

2014

Expression of O-linked N-acetylglucosamine modified proteins and production and characterization of chlamydia trachomatis CT663

Octavia Y. Goodwin

Louisiana State University and Agricultural and Mechanical College, ogoodwi1@gmail.com

Follow this and additional works at: https://digitalcommons.lsu.edu/gradschool_dissertations



Part of the [Chemistry Commons](#)

Recommended Citation

Goodwin, Octavia Y, "Expression of O-linked N-acetylglucosamine modified proteins and production and characterization of chlamydia trachomatis CT663" (2014). *LSU Doctoral Dissertations*. 3325.
https://digitalcommons.lsu.edu/gradschool_dissertations/3325

This Dissertation is brought to you for free and open access by the Graduate School at LSU Digital Commons. It has been accepted for inclusion in LSU Doctoral Dissertations by an authorized graduate school editor of LSU Digital Commons. For more information, please contact gradetd@lsu.edu.

EXPRESSION OF O-LINKED N-ACETYLGLUCOSAMINE
MODIFIED PROTEINS AND PRODUCTION AND
CHARACTERIZATION OF CHLAMYDIA TRACHOMATIS CT663

A Dissertation

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

in

The Department of Chemistry

by
Octavia Yvonne Goodwin
B. S., Georgia Southern University, 2008
December 2014

*To my best friend and other half, Steven Calloway, for his patience, love,
encouragement, and faith in me.*

*To my mother, Judy Goodwin, for her unconditional love and support over the
years.*

To my grandmother, Inez Walker, for always making my heart smile.

To my brother, Jeffery Bolton, for always believing in me.

To my entire family past, present, and future.

Acknowledgements

I would like to thank my advisor, Dr. Macnaughtan, for guidance, support, and faith in my abilities. I also would like to thank my committee members, Dr. Isiah Warner, Dr. Samuel Gilman, and Dr. Aaron Smith for their guidance through this process.

I would like to thank the Macnaughtan research group for their friendship and support. Amid Paudyal, you have been so kind and helpful. Maggie Thomasson, you are one of the best writers I know and I appreciate your willingness to proofread. Dr. Kevin Roberson, although you are as stubborn as an ox, you have been one of my best friends and I appreciate you. I would like to say a special thanks to my “skinny new post doc”, Dr. Michelle Sweeney, you have taught me so much about being professional and you have been an exceptional friend.

I would also like to thank my undergraduate advisor, Dr. Karen Welch. You believed in me when I did not believe in myself. I would have never been brave enough to apply for graduate school without you. You were my first lab mother and I will always love you.

Finally, I would not have made it through this process without my mother, grandmother, sisters, and my best friend. Judy Goodwin, Inez Walker, Nikkie Goodwin, Lakiesha Goodwin, Atiya Jordan, and Steven Calloway you are the wind beneath my wings. I love you all to pieces.

Table of Contents

Acknowledgements.....	iii
List of Tables	viii
List of Figures	ix
List of Abbreviations.....	xiv
Abstract.....	xv
Chapter 1: Background	1
1. 1 Proteins	1
1. 2 Protein Structure.....	2
1. 3 Protein function as it relates to structure.....	4
1. 4 Post translational modifications	4
1. 5 O-GlcNAc in the hexosamine biosynthesis pathway.....	5
1. 6 O-GlcNAc is not your typical glycosylation	6
1. 7 O-GlcNAc is a regulatory enzyme that is implicated in disease when it is dysregulated	8
1. 8 Inhibitors of hexosaminidase	10
1. 9 Methods for O-GlcNAc detection	11
1. 9. 1 Galactosyltransferase-based methods	11
1. 9. 2 Lectin blot	11
1. 9. 3 Western blot.....	12
1. 9. 4 Chemical Derivatization – BEMAD	12
1. 10 Methods to study O-GlcNAc-modified proteins.....	14
1. 11 Challenges.....	14
1. 12 <i>In vivo</i> expression of O-GlcNAc-modified protein	14
1. 13 Molecular cloning and gene expression.....	15
1. 14 Recombinant protein expression and purification	16
1. 15 Co-expression of target proteins with OGT.....	17
1. 16 NMR	18

