Although we present initial data here, work is presently ongoing to more completely characterize the effects of N308K on the stability of the enzyme, the length of the polySia products, and to investigate what evidence exists of these differences between human and chimpanzee brains in vivo. This human-specific coding change in a highly conserved residue of an important developmental enzyme is an interesting and unusual finding. Further exploration of
the polysialyltransferases and polySia may reveal new ways in which this glycan has impacted vertebrate brain evolution.

MATERIALS AND METHODS

Sequencing and phylogeny

Consensus peptide sequences for human, chimpanzee, bonobo, and gorilla were aligned using CLUSTALW multiple sequence alignment tool available through the Biology WorkBench. To rule out polymorphism at N308K, 44 human HapMap samples (Coriell) and 16 chimpanzee genomic DNA samples (provided by Pascal Gagneux, UC–San Diego) were further sequenced. Primers used for sequencing were GCTGACCAACAAAGTCCACA and TCTGGTTCAGGTGGGTTCTC. All sequencing was done by the company GENEWIZ. Resulting chromatograms were aligned and analyzed using Sequencher (Gene Codes Corporation).

Plasmids

pPROTA-humanST8SIA2-V5 (pPROTA-hST8SIA2) and pPROTA-chimpST8SIA2-V5 (pPROTA-cST8SIA2) encoding recombinant soluble human and chimp ST8Sia-II chimeric with protein A and V5, respectively, and pcDNA3.1-human ST8SIA2-V5/His (pcDNA-hST8SIA2) and pcDNA3.1-chimpST8SIA2-V5/His (pcDNA-cST8SIA2) encoding full-length human and chimp ST8SIA2, respectively, with V5 and (His)_6 tags, were used in this study. Mutagenesis of pPROTA-hST8SIA2 and pcDNA-hST8SIA2 was performed using the QuickChange Site Directed Mutagenesis kit (Stratagene), resulting in the construction of pPROTA-cST8SIA2-V5 and pcDNA-cST8SIA2. A construct containing a human NCAM-Fc chimera was prepared in pcDNA4. All prepared constructs were confirmed by sequencing.

Cell lines

Murine Neuro2a (N2a) neuroblastoma cells were grown in Dulbecco’s Modified Eagle’s Medium (DMEM) supplemented with 10% fetal bovine serum. Cells were transfected with
pcDNA3.1-ST8SIA2 constructs using polyethyleneimine (PEI), and positive transfectants were selected in 500 µg/ml of G418 for 10 days.

CHO cells were prepared in the laboratory of Ken Kitajima (Nagoya University) and maintained in alpha Minimum Essential Medium (alpha-MEM) with ribonucleosides and deoxyribonucleosides, supplemented with 10% fetal bovine serum. Cells were initially stably transfected with pcDNA4-NCAM-Fc under selection with Zeocin. These cells were then additionally transfected with pcDNA3.1-hST8SIA2 or cST8SIA2 using Lipofectamine and selected with G418.

iPS-derived neurons were derived from skin fibroblasts of human, chimpanzee, bonobo, gorilla, and rhesus individuals. Cells were kindly provided by Carol Marchetto in the laboratory of Fred Gage (Salk Institute) after 3 weeks of differentiation.

RT-qPCR

RNA was prepared from cells using an RNeasy Mini Kit (Qiagen). RNA from primate iPS-derived neurons was from the laboratory of Fred Gage, Salk Institute. Reverse transcription was performed on 1-2 µg of RNA using AffinityScript QPCR cDNA Synthesis Kit (Agilent) using random primers. The resulting cDNA was analyzed by quantitative PCR in triplicate using SYBRGreen master mix (Qiagen) on a CFX Connect Real-Time PCR machine (Bio-Rad).

Primer sequences used were as follows.

\[
\text{ST8SIA2 for N2a and CHO cells: TATCCTGAAGCACCACGTCA (forward)}
\]

\[
\text{TGTGGACTTTTGTGGTCAGC (reverse)}
\]

\[
\text{ST8SIA2 in primate iPS-derived neurons: GGGTCAACGAGCTTATCCTG (forward),}
\]

\[
\text{TGTGGACTTTTGTGGTCAGC (reverse)}
\]

\[
\text{NCAM-Fc: GGC ATT TAC AAG TGT GTG GTT AC (forward), TTG GCG CAT TCT TGA}
\]

\[
\text{ACA TGA (reverse)}
\]

\[
\text{Gapdh control in N2a cells: AGGTCGGTGTGAACGGATTTG (forward),}
\]

\[
\text{TGTAGACCATGTAGTTGAGGTCA (reverse)}
\]
\textit{Gapdh} control in CHO cells: TGTGGAAGGACTCATGACCA (forward), GGATGCAGGGATGATGTTCT (reverse)

\textit{HMBS} control in iPS-derived neurons: TGCCAGAGAAGGTGTGGTG (forward), GAGGTTTCCCCGAATACTCC (reverse)

Flow cytometry

N2a cells were stained with the anti-polySia antibodies 12E3 (kindly provided by the laboratory of Ken Kitajima) and 735 (generously provided by Rita Gerardy-Schahn). mAb 12E3 (mouse IgM) was used at 5 µg/ml and 735 (mouse IgG) at 2 µg/ml, followed by appropriate fluorescent secondary antibodies. All flow cytometric data was obtained on a BD FACSCalibur and data analyzed using FlowJo (TreeStar).

DMB-HPLC analysis of polySia chain length

Supernatant was collected from CHO cells stably transfected with pcDNA4-NCAM-Fc and pcDNA3.1-hST8SIA2, pcDNA3.1-cST8SIA2, or empty pcDNA3.1 vector. For each preparation, a 1 ml DEAE-Sephadex column was prepared and washed with 5 ml of PBS (150 mM NaCl). Supernatant was added to each column and allowed to pass through by gravity flow. Each column was washed with 20 ml of PBS adjusted to 200 mM NaCl. Bound anions were then eluted with 3 ml of PBS adjusted to 600 mM NaCl. The eluate was washed twice with addition of 12-15 ml of water and concentration through a 30-kD Amicon Ultra-15 centrifugal filter unit (Millipore), resulting in a concentration of 1.6 mM NaCl.

To 20 µl of each sample was added 20 µl of 0.01 N trifluoroacetic acid (TFA) and 40 µl of modified DMB reagent (7 mM DMB, 5 mM TFA, 1M 2-mercaptoethanol, 18 mM sodium hydrosulfite). Samples were incubated at 50˚C in the dark for 1 h. Lactones were removed with the addition of 4.2 µl of 2 M NaOH (final concentration 0.1 M) and incubation at 37˚C for 30 min in the dark. Derivatized samples were analyzed by HPLC on DNAPac PA-100 column in buffer A: 20 mM Tris-HCl, pH 8.0, with a gradient of buffer B: 20 mM Tris-HCl, pH 8.0, 1 M NaCl. Buffer A was
run alone for 15 min, followed by a gradient of 0-60% buffer B from 15-60 min. Fluorescence was monitored at Ex=373 nm, Em=448 nm.

**Biacore analysis**

All Biacore studies were performed in the laboratory of Ken Kitajima, Nagoya University. The Au sensor surface was washed once with acetone and after drying, the chip was immersed in 10 µM DBA in ethanol to form a self-assembly membrane (SAM) on the Au surface. After gently shaking for 30 min at room temperature, the sensor surface was washed with ethanol three times and allowed to dry. The chip was then placed in a solution of EDC and NHS (a 1:9 mixture of 130 µM EDC in water and 144 µM NHS in 1,4-dioxane) at room temperature for 30 min with gentle shaking to activate the SAM on the Au surface. After adding water, the surface was incubated for 5 min, and then washed. The Au chip containing surface-activated SAM was placed on the sensor chip support using the sensor chip assembly unit and set in a Biacore 3000 instrument. After priming the system with water for 7 min, a 0.1 mg/ml protein A solution was loaded twice, each time for 7 min at a flow rate of 10 µl/min. Immobilized streptavidin was monitored by measuring the resonance unit (RU) value, which typically reached 1300-1850 RU for protein A. To destroy excess activated groups, 1 mM ethanolamine was injected into the system for 7 min. After washing with HBS-EP (0.01M HEPES pH7.4 containing 0.15M NaCl, 3mM EDTA, and 0.0005% Surfactant P20), purified polySia-NCAM-Fc (0.1 mg/ml in 500 mM HBS-EP) was injected into the system to allow immobilization on the Au surface. Immobilization of the polySia-NCAM-Fc was monitored based on the observed RU values, which typically reached approximately 850-1300 RU. NCAM-Fc derived from a mock transfectant was used as a negative control.

