

New Tools for Studying *O*-GlcNAc
Glycosylation and Chondroitin Sulfate Proteoglycans
and
Studies on the Roles of *O*-GlcNAc Glycosylation
on the Transcription Factor CREB

Thesis by

Peter Michael Clark

In Partial Fulfillment of the Requirements

For the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

2010

(Defended December 21, 2010)

© 2010

Peter Michael Clark

All Rights Reserved

Acknowledgements

First I would like to thank Prof. Linda Hsieh-Wilson for her support, advice, guidance, and scientific mentorship. I would also like to thank Jessica Rexach, my lifelong collaborator, without whose constant encouragement, support, ideas, (and cooking) much of this work would not have been accomplished. In addition, I would like to thank members of the Hsieh-Wilson lab, past and present, who have made my time at Caltech both enjoyable and productive, and especially those who taught me many of the techniques I still use today: Nelly Khidekel, Nathan Lamarre-Vincent, and Sarah Tully. Finally, I would like to thank my committee members — Prof. Dennis Dougherty, Prof. Peter Dervan, and Prof. Mary Kennedy — for their insightful scientific ideas and advice.

Abstract

The addition and removal of the monosaccharide *N*-acetyl-D-glucosamine (GlcNAc) to serine and threonine residues of proteins has emerged as a critical regulator of cellular processes. However, studies of *O*-GlcNAc in such complex systems as the brain have been limited, in part due to the lack of tools. Here we report the development of new tools for studying *O*-GlcNAc, and the application of these and other tools for studying the roles of *O*-GlcNAc in the brain.

Working from a previously established chemoenzymatic method, we designed an isotopic labeling strategy for probing the dynamics of *O*-GlcNAc glycosylation using quantitative proteomics. With this tool, we show that *O*-GlcNAc is dynamically modulated on specific proteins by excitatory stimulation of the brain *in vivo*. Separately, we improved this chemoenzymatic strategy by integrating [3+2] azide-alkyne cycloaddition chemistry to attach biotin and fluorescent tags to *O*-GlcNAc residues. These tags allow for the direct fluorescence detection, proteomic analysis, and cellular imaging of *O*-GlcNAc modified proteins. With this strategy, we identified over 146 novel glycoproteins from the mammalian brain.

The transcription factor cAMP-response element binding protein (CREB) is critical for numerous functions in the brain, including neuronal survival, neuronal development, synaptic plasticity, and long-term memory. We show that CREB is highly glycosylated in the brain and discover new glycosylation sites on CREB in neurons. One

of these sites is dynamically modulated by neuronal activity and is important for regulating CREB. Removal of this glycosylation site accelerates axon and dendrite development *in vitro* and long-term memory consolidation *in vivo*. These studies are the first demonstration that *O*-glycosylation at a specific site on a specific protein is critical for neuronal function and behavior.

Chondroitin sulfates (CS) are sulfated linear polysaccharides important in neuronal development and viral invasion. Depending on their sulfation patterns, CS molecules differ dramatically in their functions. We developed a computational method to model the structure and function of CS. Using this approach, we show that different CS tetrasaccharides have distinct solution structures. We also modeled the CS binding site on a variety of proteins and discovered that CS may be important in modulating the interaction between specific growth factors and their receptors.

Table of Contents

Chapter 1: Chemical Approaches to Understanding <i>O</i> -GlcNAc Glycosylation in the Brain	... 1
Chapter 2: The Roles of <i>O</i> -GlcNAc Glycosylation in the Brain	... 21
Chapter 3: Probing the Dynamics of <i>O</i> -GlcNAc Glycosylation in the Brain Using Quantitative Proteomics	... 39
Chapter 4: Direct In-Gel Fluorescence Detection and Cellular Imaging of <i>O</i> -GlcNAc-Modified Proteins	... 70
Chapter 5: CREB — A Key Transcription Factor in the Nervous System	... 111
Chapter 6: Dynamic <i>O</i> -Glycosylation Regulates CREB-Mediated Neuronal Gene Expression and Memory Formation	... 128
Chapter 7: Computational Modeling of Glycosaminoglycans	... 165
Appendix I: <i>O</i> -GlcNAc Glycosylation Regulates CREB Activity in Pancreatic Beta Cells	... 210

List of Illustrations and Tables

Chapter 1: Chemical Approaches to Understanding *O*-GlcNAc Glycosylation in the Brain

Figure 1: <i>O</i> -GlcNAc glycosylation is the addition of β - <i>N</i> -acetylglucosamine to serine or threonine residues of proteins	... 2
Figure 2: Strategy for chemically tagging <i>O</i> -GlcNAc proteins	... 4
Figure 3: The BEMAD strategy for mapping glycosylation sites	... 6
Figure 4: Small molecule inhibitors of OGT and OGA	... 11
Figure 5: Chemical tools for monitoring <i>O</i> -GlcNAc dynamics	... 14