For analysis of the interactions between immobilized polySia-NCAM-Fc and the two neuroactive molecules, varying concentrations of BDNF (0-37.0 nM) or FGF2 (0-56.8 nM) in HBS-EP were injected over the polySia-NCAM-immobilized sensor chip surface at a flow rate of 20 µl/min. After 120 s, HBS-EP was flowed over the sensor surface to monitor the dissociation
phase. Following 180 s of dissociation, the sensor surface was fully regenerated by the injection of 10 µl of 3 M NaCl. The analyses were performed three times and all values are expressed as the mean ± SD.

**In vitro assays of ST8Sia-II on NCAM-Fc**

pPROTA-hST8SIA2 and cST8SIA2 were transfected into COS-7 cells using PEI and medium changed after five to eight hours. After two days, medium was collected, cellular material removed by centrifugation, and one-tenth volume of 1 M Tris-HCl buffer was added to the supernatant. To each 10 ml of medium was added 30 µl of Ig-Sepharose 6 FastFlow beads (GE Healthcare Life Sciences) 1:1 in PBS, and tubes were incubated overnight at 4˚C with rotation. Beads were collected by centrifugation at 1000 x g for 5 min, washed 3 times with ice-cold PBS, and resuspended in an equal volume of 50 mM MES buffer, pH 6.0 with NaOH. Each preparation of enzyme was quantified by SDS-PAGE and modified Coomassie Stain (GelCode Blue, Pierce).

NCAM-Fc was prepared by transient transfection of pcDNA4-NCAM-Fc into CHO-Tag cells. Medium was collected after 3 days and incubated for at least 4 h with protein A Sepharose 4 FastFlow beads (GE Healthcare LifeSciences) at 4˚C with rotation. Beads were collected in a column and washed with 5 ml of PBS. NCAM-Fc was eluted with 100 mM glycine-HCl, pH 3, and neutralized with sodium hydroxide.

Reactions were allowed to proceed at 37˚C with agitation in a total volume of 20 µl. Reaction buffer was 50mM sodium cacodylate, 5mM MnCl₂ pH 6.7. Each reaction contained 15 pmol NCAM-Fc, 30 pmol CMP-[9-³H]Neu5Ac (300,000 cpm), 720 pmol CMP-Neu5Ac (Nacalai), and 5 µl of enzyme bead preparation (0.5-1 µg). At appropriate timepoints, reactions were quenched with the addition of 180 µl of 100mM Tris-HCl, 20mM EDTA pH 8, incubation at 50˚C for 10 min, centrifugation at 1000 x g for 5 min, and separation of the supernatant from the enzyme beads.

50 µl of each reaction supernatant was spotted on 3MM Whatman chromatography paper and run to approximately 14 cm in a solvent ethanol and 1 M ammonium acetate (7:3, v/v), pH
7.5. Paper was allowed to air dry. The origin was cut out, added to 500 µl water and 5 ml of Scintiverse BD cocktail (Fisher). Vials were left for at least 1 hour to equilibrate before counting.

**In vitro assays on DMB-DP3**

DMB-labeled Neu5Ac trimer (DMB-DP3) was prepared as described elsewhere \(^{15}\). Briefly, commercial \((\text{Neu5Ac})_3\) (Nacalai) was DMB-derivatized in 7 mM DMB, 0.25M Na\(_2\)S\(_2\)O\(_4\), 0.5 M 2-mercaptoethanol, 20 mM trifluoroacetic acid. DMB-DP3 was separated from unlabeled trimer by HPLC on a DNAPac PA-100 column using a gradient of 1 M sodium nitrate from 0-3% over 30 min. Elution was tracked by fluorescence detection at Ex=373 nm, Em=448 nm. Pooled fractions were dialyzed in water, lyophilized, and weighed to quantify yield of reagent.

Reactions of total volume 20 µl contained 400 pmol DMB-DP3, 80 nmol CMP-Neu5Ac (Nacalai), and 5 µl of enzyme preparation. The reaction buffer was 10 mM sodium cacodylate with 10 mMnCl\(_2\), pH6-7. In some cases, 20 pmol of NCAM-Fc were treated with PNGase F (New England Biolabs) and added to the reaction also. Tubes were incubated overnight at room temperature with rotation. Reactions were quenched as described above and an aliquot of supernatant analyzed by anion-exchange HPLC on a DNAPac PA-100 column. The buffers used were water with added 1 M sodium nitrate, at a concentration of 0% (0-5 min), 1% (10 min), 10% (20 min), 50% (65 min), followed by 100% for 10 min to clean remaining material from the column.

**Detection of polySia and NCAM by immunoblot**

Snap-frozen frontal cortex samples from three human and three chimpanzee individuals, as well as iPS-derived neurons, were lysed in 20 mM Tris-HCl, pH 8.0, 150 mM NaCl, 5 mM EDTA, 2% Triton X-100, and 1x protease inhibitor cocktail added just before use. Protein content was quantified using the BCA Protein Assay kit (Pierce). To minimize degradation of polySia, samples were denatured for 10 min at 60˚C instead of boiling. 100 µg of protein were loaded per well on a 7% SDS-PAGE gel. For each sample, one aliquot was pretreated with endoN (a gift of Rita Gerardy-Schahn). After transfer to PVDF membrane, blots were stained using standard
immunoblot procedure. mAb 12E3 was used at 1 \( \mu g/ml \) followed with anti-mouse IgM conjugated to IRDye800 (Rockland Immunochemicals). mAb 735 was used at 0.2 \( \mu g/ml \) followed by anti-mouse IgG conjugated to IRDye800 (LI-COR Biosciences). A monoclonal rabbit anti-NCAM (EPR2566, Abcam) was used at 1:2000 followed with anti-rabbit secondary conjugated to IRDye680. All blots were imaged and analyzed using a LI-COR Odyssey system.

**Homology modeling**

The program Phyre2 was used to map consensus peptide sequences of human and chimpanzee ST8Sia-II onto known structures\(^\text{16}\). The primary crystal structure used was porcine ST3Gal-I\(^\text{17}\). MacPyMOL (The PyMOL Molecular Graphics System, Schrödinger, LLC) was used to align resulting models and to create all structural images.

**Phylogenetic testing for positive selection**

Initial analysis for positive selection was conducted via the Phylogenetic Analysis by Maximum Likelihood software\(^\text{18}\).

Additional analyses were performed with the suite of online tools available through Data Monkey\(^\text{19}\). Specific tools used were GA-Branch\(^\text{20}\) and Mixed Effects Model of Evolution (MEME)\(^\text{21}\).

**RESULTS**

*Uniquely human genomic single base pair changes alters a highly conserved amino acid residue of the enzyme ST8Siall*

With more genomic data from a variety of species becoming continually available, we initially wanted to confirm that N308K is present only in the human lineage and that the site is not polymorphic in human or ancestral populations. We extended the previous phylogenetic alignment using more recently available consensus genomic sequence data from bonobo and
gorilla. These results confirm that the predicted human-specific change is not found in these other species (Figure 5-1).

To verify that this site is not polymorphic in human or chimpanzee populations, we also sequenced 44 HapMap human individuals and 19 chimpanzees at this locus. The asparagine residue was encoded in all chimpanzees, whereas the lysine was present in all humans tested. After the recent publication of over 1000 human genome sequences from 14 global populations, we further verified that no coding changes were identified at this site among those individuals. Thus N308K is fixed within the human population, and not present in other species.

Importantly, although there are seven differences in the nucleotide coding sequence between humans and chimpanzees, only one is nonsynonymous. Aside from N308K, the peptide sequence of ST8Sia-II is identical among humans, chimpanzees, bonobos, and gorillas.

**Structural modeling indicates that human change may expand a basic patch on surface of protein**

Structural analysis of ST8Sia-II can help us to understand what effect N308K might have on the structure and function of ST8Sia-II. Unfortunately, no laboratory has successfully crystallized this protein. Reports via personal communication indicate that it is highly challenging to crystallize due to its instability and propensity to aggregate. Thus, we used homology modeling to map ST8Sia-II onto the crystal structure of the only vertebrate sialyltransferase crystallized to date, pig ST3Gal-I. The resulting crystal structures are grossly similar—however, given that we are building these models over an underlying structure, this may be the expected finding (Figure 5-2A).

Interestingly, although N308K is within the helix of sialylmotif S as predicted, it is on the distal portion of the helix away from the area predicted to interact directly with the donor and acceptor molecules (Figure 5-2A). It is important to note that only very limited inferences can be made about ST8Sia-II from this comparison to a single enzyme, ST3Gal-I, with different activity and specificity. Despite this distance, mutation at the equivalent residue has been found to affect
Figure 5-1. Alignment of peptide sequences of ST8Sia-II. Consensus peptide sequences for ST8Sia-II for the indicated species were aligned using CLUSTALW. The peptide sequence is identical among this group of primates, with the exception of N308K, indicated with an arrow.
Figure 5-2. Homology modeling of ST8Sia-II to the crystal structure ST3Gal-I suggests N308K is distant from the catalytic site. Human and chimp peptide sequences of ST8Sia-II were homology modeled to known crystal structures, primarily to porcine ST3Gal-I, using the program Phyre2. A: human (left) and chimp (right) structures are relatively similar. Sialylmotifs are color-coded as follows: sialylmotif L (red), sialylmotif S (cyan), sialylmotif 3 (green), sialylmotif VS (blue). Donor and acceptor molecules from ST3Gal-I are included (yellow) to indicate the presumed catalytic site. B: Homology model of human ST8Sia-II demonstrates N308K (red) distant from the presumed catalytic site, indicated by the ST3Gal-I donor and acceptor (yellow).

Other domains, both enriched for basic residues, that have been identified as important in polysialyltransferase function include the polysialyltransferase domain (blue), thought to interact with a growing polySia chain, and the polybasic domain (green), thought to bind to the carrier protein.
the activity of the related enzyme ST8Sia-I on substrate ganglioside GM3. The position of the residue, therefore, does not rule out its involvement in the catalytic activity of the enzyme.

To further examine how N308K might affect the function of ST8Sia-II, we looked at its position relative to other important domains within the protein. Two additional regions have been identified in the two polysialyltransferases; these regions are thought to be specifically important to the formation of polySia. The polysialyltransferase domain, near sialylmotif S in the peptide sequence, is a region enriched for basic residues, thought to interact with the growing polymer as the enzyme forms the polySia chain. The polybasic domain, closer to the N-terminal transmembrane region of the protein, is also enriched for basic residues, but this has been shown to interact instead with an acidic patch on NCAM. On the surface of the ST8Sia-II model, N308K, itself a change to a basic residue, appears to be positioned close to the polysialyltransferase domain (Figure 5-2B). Thus the mutation may expand this basic patch, perhaps allowing stronger interaction with a growing polySia chain. We suggest this may affect either the total enzymatic activity or the length of the synthesized polySia.

**Human amino acid change in ST8Sia-II has no clear effect on cell surface polySia production**

To assess the effect of N308K on enzymatic activity, we initially used cell-culture based assays. A plasmid of pcDNA3.1 carrying the cDNA of human ST8SIA2 allowed us to express exogenous human enzyme in transfected cells. To express chimpanzee enzyme, we used targeted mutagenesis on this plasmid to revert the human lysine back to the chimpanzee arginine. Although this “chimpanized” version of ST8SIA2 (hereafter referred to as “chimp”) does not perfectly match the chimpanzee genomic coding sequence, all differences are synonymous. The resulting protein, therefore, is identical to that encoded by the genomic chimpanzee sequence.

Neuro2a (N2a) is a murine neuroblastoma line that expresses endogenous NCAM but does not polysialylate this protein. We transfected a population of these cells with human and chimp pcDNA3.1-ST8SIA2 and selected with the antibiotic G418, producing mixtures of stably
Figure 5-3. Human and chimp ST8SIA2 produces equal amounts of polySia in N2a transfectants. N2a cells, which normally express NCAM but not polySia, were stably transfected with pcDNA3.1 containing the human or chimp ST8SIA2 cDNA. A: The two stable lines express similar amounts of ST8SIA2 when analyzed by RT-qPCR. Vector-only negative control does not express any ST8SIA2. Error bars indicate standard deviation across three replicates. B: Flow cytometric analysis with two anti-polySia antibodies, 12E3 (left) and 735 (right) demonstrates very similar expression of surface polySia-NCAM in both human and chimp transfected lines. As a negative control, cells were pretreated with endoNF, a phage endoneuraminidase that selectively removes surface polySia.
transfected cells. RT-qPCR analysis of RNA from the resulting cell lines showed similar
expression of ST8SIA2 in both human and chimp transfected cells (Figure 5-3A).

We then analyzed surface expression of polySia on these cells using two monoclonal
antibodies to polySia (Figure 5-3B). mAb 12E3, a mouse IgM antibody, requires a minimal degree
of polymerization (DP) of 5 Neu5Ac residues, and binds only to the nonreducing end of a polySia
chain \(^27\). mAb 735, a mouse IgG antibody, requires a longer DP of 11, and its internal binding is
responsive to increasing length of polySia chains \(^{28,29}\). As a negative control, cells were
pretreated with endoN, a phage endoneuraminidase that selectively degrades polySia. Flow
cytometric analysis of the N2a lines using both mAb 12E3 and mAb 735 demonstrated no
difference between the human and chimp transfectants (Figure 5-3B). Therefore, within the
context of this cell line, there is no obvious difference in total activity between human and chimp
ST8Sia-II.

**Human ST8Sia-II produces longer polySia in culture**

To examine the chain length of synthesized polySia with higher resolution, CHO cells
were stably transfected with NCAM-Fc and human or chimp ST8SIA2, resulting in the secretion of
polySia-NCAM-Fc into the culture medium. We confirmed that all cells express NCAM-Fc and
ST8SIA2 as expected (Figure 5-4A). The expression of NCAM-Fc, the product of the intial
transfection, was highly variable among the three lines. The ratio of ST8Sia-II to NCAM-Fc is
much higher in the chimp ST8SIA2 transfectants than the human, potentially biasing the studies
involving these cells towards detecting higher activity from the chimp enzyme.

The supernatant from the cells was collected and polySia-NCAM-Fc was purified using a
DEAE-Sephadex column. The resulting polySia-NCAM-Fc was then DMB-derivatized using a
modified method. This technique uses mild acid hydrolysis to preferentially cleave the inner most
\(\alpha 2-3\) linkage. The relatively intact polySia chain is then labeled at the reducing end with DMB.
The samples were analyzed by anion-exchange HPLC using a DNAPac PA-100 column with
NaCl gradient.
Figure 5-4. Human ST8Sia-II produces longer polySia chains. CHO cells were stably transfected with both human NCAM-Fc and human or chimp ST8SIA2, resulting in the secretion of polySia-NCAM-Fc into the supernatant. A: RT-qPCR analysis indicates expression of NCAM-Fc and ST8SIA2 in transfected cells. Error bars indicate standard deviation across three replicates. B: DMB-HPLC analysis of secreted polySia-NCAM-Fc demonstrates longer polySia chains synthesized by the human enzyme, top panel, than those synthesized by the chimp enzyme, bottom panel. Maximal chain length detected was 36±2.3 for human, and 29±1.2 for chimp transfectants. C: Binding of polySia-NCAM-Fc to the trophic factors BDNF and FGF-2 was measured by surface plasmon resonance. PolySia-NCAM-Fc was immobilized and BDNF or FGF-2 used as an analyte. In both cases, the human product binds more strongly than the chimp equivalent. Data shown in parts B and C is from Masaya Hane, Chihiro Sato, and Ken Kitajima of Nagoya University.
In preliminary results, polySia-NCAM-Fc secreted from the human ST8SIA2 transfectants is slightly longer than that secreted from the chimp transfectants. The overall profiles are quite different, with the maximum DP detected 36±2.3 for human ST8SIA2-transfected cells, and 29±1.2 for chimp ST8SIA2-transfected cells (Figure 5-4B). Importantly, since the ST8SIA2 to NCAM-Fc expression ratios actually put the human transfectants at a disadvantage (Figure 5-4A), this may be an underestimate of the true difference between the two enzymes.

PolySia has previously been demonstrated to bind directly to BDNF and to FGF-2.10,11 To determine whether the observed length difference affects this binding, the polySia-NCAM-Fc secreted from stable CHO lines was analyzed for BDNF and FGF-2 binding by surface plasmon resonance (SPR) in the laboratory of Ken Kitajima. The polySia-NCAM-Fc was immobilized, and the respective factors used at various concentrations as analytes. Both BDNF and FGF-2 demonstrate strikingly higher binding to both factors (Figure 5-4C).

This finding is surprising given the relatively minor difference in length that was observed by DMB-HPLC (Figure 5-4B). It may indicate a threshold effect, by which the slightly longer polySia synthesized by the human enzyme binds much more strongly to these trophic factors. It may also indicate alterations in the branching pattern, which is not recognized by HPLC. We are currently working to confirm the length difference in synthesized polySia observed here. Such a difference would have fascinating implications for the human and chimpanzee brains, through potential effects on the signaling of these important trophic factors.