Chapter 2: The Roles of *O*-GlcNAc Glycosylation in the Brain

Figure 1: <i>O</i> -GlcNAc glycosylation is the addition of β - <i>N</i> -acetylglucosamine to serine or threonine residues of proteins	... 22
Figure 2: <i>O</i> -GlcNAc proteome from rodent brain	... 27

Chapter 3: Probing the Dynamics of *O*-GlcNAc Glycosylation in the Brain Using Quantitative Proteomics

Scheme 1: QUIC-Tag strategy for quantitative <i>O</i> -GlcNAc proteomics	... 41
Figure 1: Accurate quantification of known <i>O</i> -GlcNAc peptides from complex mixtures using the QUIC-Tag approach	... 43
Table 1a: Mean ratios of individual peptides from α -crystallin and OGT	... 44
Table 1b: Mean ratios of all peptides	... 44
Figure 2: <i>O</i> -GlcNAc glycosylation is reversible in cultured cortical neurons	... 45
Figure 3: Sequencing of tagged <i>O</i> -GlcNAc peptides regulated by	... 46

PUGNAc treatment using CAD	
Figure 4: Sequencing of tagged <i>O</i> -GlcNAc peptides regulated by PUGNAc treatment using ETD	... 49
Table 2: Identification and quantification of changes in <i>O</i> -GlcNAc glycosylation induced by PUGNAc	... 50
Figure 5: Quantification of <i>O</i> -GlcNAc glycosylation on intact proteins by immunoblotting and infrared imaging detection	... 51
Figure 6: <i>O</i> -GlcNAc glycosylation is dynamically modulated by robust excitatory stimulation of the brain <i>in vivo</i> using kainic acid	... 52
Table 3: Identification and quantification of changes in <i>O</i> -GlcNAc glycosylation induced by kainic acid	... 54
Figure 7: Expression levels of EGR-1, GRASP55, and eIF4G following kainic acid treatment of rats	... 54
Table 4: <i>O</i> -GlcNAc glycosylated proteins identified from the cerebral cortex of kainic acid-stimulated rats	... 55
Chapter 4: Direct In-Gel Fluorescence Detection and Cellular Imaging of <i>O</i>-GlcNAc-Modified Proteins	
Figure 1: Chemoenzymatic labeling of <i>O</i> -GlcNAc proteins using [3+2] cycloaddition chemistry	... 72
Figure 2: Selective labeling of α -crystallin	... 73
Figure 3: Detection sensitivity of the chemoenzymatic approach	... 73
Figure 4: Enrichment and in-gel fluorescence detection of <i>O</i> -GlcNAc-modified proteins in 1D (top and middle) and 2D (bottom) gels	... 74
Figure 5: Functional classification of <i>O</i> -GlcNAc proteins from rat brain identified by MS	... 76
Figure 6: Direct detection of changes in <i>O</i> -GlcNAc glycosylation levels upon cellular stimulation	... 76
Figure 7: Detection of <i>O</i> -GlcNAc levels on chromatin	... 78

Figure 8: <i>O</i> -GlcNAc levels are distributed evenly across the mouse chromosomes	... 78
Figure 9: Fluorescence imaging of <i>O</i> -GlcNAc proteins in HeLa cells or cortical neurons	... 79
Table 1. <i>O</i> -GlcNAc glycosylated proteins identified by mass spectrometry	... 81
Table 2: Promoters enriched in <i>O</i> -GlcNAc levels	... 90

Chapter 6: Dynamic *O*-Glycosylation Regulates CREB-Mediated Neuronal Gene Expression and Memory Formation

Figure 1: CREB is <i>O</i> -GlcNAc glycosylated in neurons.	... 131
Figure 2: CREB is <i>O</i> -GlcNAc glycosylated at Ser40 and Thr227 or Thr228.	... 131
Figure 3: CREB glycosylation sites	... 132
Figure 4: CREB is <i>O</i> -GlcNAc glycosylated at Ser40 in response to neuronal activity	... 132
Figure 5: Kinetics of endogenous CREB glycosylation and Ser133 phosphorylation upon depolarization of cortical neurons	... 133
Figure 6: The protein synthesis inhibitor cycloheximide does not block depolarization-induced glycosylation of CREB	... 133
Figure 7: CREB glycosylation is modulated by specific kinase pathways	... 134
Figure 8: Glycosylation is not induced on S133A CREB	... 134
Figure 9: Forskolin induces CREB phosphorylation but not CREB glycosylation in cortical neurons	... 136
Figure 10: Chemoenzymatic labeling of endogenous CREB for visualizing phosphorylation and glycosylation within the same protein molecule and for quantifying the levels of each modification within distinct posttranslationally modified subpopulations	... 136