**In vitro comparison of enzymes suggests human ST8Sia-II is less stable**

We attempted to use *in vitro* assays to confirm the differences produced by the enzymes in cell culture. Since ST8Sia-II is typically bound within the Golgi membrane, we created protein A chimeras, in which the transmembrane domains of the enzymes were replaced with protein A sequence. Plasmids encoding these fusion proteins were transiently transfected into COS-7 cells, and the enzyme was collected from the supernatant on Ig-Sepharose beads. As an acceptor protein for polysialylation, we used NCAM-Fc fusion produced in CHO-Tag cells, collected and
eluted from protein A-Sepharose beads. Reactions were conducted at 37°C using CMP-[³H]Neu5Ac as a Sia donor. The reaction components were separated from the resulting polySia-NCAM-Fc product by paper chromatography.

In initial studies, chimp ST8Sia-II consistently demonstrated higher activity than the human enzyme. We studied the activity of the enzymes across time. Within the first hour, both enzymes appear to produce polySia linearly, with the rate of chimp enzyme approximately 1.4 times that of the human enzyme (Figure 5-5A). As the reactions continue, however, they rapidly lose linearity, indicating that the enzymes are unstable within the time period tested. This instability appears to be much more severe in the human enzyme than in the chimp (Figure 5-5A).

We attempted to repeat this experiment using a DMB-labeled trimer of Neu5Ac (DMB-DP3). This fluorescent molecule has been described elsewhere as forming a universal primer for all known bacterial and mammalian polysialyltransferases. We were able to achieve only minimal extension of DMB-DP3 with an overnight reaction of either ST8Sia-II (Figure 5-5B). To attempt to further the reaction, we added NCAM-Fc that was treated with PNGase, an enzyme that removes all N-glycans. This NCAM-Fc, which would be unable to be polysialylated itself, could presumably still interact with the enzymes and induce conformational changes. Interestingly, this did enable the enzymes to proceed further with the reactions (Figure 5-5C). Once again, the chimp enzyme exhibited a higher degree of activity and/or stability than the human enzyme. However, since the enzymes’ instability still prevents the reactions from reaching completion, it is not possible to compare the maximal chain length of their products.

Further formal tests of the stability of human and chimp ST8Sia-II are currently underway. This work will allow us to further understand the degree of instability produced by the human mutation and the significance it may have in a physiological system.

**Differences in expression of polySia in human and chimpanzee neurons**

The initial results from cell culture are promising, but for them to be meaningful *in vivo*, N308K must cause significant changes in human brain itself. The challenge of addressing this
Figure 5-5. N308K confers instability to human ST8Sia-II. A: *In vitro* reactions measure the incorporation of radiolabeled Neu5Ac into polySia on NCAM-Fc by human and chimp protein A ST8Sia-II chimeras. Initial reaction rates are approximately 1.4-fold faster with the chimp enzyme. However, both enzymes are highly unstable and lose linearity within a few hours; this seems to be exaggerated in the human enzyme. B: A DMB-labeled trimer of Neu5Ac (DMB-DP3) (top panel) was used to prime both ST8Sia-II reactions. Because the enzymes are highly unstable, neither human (middle) nor chimp (bottom) reaction was able to proceed to completion. Again, the higher stability of chimp ST8Sia-II produces higher activity than the human enzyme. C: Identical reactions to B were conducted in the presence of PNGase-treated NCAM-Fc at an equimolar concentration to the ST8Sia-II enzyme. Interestingly, the addition of this molecule to the reaction mixture induces a beneficial conformational change in both enzymes, allowing them to achieve higher activity on DMB-DP3.
properly is that ST8Sia-II is most relevant to embryonic and very early postnatal development; most polySia present in adulthood is synthesized by ST8Sia-IV. We analyzed polySia expression in frontal cortex samples from three human and three chimpanzee individuals. All samples were from adults; chimpanzees were aged 44-49 years at time of death, and humans were 66-79 years. Immunoblots for both polySia and NCAM expression were performed with or without treatment with endoN, which cleaves all polySia. PolySia expression normalized to NCAM expression was found to be approximately equal across all individuals, perhaps the expected result for adult samples (Figure 5-6).

To properly query the activity of ST8Sia-II, this experiment should ideally be conducted on age-matched samples from embryonic or very early postnatal life. Of course, the availability of such samples is quite limited. To circumvent this problem, we are exploring the use of early neurons differentiated from induced pluripotent stem (iPS) cells derived from primate fibroblasts. These samples will allow us a look into the early neuronal development of humans and chimpanzees.

**Differences in expression patterns of ST8SIA2 between humans and other great apes**

More complete understanding of the significance of N308K in vivo requires an analysis of not just the enzymatic activity, but also the level of its expression. A recent paper used RNA sequencing to study developmental differences in gene expression between humans, chimpanzees, and rhesus macaques in the frontal cortex and cerebellum. The authors identified a number of genes with “human-specific” expression profiles; i.e. the pattern of expression across development in humans was distinct from that found in the other two primate species. One of those identified was in fact ST8SIA2. Levels of ST8SIA2 mRNA are significantly higher in neonatal humans than in chimpanzees or rhesus macaques. Over time, the human level declines to levels similar to those found in the other primates (Figure 5-7A).

To support this finding, we analyzed expression of ST8SIA2 in neurons derived from induced pluripotent stem cells (Figure 5-7B). These results also trend towards higher
Figure 5-6. PolySia expression does not significantly differ in human and chimpanzee adult frontal cortex. Frontal cortex samples from adult humans and chimpanzees were lysed and analyzed for polySia expression using mAb 735 or 12E3, and for NCAM expression after treatment with endoNF. A: Immunoblots demonstrate similar expression of polySia and NCAM across all samples. B: Quantification of the immunoblots shown in A demonstrate no significant difference between human and chimpanzee individuals. Error bars indicate standard deviation across all three samples of each species.
Figure 5-7. Evidence for increased expression of ST8SIA2 in human development. A: Data taken from a recent study of RNA in primate frontal cortex determined that ST8SIA2 has a “human-specific” expression pattern. Compared to chimpanzee and rhesus individuals, humans appear to have high expression of ST8SIA2 in very early postnatal life that declines with age. B: Expression of ST8SIA2 was analyzed by RT-qPCR on RNA from iPS-derived neurons created from fibroblasts. Although the small sample size is limiting for statistical analysis, there is a trend towards increased expression in the human derived samples. Error bars indicate standard deviation across 3 replicates for each samples. p-values were calculated by two-tailed t-test across the groups indicated in brackets.
ST8SIA2 expression in humans compared to other primates. The sample size is too small to properly determine whether this difference is statistically significant \((p = 0.22\) for human versus chimpanzee samples). Grouping all nonhuman primates together \(i.e.\) all samples that have the ancestral lysine at N308K gives a borderline \(p\)-value of 0.05.

The finding of higher expression of ST8SIA2 in human neurons suggests that there may be more significance to this enzyme in human evolution than the effect of the N308K coding change itself. The increased expression may partly occur as a compensatory mechanism, countering the instability caused by N308K. It is interesting to consider, however, that the increased expression and decreased stability may together allow for higher enzymatic activity with the ability to more rapidly turn off the enzyme.

**Timing of N308K change during human evolution**

Recent advances have allowed sequencing of genomes from Neandertals and Denisovans, two related hominid species thought to have diverged from human ancestors about 500,000 years ago \(^{31,32}\). These sequences are extremely useful in aiding the timing of human-specific mutations such as N308K. Indeed, both species’ genomes match the human genome at the base in question (Figure 5-8). We can therefore date the mutation’s occurrence after the last common ancestor with chimpanzees but before the last common ancestor with Neandertals and Denisovans, so between about 6 million years ago and 500,000 years ago. Although this window remains quite wide, it is the period during which most neural changes in human evolution occurred, including the major brain expansion in *Homo erectus*.

Unfortunately, the age of this mutation limits our ability to detect positive selection at the locus. A number of signatures of positive selection have been identified, for example the enrichment of rare alleles and the presence of unusually long haplotypes in the region of a selective sweep. However, the various signatures of positive selection decay over time as mutation and recombination continue to occur. As a result, the only analyses available for a
Figure 5-8. Alignment of human, Neandertal, and Denisovan sequences allow dating of N308K to at least 500 kya. A: Alignment of human peptide sequence with that predicted by Neandertal and Denisovan genomic sequence. Dots indicate identity to reference (human) sequence. Dashes indicate absence of sequence data at that locus. The location of N308K is marked with an arrow.
mutation of this age rely on the rate of functional changes occurring along a given lineage, using the rate of synonymous changes as an approximation of the neutral rate of change\textsuperscript{33}.

In the case of \textit{ST8Sia2}, such an analysis is limited by the fact that there is only one nonsynonymous change along the human lineage. Nevertheless, we used a number of analyses to determine whether any evidence of positive selection remains in the human genome. We initially used PAML, a program that measures phylogenetic analysis by maximum likelihood\textsuperscript{18}. Most notably, this analysis has been used to demonstrate rapid evolution of the transcription factor \textit{FOXP2} along the human lineage\textsuperscript{34}. Using consensus sequences of human, chimpanzee, orangutan, rhesus, and marmoset, we tested the hypothesis that the human rate of functional change was unique against the null hypothesis that all species shared the same neutral rate of change. This resulted in a \textit{p}-value of 0.109.

We pursued two other analyses available through the online suite of tools Data Monkey\textsuperscript{19}. GA Branch uses a similar method to PAML. Specifically, it compares the rates of functional change along a number of lineages, and fits a phylogenetic model to that analysis. GA Branch calculates a \textit{dN/dS} of 0.189 on the human branch, in a much lower background of \textit{dN/dS} of 0.007 in chimpanzee, bonobo, and gorilla\textsuperscript{20}. Lastly, the Mixed Effects Model of Evolution (MEME) examines evidence for selection at a specific site; it finds marginal evidence for selection at position 308 with \textit{p}=0.156\textsuperscript{21}.

None of these analyses reach the traditionally accepted statistical significance at \textit{p}<0.05. These tests do, however, suggest a reasonable possibility that positive selection occurred at N308K along the human lineage. Unfortunately, given the age of this mutation, it is not possible to investigate this question further using existing sequence data and current techniques.

\textbf{DISCUSSION}

Here we describe our ongoing efforts to characterize the effects of a human-specific change in the polysialyltransferase \textit{ST8Sia-II}. The peptide sequence of \textit{ST8Sia-II} is completely identical among chimpanzees, bonobos, gorillas, and humans, with the exception of a single
change in humans. This change, N308K, is within a highly conserved residue, in which the ancestral asparagine is present as distantly as *Xenopus*, and is predicted to alter the secondary structure of the protein. With new genomic data now available, we are able to date N308K to between about 6 million and about 500,000 years ago, and suggest that it may be under positive selection. The data we have obtained to date on the effect of this change are limited; however, they suggest that N308K may alter both the product of the enzyme and the stability of the protein itself.

Through homology modeling to ST3Gal-I, we show here that N308K is actually distant from the catalytic site of the enzyme, yet this site in sialylmotif S has been shown in other sialyltransferases to be important to the activity of the enzyme. Clearly, much more of the protein than the catalytic site is relevant to its activity. We propose that the proximity of N308K to the basic polysialyltransferase domain may expand this region on the surface of the protein, resulting in stabilization of the growing polySia chain. Thus N308K may increase the processivity of the enzyme.

This hypothesis would predict that the human enzyme, then, might be able to produce longer polySia chains. Indeed, we demonstrate here that in cell culture the mutation does not appear to drastically alter the overall activity of polySia by the enzyme in N2a cells. However, importantly, preliminary work on polySia-NCAM-Fc secreted from transfected CHO cells, has provided evidence that it does indeed lengthen the synthesized polySia, with the human enzyme producing a slightly longer degree of polymerization.

Notably, this small difference in polySia chain length causes a remarkable difference in its ability to bind trophic factors BDNF and FGF-2. This may result from a previously unknown threshold effect, where the binding is much increased in the slightly longer human product, or from differences in glycan branching that cannot be detected with HPLC. As there are examples of signaling effects of both of these factors being modulated by their interaction with polySia, this finding hints at mechanisms by which N308K may be functionally significant *in vivo*. We are
currently working to replicate these early results, and to investigate the implications that N308K may have on growth, migration, and synaptic plasticity modulated by these factors.

Efforts to confirm these results and to better characterize these enzymes in vitro revealed a new finding: dramatic instability of both enzymes, notably much exaggerated in the human enzyme. As this residue is not glycosylated or integral to the structure of the molecule, the mechanism producing this instability is not clear at present. Experiments are now underway to formally compare the stability of the two enzymes, including assays of enzymatic activity after preincubation at 37°C and tests of the thermal stability of the two proteins.

In understanding the role that this mutation may have played in human evolution, it is important to also consider the finding that ST8SIA2 gene expression appears to be higher in the developing brain of humans than in that of other primates. This may be in part a compensatory mechanism to counteract the loss of stability of the encoded protein. Taken together, however, the increased expression and decreased stability of ST8Sia-II may allow for more rapid regulation, and specifically rapid downregulation, of the protein and the produced polySia.

The idea that rapid up- and down-regulation of polySia in vivo may be necessary or advantageous is still speculative, since it is not yet fully clear how dynamic polySia expression can be. There are indications, though, that polySia levels can fluctuate in response to various stimuli and induce consequent changes. For example, it is well established that polySia is upregulated in response to various central nervous system injuries. Further, its presence or absence influences dendritic spine remodeling in interneurons, and alterations in the levels polySia in the hypothalamus have been associated with the onset and termination of lactation. Such processes may benefit from the ability to more tightly and more dynamically alter polySia expression.

The possible interaction of polySia with Siglec-11 creates a new interesting avenue to explore. Siglec-11 is expressed on microglia in human, but not chimpanzee brains, and it has been proposed to protect neurons from inflammatory and phagocytic actions of microglia. If, like BDNF and FGF-2, Siglec-11 binds polySia-NCAM-Fc from human ST8Sia-II better than that
from the chimp enzyme, N308K may affect inflammation or pruning in human development. Perhaps the two changes together allow fine-tuned modulation of microglial processes not possible in Siglec-11-negative chimpanzee brains.

The confirmation of these ideas and the better characterization of the effects of N308K properly require a controlled environment in which to study the mutation in vivo. Accordingly, we are currently developing a mouse line that contains this single mutation within its genome. This allows careful evaluation of this particular change without the numerous other changes that have occurred since humans and chimpanzees have diverged.

This project is still ongoing, and future work will be necessary to properly confirm the initial indications that N308K produces instability of the enzyme along with an increase in the length of the synthesized polySia product. Regardless, this early work provides interesting evidence that polySia and its synthetic enzymes may have played important and previously unexplored roles in brain evolution.

REFERENCES


CHAPTER 6

Conclusions and Future Directions
The work completed in this thesis explores the roles of sialic acids in vertebrate brain evolution. The studies described here have focused on multiple specific areas of sialic acid biology, from an individual sugar, N-glycolylneuraminic acid (Neu5Gc); to a particular unusual sialylated glycoprotein, polysialylated neural cell adhesion molecule (polySia-NCAM); to the polysialyltransferases that build and the sialidases that digest that glycoprotein.

My initial and major focus has been on the evolutionarily conserved absence of the common sialic acid Neu5Gc from the vertebrate brain. I have approached this from two angles, identifying a possible biochemical mechanism by which Neu5Gc may be detrimental to the brain, and developing a murine model to characterize the phenotypes that result from transgenic overexpression of Neu5Gc in that same tissue.

As we confirm in Chapter 3, Neu5Gc is expressed at extremely low levels in the vertebrate brain throughout all species thus tested, found at less than 3% of total sialic acids. Immunohistochemical analysis reveals Neu5Gc to be present very sparsely within the brain tissue of numerous vertebrates. In the case of mouse, the localization of Neu5Gc to endothelial cells in the brain vasculature is so complete that it appears that Neu5Gc may be completely absent from the parenchyma of the brain itself (Chapter 4).

These findings make a good initial case for selective pressure maintaining exclusion of Neu5Gc from the brain. As a candidate mechanism causing this selective pressure, in Chapter 3, we demonstrated that α2-8 linked Neu5Gc exhibits a remarkable resistance to degradation by both vertebrate and bacterial sialidases. This finding appears to be specific to the α2-8 linkage, as the enzymes do not differentiate between Neu5Ac and Neu5Gc in a nonspecific α-linkage to 4-methylumbelliferone. Further, since the resistance is not present in conditions of acid-catalyzed hydrolysis, it presumably involves some mechanism specific to the catalytic activity of sialidase. Since this linkage is enriched in brain, and a single terminal Neu5Gc could dramatically affect the entire underlying molecule.

This result may help to explain the exclusion of Neu5Gc from the vertebrate brain, but it has additional broader implications for sialic acid metabolism. The difference between Neu5Ac
and Neu5Gc in this context is surprising. The molecules differ by only a single oxygen atom, and yet clear differences appear between them. It is not unusual for proteins involved in sialic acid metabolism to exhibit a relative binding preference for Neu5Ac or Neu5Gc, but this study demonstrates a structural difference in the actual conformations of the molecules, with consequent biological effect. This work is the first, to our knowledge, to demonstrate a biophysical difference between these two very similar molecules. It remains to be seen whether the replacement of Neu5Ac with Neu5Gc, or vice versa, has biophysical effects in any other contexts. Such effects would be particularly relevant for the few evolutionary cases, including humans, in which endogenous Neu5Gc has been lost.