Figure 11: CREB is simultaneously phosphorylated at Ser133 and glycosylated at Ser40	... 137
Figure 12: Kinetics of Ser133 phosphorylation for specific post-translationally modified subpopulations of endogenous CREB	... 137
Figure 13: Kinetics of glycosylation for specific post-translationally modified subpopulations of endogenous CREB	... 137
Figure 14: Knockdown of endogenous CREB and overexpression of shRNA-resistant, Flag-tagged WT or S40A CREB in neuro2a cells	... 138
Figure 15: Glycosylation at Ser40 represses CREB activity	... 139
Figure 16: Glycosylation at Ser40 represses CREB activity	... 139
Figure 17: WT and S40A CREB show similar binding to the CRE promoter <i>in vitro</i>	... 139
Figure 18: Glycosylation at Ser40 blocks the interaction between CREB and CRTC2	... 139
Figure 19: Glycosylation at Ser40 represses CREB activity via a CRTC-dependent mechanism in neurons	... 140
Figure 20: S40A CREB produces approximately a 2.5–3.6-fold increase in mRNA expression relative to WT CREB	... 141
Figure 21: Both CREB and OGT occupy the <i>BDNF</i> exon IV, <i>Arc</i> , <i>Cdk5</i> , <i>c-fos</i> , and <i>Wnt2</i> promoters	... 142
Figure 22: WT and S40A CREB show comparable levels of promoter occupancy on <i>BDNF</i> exon IV, <i>Arc</i> , <i>Cdk5</i> , <i>c-fos</i> , and <i>Wnt2</i> promoters	... 142
Figure 23: S40A CREB regulates a subset of CREB-mediated genes	... 142
Figure 24: Blocking glycosylation enhances <i>BDNF IV</i> and <i>c-fos</i> expression even more after depolarization	... 143
Figure 25: WT and S40A Flag-tagged CREB are expressed at similar levels in cortical neurons	... 143

Figure 26: CREB glycosylation at Ser40 represses dendritic growth	... 144
Figure 27: Expression of S40A CREB enhances axon outgrowth	... 144
Figure 28: OGT expression represses while OGT knockout enhances axon length	... 145
Figure 29: OGT knockout decreases <i>O</i> -GlcNAc levels, and OGT overexpression increases <i>O</i> -GlcNAc levels	... 145
Figure 30: CREB glycosylation at Ser40 represses dendritic growth through overexpression on Wnt2	... 146
Figure 31: Depolarization-induced dendrite growth is blocked by Wnt2 siRNA, Dkk-1, and Ncad(intra)	... 147
Figure 32: The enhanced axon outgrowth caused by expression of S40A CREB is blocked by TrkB-Fc or CRTCI knockdown	... 148
Figure 33: Herpes simplex virus (HSV) infection of the lateral amygdala of mice	... 150
Figure 34: CREB glycosylation at Ser40 modulates long-term conditioned fear memory	... 150
Figure 35: Inhibition of protein synthesis immediately after training blocks the memory enhancement of S40A CREB at 2 h	... 150
Figure 36: Inhibition of protein synthesis 2 h after training blocks the memory enhancement of WT CREB at 24 h, but not that of S40A CREB	... 151
Figure 37: CREB glycosylation is induced following activation of neurons <i>in vivo</i>	... 151
Table 1: Primers used for RT-PCR	... 159
Table 2: Primers used for ChIP PCR	... 159
 Chapter 7: Computational Modeling of Glycosaminoglycans	
Figure 1: CS-A, CS-C, CS-E, and CS-R tetrasaccharides	... 169

Figure 2: <i>Top:</i> The lowest energy structures of CS-A, CS-C, CS-E, and CS-R tetrasaccharides. <i>Bottom:</i> Electrostatic representations of these structures	... 172
Figure 3: a, Predicted heparin binding site of FGF-2 from 1BFB using default ScanBindSite parameters. b, Predicted heparin binding site of FGF-2 from 1BFB using modified ScanBindSite parameters	... 173
Figure 4: a, Heparin binding site on FGF-2 (from 1BFB). b, Predicted heparin binding site (from 1BFB crystal structure). c, Predicted heparin binding site (from 1BLA crystal structure)	... 174
Table 1: Residues that interact with heparin from the 1BFB crystal structure and in the predicted heparin binding site from FGF-2 in the 1BFB and 1BLA crystal structure	... 175
Figure 5: a, Heparin binding site on FGF-1 (from 2AXM). b, Predicted heparin binding site	... 177
Table 2: Residues that interact with heparin from the 2AXM crystal structure and in the predicted heparin binding site from FGF-1 in the 2AXM crystal structure	... 178
Figure 6: a, Residues important for CS-A binding to DLB3X (left) and DLB6 (right), as previously determined by mutagenesis experiments. b, Predicted CS-A binding site on DLB3X (left) and DLB6 (right)	... 179
Table 3: Predicted CS-A binding site on DBL3X and residues experimentally determined as important for CS-A binding	... 180
Table 4: Predicted CS-A binding site on DBL6 and residues experimentally determined as important for CS-A binding	... 180
Figure 7: a, CS-A binding site on Cathepsin K (from 3C9E). b, Predicted CSA binding site	... 181
Table 5: Residues that interact with CS-A from the 3C9E crystal structure and in the predicted CS-A binding site from Cathepsin A in the 3C9E crystal structure	... 182
Figure 8: a, Crystal structure of TNF trimer (from 1TNF). b, Predicted CS-E binding site on TNF trimer. c, Overlay between predicted CS-E binding site (slate) and predicted TNF-R1 (green, from 1TNR) complex.	... 183