The results of the molecular dynamics modeling, which demonstrate that certain conformations are completely inaccessible to Neu5Gc polymers, are particularly interesting. Importantly, these conformations are accessible to sialidase-sensitive Neu5Pr polymers. Given then that Neu5Gc has a hydrophilic hydroxyl at the position where Neu5Pr has a hydrophobic methyl group, the conformational restriction and resistance of Neu5Gc is likely to result from hydrogen bonding. Unfortunately, neither the molecular dynamics modeling nor initial NMR studies of synthetic Neu5Gc-containing polymers revealed any significant differences in hydrogen bonding. Further efforts to better characterize atomic interactions may be useful in revealing the biochemical mechanism that restricts terminal α2-8 linked Neu5Gc from accessing certain conformations.

The idea that α2-8 linked Neu5Gc may conformationally differ from Neu5Ac or Neu5Pr is supported by the studies of breakdown by acid-catalyzed hydrolysis. The relative resistance to enzymatic hydrolysis is abolished when the reaction is instead catalyzed by environmental protons. α2-8-linked Neu5Gc, then, is not ultimately resistant to the hydrolysis reaction itself, but presumably cannot form the appropriate conformer necessary for an enzyme to bind and/or cleave the molecule. Although we demonstrated enzymatic resistance only for three vertebrate and one bacterial sialidase, the common mechanism used by sialidases may mean that this finding extends throughout many enzymes.
What implications might a broad sialidase resistance of α2-8 linked Neu5Gc have for mammals? This linkage is enriched in brain tissue and relatively rare in other tissues, the exact opposite of the distribution of Neu5Gc-containing glycans, and so α2-8 linked Neu5Gc is not expected to appear in normal brain tissue. This supports the idea that α2-8 linked Neu5Gc should be avoided in biological systems. There is, though, an ever-expanding array of nonneural tissues identified as expressing polySia—for example, dendritic cells, myoblasts, and placenta. Many of these tissues do express Neu5Gc, but its presence in polySia is unexplored. If Neu5Gc does in fact appear in an α2-8 linkage in these tissues, this provides a convenient natural experiment in which to study the consequences of its sialidase resistance.

The only place that α2-8 linked Neu5Gc is conclusively known to be present is the capsule of the eggs of some salmonid fish, where polysialylated glycoproteins containing both Neu5Ac and Neu5Gc have been identified. Notably, this protective covering of the egg should not be desialylated; in fact, it would be preferable to the developing embryo to avoid it. Other rare polysialosyl linkages have been identified in marine species. For example, α2-5 linked O-glycolylated Neu5Gc is found in sea urchin eggs, and an unusual sialidase highly active on these polymers has been identified in oyster hepatopancreas. The resistance of α2-8 linked Neu5Gc to sialidase, then, may be more relevant in species that would benefit from digesting it. Indeed, it would be interesting to see whether some such species have developed other unusual sialidases with higher specificity for α2-8 linked Neu5Gc, or whether the observed resistance indeed impacts some universal constraint of sialidases.

Returning to our original hypothesis, does the resistance of α2-8 linked Neu5Gc to sialidase digestion explain the exclusion of Neu5Gc from the vertebrate brain? Much work remains to be done to answer this question. First, we need to demonstrate that vertebrate sialidases are necessary for the degradation of polySia in vivo. Second, we need to show that impairment of this process by Neu5Gc has negative consequences upon which selection may act.

One study exploring the endogenous degradation of polySia implicated NEU4 as the responsible enzyme. Yet while the authors showed it can degrade polySia (similar to our
demonstration that NEU1 can perform the same reaction), they did not demonstrate that it actually occurs in brain tissue. We have attempted to pursue the question of how polySia is degraded in vivo in NEU1 and NEU4 deficient mice (discussed in Chapter 7, Appendix). Our results demonstrate a very slight increase in total polySia in both NEU1 and NEU4 deficient mice, perhaps indicating that both enzymes are necessary for its proper degradation in vivo. However, these preliminary results require more sensitive analysis of rapid turnover of polySia and a study of earlier embryonic time points to conclusively determine the importance of NEU1 and NEU4 in polySia breakdown.

Importantly, even if we demonstrate that vertebrate sialidases are required to properly degrade polySia, and that Neu5Gc does critically impair this process, we still cannot rule out additional reasons that Neu5Gc is excluded from the vertebrate brain. To properly address other phenotypes produced by Neu5Gc, then, we need a model that overexpresses Neu5Gc in brain.

The murine model of CMAH and Neu5Gc overexpression, discussed in Chapter 4, provides such an opportunity. What are the detrimental effects of neural Neu5Gc—polySia-related or otherwise—that could be responsible for negative selection against its expression? In Chapter 4, I discussed the current status of our numerous attempts to overexpress CMAH and Neu5Gc in the murine brain. Initial efforts were strongly indicative of embryonic lethality, with no transgenic mice surviving until birth, and embryonic abnormalities observed at E13.5. Although the subsequent use of the inducible construct CFE-CMAH has allowed us to establish transgenic lines for the overexpression of Neu5Gc, these lines do not replicate the previous abnormalities and embryonic lethality. Indeed, we are able to achieve over 90% Neu5Gc in the adult mouse brain, in mice that are viable, fertile, and grossly normal in appearance.

It is not presently clear why the CMV-Cmah and NSE-Cmah lines generated initially demonstrated embryonic lethality. One hypothesis is that injection of a transgene is followed by a burst of transcription prior to its integration in the genome, a phenomenon that has previously been reported to occur with transgenic microinjection. The initial CMV-Cmah and NSE-Cmah constructs may have permitted transcription of Cmah at this point. On the other hand, the
inducible CFE-CMAH constructs, having not yet encountered a Cre recombinase, would have instead resulted in transcription only of the floxed GFP. If Cmah or Neu5Gc is toxic to a single-cell embryo, then, this could explain the very different findings.

We are pursuing this hypothesis currently by culturing fertilized mouse eggs in the presence of Neu5Gc, using a corresponding concentration of Neu5Ac as a negative control. Previous work has shown that cells in culture can take up Neu5Gc from the culture medium by macropinocytosis and incorporate it into endogenous glycans \(^9\). This experimental set up, therefore, is expected to mimic the conditions produced by transgenic overexpression of Cmah in a fertilized egg. Alternatively, CMAH could be responsible for cellular toxicity through a different mechanism. For example, its activity depends on NADH for the reduction of its cofactor cytochrome b\(_5\) \(^10\). If the approximately 1000 copies of the injected transgenic construct cause a large burst of CMAH activity, this could dramatically deplete the cell’s stores of NADH, potentially affecting the redox state of the cell and the availability of NADH for other essential reactions.

This is a fascinating new line of questioning on the roles of sialic acids in very early embryonic development. However, it does not address our original question of why Neu5Gc is excluded from the vertebrate brain. It is clear from the surviving CFE-CMAH mice, with very high expression of neural Neu5Gc, that Neu5Gc is not grossly toxic to the brain within the time frame tested. The question of the reason for its conserved suppression in brain remains unanswered.

Using the CFE-CMAH inducible lines, we can investigate phenotypes and test hypotheses as to why Neu5Gc is absent from the brain. Comprehensive screening tools are a useful place to start in investigating potential phenotypes of the mice. We are currently pursuing anatomic analysis of brain tissue, to determine whether any structural abnormalities are present. This will be coupled with cognitive and behavioral analysis of the mice, which can help to elucidate the functional significance of any abnormalities detected, as well as to reveal additional problems that are not associated with obvious structural defects. We are further considering a microarray analysis of gene expression in the brain tissue of wild type and CFE-CMAH transgenic
mice. Such a study will be useful in understanding both the consequences and the compensatory changes that result from Neu5Gc overexpression.

Our existing hypotheses and knowledge of sialic acid metabolism in the brain predict certain findings within these general screens. If the hypothesis discussed in Chapter 3 is correct, i.e. that Neu5Gc presence interferes with appropriate degradation of polySia, we expect there to be enrichment of terminal alpha2-8 linked Neu5Gc in polySia. This is easily tested by using sodium periodate to modify terminal sialic acids, a process that slightly alters their elution time on a C18 HPLC column. Total hydrolysis of glycans followed by DMB-HPLC analysis will then reveal whether more Neu5Gc than expected is present at this terminal position. Further, we may see phenotypes that are the opposite of those already characterized in ST8Sia-II and ST8Sia-IV knockout models. For example, hyperplasia, rather than the hypoplasia found in knockouts, of tracts such as the anterior commissure or corpus callosum may be present. Since diseases such as depression and schizophrenia are associated with dysregulated (typically decreased) polySia (see Chapter 2), interesting phenotypes may appear in behavioral screens.

Additionally, we expect these mice to exhibit severe problems with myelination. Overexpression of polySia in mouse brain causes delayed and impaired myelination. Further, myelin-associated glycoprotein (MAG, Siglec-4), which recognizes Neu5Ac-containing gangliosides, is unable to bind to ligands containing Neu5Gc. We therefore expect to see progressive phenotypes of axonal degeneration and abnormal motor activity as the mice age, similar to those seen in Mag-/- mice. As this phenotype is expected to appear relatively late in life, and since a small percentage of Neu5Gc is unlikely to significantly impair MAG function, however, we do not think it a likely explanation for Neu5Gc exclusion from the brain.

Ultimately, our hypothesis predicts that at least one abnormal phenotype of Neu5Gc in the brain must be acted upon by negative selection in order to preserve the suppression seen throughout all vertebrates. This requires that reproductive fitness be impaired. It appears that neural Neu5Gc does not cause death before reproductive age; presumably, therefore, it causes a reduction in fertility by some other mechanism. We are currently investigating this question by
examining fertility of CFE-CMAH Neu5Gc-overexpressing mice compared to their wild type littermates. If even a minor reduction in fertility is found, negative selection can maintain the exclusion of Neu5Gc from the brain across vertebrates. Of course, reproductive fitness involves an enormous number of individual processes. For example, reproductive hormone secretion, interest in mating behavior, social interaction, motor coordination, and instinctive maternal care could all possibly affect reproductive fitness. Even after identifying a deficit, pinning it to a molecular mechanism poses a further challenge. The screening techniques discussed above will be essential in sorting out this problem. Nevertheless, studies of fertility will be very informative in demonstrating whether there is a selective disadvantage of brain Neu5Gc expression.

With the development and subsequent survival of the CFE-CMAH mouse lines, it became clear that my initial study of the absence of Neu5Gc from the vertebrate brain was a much larger project than it had appeared. Having become interested in polySia, then, I additionally pursued other implications of this molecule in vertebrate brain evolution. Specifically, I wondered whether it has any significance within the human context. Given that polySia is such an important molecule with diverse roles in brain development, it is an interesting glycan to consider within the context of evolutionary changes in the human brain. In Chapter 5, I examined the effects of a human-specific mutation in one of its biosynthetic enzymes, ST8Sia-II. The mutation, N308K, is found in a domain, sialylmotif S, that is predicted to be important for protein function, and at a residue that is highly conserved from chimpanzee to Xenopus.

We are characterizing the effects of this mutation on the structure and function of the protein, as well as its significance to the human species. Work on this question is still ongoing. We have been able to use homology modeling to locate the mutation on the surface of the protein, close to the basic polysialyltransferase domain but relatively distant from the presumed catalytic site. We have also timed this mutation to at least 500,000 years ago, given its presence in both Neandertal and Denisovan genomes, and found limited evidence suggestive of positive selection in the genome.
There is early evidence from studies of ST8Sia-II-mediated polysialylation of NCAM that N308K lengthens the synthesized polySia, as well as destabilizes the enzyme itself. These initial results are promising, but need to be validated and confirmed in further experiments. Specifically, the CHO cell transfectants initially used to analyze the length of synthesized polySia-NCAM-Fc are not optimal in that they have quite different levels of expression of NCAM-Fc. Ideally, human and chimp ST8Sia-II should be compared in conditions as close to identical as possible. We are therefore planning to recreate these lines, possibly using an internal ribosomal entry site (IRES) to ensure equivalent expression of both NCAM-Fc and ST8SIA2 within each cell. Additionally, DMB-HPLC and other techniques available to analyze polySia length to date all use acid, which releases the α2-8-linked chain from the underlying structure, but can also cause internal cleavage of the fragile polymer\textsuperscript{15}. Therefore, we are also investigating anion-exchange and electrophoretic techniques to evaluate chain length of polySia without requiring acidic conditions.

We are further formally assessing the decreased stability we detected in soluble human ST8Sia-II protein. Specifically, we are testing the activity of the human and chimp enzymes after storage at various temperatures, to evaluate differences in their thermal stability. Pulse-chase analysis will also be pursued to study the turnover of the enzyme in cell culture.

If our initial results are confirmed in later experiments, N308K has very interesting implications for human evolution. Extensive work to date has examined the effects of the presence of absence of polySia in a number of anatomical areas or cell types; however, the effects of minor differences in polySia length are completely unknown. The remarkable increase in binding to BDNF and FGF-2 found in the slightly longer polySia synthesized by human ST8Sia-II suggests an underlying threshold effect. The downstream phenotypes resulting from N308K, then, may lie in the interaction of these molecules. Diverse but important processes such as growth, migration, and synaptic plasticity have all been implicated as involving the interaction between polySia and BDNF or FGF-2.

Siglec-11, for which polySia is a putative ligand, may also be affected by N308K. Notably, Siglec-11 has human-specific expression in microglia\textsuperscript{16}. Although its function there is not well
understood, it has been implicated in polySia-dependent inhibition of microglial neurotoxicity\textsuperscript{17}. We have not yet examined whether Siglec-11 also demonstrates differential binding to polySia synthesized by human and chimp ST8Sia-II. If the results are similar to those seen with BDNF and FGF-2, this has interesting implications for human evolution. Perhaps the longer polySia resulting from N308K, in combination with the emergence of neural Siglec-11 expression, allow human-specific regulation of microglial pruning, phagocytosis, or inflammation within the brain.

The combination of decreased enzymatic stability and a longer polySia product caused by N308K may be relevant to the regulation of polySia in the brain. Perhaps the mutation allows for tighter up- and down-regulation of polySia within the developing brain. Interestingly, this speculation that rapid turnover of polySia is an important developmental process has been raised by both major portions of my thesis. If it does occur, it may be regulated in part by the removal of polySia itself from the cell surface with a vertebrate sialidase such as NEU1 (Chapter 3), and by rapid downregulation of ST8Sia-II (Chapter 5). To put these studies in physiological context, it will be important to verify whether such rapid downregulation does in fact occur \textit{in vivo}.

Evidence on this point is limited. A recent study has demonstrated that dendritic spines on interneurons can be dynamically regulated by the presence or absence of polySia\textsuperscript{18}. Additional work from the laboratory of Ken Kitajima has discovered rapid loss of surface polySia on microglia in response to lipopolysaccharide stimulation (unpublished). No study has yet revealed rapid regulation of polySia occurring \textit{in vivo}, however. Nevertheless, this work is exciting for its indication that neural sialic acid metabolism may be a highly dynamic process, part of the brain’s response to stimuli or its ability to maintain plasticity.

It is worth mentioning that in our study of ST8Sia-II to date we exclusively focus on NCAM, ignoring another, less common polysialylated neural molecule, SynCAM 1. This has been only recently characterized as a polySia carrier protein, and its function is so far unknown\textsuperscript{19}. It is expressed on a subset of glial progenitor cells called NG2 cells, where it is speculated to modulate synaptic formation by inhibiting SynCAM 1 adhesion. Importantly, unlike NCAM, SynCAM 1 is exclusively polysialylated \textit{in vivo} by ST8Sia-II, not by ST8Sia-IV\textsuperscript{20}. Modification to
ST8Sia-II function may therefore also affect polySia-SynCAM 1 function. As more is learned about the function of polySia on SynCAM 1, the possible implications of N308K on this molecule may become clearer.

As with any line of scientific inquiry, the work completed so far only reveals more that remains to be done, both in exploring the reasons Neu5Gc is excluded from the vertebrate brain, and in understanding the effects of the human-specific mutation N308K on the ST8Sia-II. Sialic acid is a molecule with numerous evolutionary lines of significance. Through the course of my thesis work described here, I have explored this idea within the brain, studying the roles of neural glycans, sialyltransferases, and sialidases. The work presented here, and further studies examining these glycans within an evolutionary context, reveal new ways in which sialic acid plays diverse and important roles in vertebrate brain development.