d, Overlay between predicted CS-E binding site (slate) and residues important for TNF-R2 (green) binding

Figure 9: **a**, BDNF monomer crystal structure (from 1BND). **b**, CS-E binding site (slate). **c**, Homology model of BDNF dimer crystal structure (wheat and cyan). **d**, CS-E binding site (slate) ... 184

Figure 10: **a**, NT-3 monomer crystal structure (from 1BND). **b**, CS-E binding site (slate). **c**, NT-3 dimer crystal structure (from 1NT3). **d**, CS-E binding site (slate) ... 185

Table 6: Predicted CS-E binding sites on BDNF, NGF, NT-4/5, and NT-3 ... 186

Figure 11: Cationic amino acids (yellow) in the respective CS-E binding sites CS-E binding site comparison ... 187

Figure 12: **a**, NGF monomer crystal structure (from 2IFG). **b**, CS-E binding site (slate). **c**, NGF dimer crystal (from 2IFG). **d**, CS-E binding site (slate) ... 187

Figure 13: **a**, NT-4/5 monomer crystal structure (from 1HCF). **b**, CS-E binding site (slate). **c**, NT-4/5 dimer crystal (from 1HCF). **d**, CS-E binding site (slate) ... 188

Figure 14: CS-E binding sites on the Trk family of receptors. **a**, TrkA. **b**, TrkB. **c**, TrkC ... 190

Table 7: Predicted CS-E binding sites on TrkA, TrkB, and TrkC ... 191

Figure 15: CS-E binding sites for the neurotrophins (slate) and Trk receptors (green), projected onto the neurotrophin (wheat) - receptor (cyan) complexes. **a**, NGF–TrkA. **b**, BDNF–TrkB. **c**, NT-4/5–TrkB. **d**, NT-3–TrkC ... 192

Figure 16: **a**, Midkine (from 1MKN). **b**, Predicted CS-E binding site (slate) ... 194

Table 8: Residues that interact with heparin as previously determined by NMR and the predicted CS-E binding site ... 194

Figure 17: **a**, GDNF crystal structure (from 2V5E). **b**, Predicted CS-E binding site on GDNF from **a** (slate and yellow). **c**, GDNF crystal structure (from Chain D 3FUB). **d**, Predicted CS-E binding site on GDNF ... 196

(slate and yellow) from **c**. Residues predicted to be in the CS-E binding site for both GDNF crystal structures are yellow; those predicted to be in the CS-E binding site for only one GDNF crystal structure are colored slate

Figure 18: **a**, GDNF protein–receptor complex (wheat and cyan, ... 197 respectively, from 3FUB). **b**, Predicted CS-E binding site (slate). **c**, Binding site of heparin mimic sucrose octasulfate (slate) as determined from the co-crystal

Table 9: CS-E binding site on GDNF monomer (3FUB, 2V5E ... 198 crystal structure), GDNF receptor, and Nogo

Figure 19: CS-E binding sites (slate) on GDNF (wheat) and CS-E binding ... 199 sites (yellow) on the GDNF receptor (cyan) as mapped onto the GDNF–GDNF receptor complex (3FUB)

Figure 20: **a**, Nogo receptor crystal structure (from 1P8T). **b**, CS-E ... 200 binding site (slate)

Appendix I: O-GlcNAc Glycosylation Regulates CREB Activity in Pancreatic Beta Cells

Figure 1: Hyperglycemia induces CREB glycosylation ... 214

Figure 2: CREB is dynamically glycosylated at Ser40 in HIT-T15 cells ... 216

Figure 3: O-GlcNAc glycosylation represses CREB transcriptional ... 217 activity

Figure 4: The *IRS-2* gene is a direct target for regulation by CREB and ... 219 OGT