REFERENCES


APPENDIX

Changes in Developmental Expression of Polysialic Acid in Sialidase Deficient Mice
ABSTRACT

We have previously demonstrated that Neu5Gc in an α2-8 linkage, such as that found in the important brain glycoconjugate polysialic acid (polySia), is relatively resistant to degradation by vertebrate sialidases. This finding may help explain the requirement for the evolutionarily conserved exclusion of Neu5Gc from the vertebrate brain. This hypothesis requires that sialidase degradation of polySia be necessary in development. Very little is known about the in vivo turnover of polySia; however, the vertebrate sialidases NEU1 and NEU4 are the strongest candidates for catalyzing its hydrolysis. Here we examine the total polySia concentration in early postnatal life of mice with 10% of normal NEU1 activity or complete loss of NEU4 activity. Expression was normalized to NCAM, the predominant carrier protein of polySia. The expression of polySia from P1 to P20 in sialidase mutants is similar, although slightly higher, than that of wild type mice, indicating that the majority of postnatal loss of polySia does not depend on these two enzymes. Further work is necessary to understand the turnover of individual polySia chains and the potential role of vertebrate sialidases in this process.

INTRODUCTION

In Chapter 3, I described a hypothesis for the conserved absence of Neu5Gc from the vertebrate brain: that α2-8 linked Neu5Gc, such as that present in polysialic acid, causes a resistance of the glycan to degradation by sialidase. Although the relative resistance is clearly demonstrated in vitro in this study, the hypothesis relies on unproven assumptions about the normal degradation of polySia in vivo. For the relative resistance to be physiologically relevant, there must be a requirement for rapid turnover of polySia. Further, this degradation must rely on the activity of sialidases, possibly at the cell surface, rather than on hydrolysis under intracellular conditions.

Very little is known about the degradation of polySia in vivo. There are a few indications that rapid turnover of surface polySia may be useful in development. For example, sudden alterations in polySia can cause remodeling in the neurites of interneurons. Additionally,
unpublished evidence indicates that microglia rapidly lose surface polySia in response to stimulation by lipopolysaccharide (LPS), although the underlying mechanism of this is unknown (Kitajima lab). However, no studies have yet examined the turnover of endogenous polySia within the brain.

There are four vertebrate sialidases, NEU1 – NEU4, that are potentially responsible for polySia breakdown. NEU1 is the most highly expressed sialidase in the brain. It is typically localized to the lysosome, but a growing body of evidence suggests that NEU1 can move to the cell surface via exocytosis to desialylate cell surface glycans. It has additionally been suggested that NEU4 is responsible for polySia degradation. This sialidase is predominantly expressed in brain, albeit at levels still much lower than NEU1.

If breakdown of polySia by either of these sialidases is highly prevalent, we might expect that their loss would increase the total cell-surface polySia present in the brain. We therefore examined the total expression of polySia by western blot in mice that were deficient in NEU1 (10% activity) or completely lacking NEU4.

METHODS

Mice

The brains of mice were collected in the laboratory of Alexey Pshezhetsky (University of Montreal). Three mouse strains were tested in addition to wild type. CathAS190A-Neo (NEU1 kd) has a pGK-Neo cassette in the cathepsin A gene, resulting in 10% of normal NEU1 activity. Neu4 knockout mice (Neu4 -/-) completely lack the sialidase NEU4. Lastly, mice that had both of these mutations in addition to a knockout of hexosaminidase A (HexA) (NEU1 kd; Neu4 -/-; HexA -/-). Brains were collected at postnatal days 1, 5, 10, 15, and 20 and snap-frozen. On thawing, they were homogenized in lysis buffer (20 mM Tris-HCl, pH 8.0, 150 mM NaCl, 5 mM EDTA, 2% Triton X-100, and 1X protease inhibitor cocktail added just before use).

Immunoblot
For analysis by western blot, 100 µg of protein were stored on ice in the presence or absence of endoN (a gift of Rita Gerardy-Schahn) for 45 min. After addition of sample buffer, samples were not boiled but heated at 60°C for 10 min. All were run by SDS-PAGE using a 7% polyacrylamide gel. After transfer, blots were stained with the 12E3 anti-polySia antibody (generously provided by the laboratory of Ken Kitajima) at 1 µg/ml or with the H28 monoclonal NCAM antibody, followed by appropriate LI-COR infrared secondary antibody. All imaging and quantification was done using a LI-COR Odyssey system.

RESULTS

Mice deficient in NEU1 and NEU4 exhibit grossly normal developmental polySia

To determine whether any gross differences in polySia content were present, we examined four strains of mice: wild type, NEU1 knockdown (NEU1 kd, 10% activity), NEU4 knockout (Neu4-/-), and a triple mutant for NEU1 kd, Neu4-/-, and knockout of hexosaminidase A (HexA). HexA-/- mice lack lysosomal beta-hexosaminidase, which normally converts ganglioside GM2 to GM3. In the presence of additional Neu4 deletion, they exhibit a severe neural phenotype mimicking human Tay-Sachs disease ⁴.

Brain tissue was collected from mice at P1, P5, P10, P15, and P20, lysed, and analyzed by immunoblot. Membranes were stained with an antibody to polySia. To properly quantify NCAM expression, an aliquot of each sample was treated with the phage endoneuraminidase endoN, which selectively cleaves polySia, and stained with an antibody to NCAM.

As expected from previously published data, very young mice from all strains exhibited exclusively polysialylated NCAM, progressing to low levels of polySia and predominantly polySia-free NCAM in the first 3 weeks of life (Figure 7-1)⁵. No gross differences are apparent among the four strains.

The stained blots were then quantified and polySia expression was calculated relative to the total amount of NCAM (Figure 7-2). Again, no major differences are apparent among the four strains. As a general trend, however, the wild type mice do have slightly less polySia than the
Figure 7-1. Immunoblots for polySia and NCAM in sialidase deficient mice are similar to wild type. Lysates from the brains of wild type, NEU1 kd, Neu4 -/-, and NEU1 kd; Neu4 -/-; HexA -/- mice were analyzed by immunoblot for NCAM (above) and polySia (below). Each lane represents one individual of the indicated age, ranging from postnatal day 1 (P1) through adult (A). An aliquot of each sample was treated with endoN, which removes polySia from NCAM.
Figure 7-2. Quantification of polySia expression suggests slight increase of polySia in sialidase deficient mice. Quantification of the immunoblots shown in Figure 7-1 was performed using LI-COR imaging software. PolySia was quantified in intact lysates (left). NCAM was quantified in endoN-treated lysates and the total intensity of all three isoforms was summed (center). Intensity of polySia staining was normalized to intensity of NCAM staining (right).
other three genotypes, NEU1 kd and Neu4 -/- are intermediate, and the triple knockout has the highest polySia expression. This is most exaggerated at P1, suggesting that embryonic timepoints may be even more affected by the loss of these sialidases. However, these differences are so slight that conclusions cannot be drawn at present.

DISCUSSION

We show here that mice deficient in NEU1 and NEU4 have very similar postnatal polySia expression to that of wild type. Although there is a very slight indication of excess polySia present in the mutant lines, the current data is not sufficient to conclude a true difference. This suggests that the majority of postnatal polySia degradation in vivo does not depend on these enzymes. This limited work to date, however, leaves many avenues unexplored.

The slightly exaggerated difference between wild type and sialidase mutant mice present at P1 may indicate that a greater difference is present in embryonic life, when polySia is highly expressed during neural development. It would be very interesting to extend this analysis to earlier periods of development to properly address this question.

We have here queried NEU1 and NEU4, the strongest candidates for neural polySia breakdown. There are two remaining vertebrate sialidases that are not examined in this study. NEU2, a cytosolic sialidase, has not been demonstrated to reach the cell surface. NEU3 is localized to the plasma membrane; however, its activity is believed to be exclusive to gangliosides. Although we believe these are unlikely to be involved in polySia breakdown, it is possible that they in fact are, or that they compensate for the lack of NEU1 and NEU4 in the current study.

We are examining the total amount of surface polySia here. However, polySia degradation may be affected by the loss of NEU1 and NEU4 in ways that are not detectable by this method. First, we do not know for certain that polySia is removed directly from the cell surface, but it may be digested after endocytosis by lysosomal NEU1. Lysosomal polySia fragments would not be detectable by immunoblot. Additionally, we have hypothesized that there
may be very rapid cell-surface degradation of polySia by sialidase. This process may not affect the total polySia concentration if it occurs on a small scale or if it requires a particular trigger, such as LPS stimulation.

As discussed in Chapter 4, we are additionally studying the expression of polySia in transgenic mice that overexpress neural Neu5Gc. Since Neu5Gc in polySia is relatively resistant to sialidase degradation (Chapter 3), studies of these mice will also aid our understanding of the roles of endogenous sialidase in polySia turnover.

This initial study provides limited indication that the loss of NEU1 and NEU4 increases total polySia in the developing brain. Further work will be necessary to conclusively demonstrate which vertebrate sialidases, if any, are involved in the degradation of endogenous polySia, and to understand the importance of this process in brain development.

REFERENCES