Chapter 2: <i>E. coli</i> sabotages the <i>in vivo</i> production of O-linked-N-acetylglucosamine (O-GlcNAc) modified proteins.....	21
2. 1 Introduction.....	21
2. 2 Materials and methods	22
2. 2. 1 Materials	22
2. 2. 2 Cloning and expression vectors.....	24
2. 2. 3 <i>E. coli</i> cultures and protein co-expression	25
2. 2. 4 O-GlcNAc detection and quantification	26
2. 2. 5 Partial purification of <i>E. coli</i> β -N-acetylglucosaminidase.....	27
2. 2. 6 <i>E. coli</i> β -N-acetylglucosaminidase detection assays	28
2. 2. 7 Kinetic assays.....	29
2. 2. 8 Kinetic analysis	30
2. 2. 9 Inhibition of <i>E. coli</i> β -N-acetylglucosaminidase activity.....	31
2. 2. 10 NMR experiments to detect <i>E. coli</i> β -N-acetylglucosaminidase hydrolysis of an O-GlcNAc-modified peptide.....	31
2. 3 Results and discussion	33
2. 3. 1 Co-expression of OGT with substrate proteins using <i>E. coli</i>	33
2. 3. 2 Inhibiting NagZ activity in <i>E. coli</i> increases <i>in vivo</i> O-GlcNAc modification	37
2. 3. 3 <i>E. coli</i> NagZ cleaves O-GlcNAc from a peptide	41
2. 3. 4 Identifying the <i>E. coli</i> β -N-acetylglucosaminidase as NagZ	43
2. 4 Conclusions	44
2. 5 Acknowledgments.....	46
 Chapter 3: Is NagZ the saboteur?.....	 47
Chapter 3. 1 Introduction	47
3. 1. 1 The Hexosamine biosynthesis pathway in <i>E. coli</i>	47
3. 1. 2 The role of β -N-acetylglucosaminidase (NagZ) in <i>E. coli</i>	48
3. 2 Materials and Methods	50
3. 2. 1 Materials	50
3. 2. 2 Knocking out the <i>nagZ</i> gene in the Rosetta(DE3) pLysS strain.....	51
3. 2. 3 Making the Δ <i>nagZ</i> Rosetta(DE3) pLysS cells competent	51

3. 2. 4	Transformation.....	52
3. 2. 5	Expression conditions.....	52
3. 2. 6	Ni-affinity purification conditions	54
3. 2. 7	SDS-PAGE and western blot analysis	54
3. 2. 8	WGA agarose affinity chromatography	55
3. 2. 9	Coomassie protein concentration assay	56
3. 2. 10	Supplementing $\Delta nagZ$ <i>E. coli</i> cultures with sugar	56
3. 2. 11	Resolubilized O-GlcNAc-modified ABL2.....	57
3. 2. 12	OGT activity assay.....	57
3. 3	Results.....	57
3. 3. 1	Expression of O-GlcNAc-modified ABL2 in $\Delta nagZ$ <i>E. coli</i>	57
3. 3. 2	$\Delta nagZ$ Rosetta(DE3)pLysS sugar trials	59
3. 3. 3	Resolubilized ABL2.....	61
3. 3. 4	Optimized expression of O-GlcNAc-modified ABL2 in $\Delta nagZ$	63
3. 3. 5	Comparison of OGT activity produced from human OGT (hOGT) and codon-optimized OGT (coOGT)	65
3. 4	Conclusion.....	66

Chapter 4: *Chlamydia trachomatis* CT663: Structure and Ligand Interaction

Model.....	69
4. 1 Introduction.....	69
4. 1. 1 <i>Chlamydia trachomatis</i>	69
4. 1. 2 Type III secretion	70
4. 1. 3 CT663/Scc1-CopN complex	72
4. 1. 4 Research goal.....	73
4. 2 Materials and Methods	74
4. 2. 1 Materials	74
4. 2. 2 Structure predictions.....	75
4. 2. 3 CT663 expression and purification	75
4. 2. 4 CT663 and Scc1 co-expression and co-purification:.....	75
4. 2. 5 CopN expression and purification	76
4. 2. 6 Separation of CT663 and Scc1	76
4. 2. 7 Gel electrophoresis analysis	77

4. 2. 8	Gel filtration analysis.....	77
4. 2. 9	SPR Analysis.....	78
4. 2. 10	NMR buffer optimization for CT663.....	79
4. 2. 11	¹⁵ N-CT663/Scs1 NMR analysis.....	80
4. 3	Results.....	81
4. 3. 1	Structural predictions of the <i>Chlamydia trachomatis</i> T3SS effectors-chaperone complex CT663:SCC1:CopN.....	81
4. 3. 2	Expression and Purification.....	82
4. 3. 3	Separation of CT663/Scs1 complex.....	83
4. 3. 4	Buffer dependent gel filtration analysis.....	85
4. 3. 5	Detergent assisted Gel filtration.....	86
4. 3. 6	SPR analysis of Ct663/Scs1 binding to CopN.....	88
4. 3. 7	The magical CT663 sample from never never land.....	89
4. 3. 8	NMR buffer optimization for CT663.....	90
4. 3. 8	¹⁵ N-CT663:Scs1 NMR analysis.....	91
4. 3. 9	NMR analysis of the CT663/Scs1 complex.....	92
4. 4	Conclusions.....	93
	References.....	96
	Appendix.....	106
	Letter of Permission.....	106
	Supplementary Figures.....	107
	Vita.....	112

List of Tables

Table 3-1 The recipe for 1 L of minimal growth media, 1 L of 10X M9 salts, and 100 mL of 1000X trace mineral.....	53
Table 4-1 This table summarized the resulting oligomeric state of CT663 and Scc1/CT663 complex in different buffer conditions.	87

List of Figures

Figure 1-1 Synthesis and structure of the peptide backbone. http://www.molecularsciences.org/book/export/html/253	1
Figure 1-2 Transcription and translation of proteins. http://www.tokresource.org/tok_classes/biobiobio/biomenu/transcription_translation/	2
Figure 1-3 Protein structure. https://www.mun.ca/biology/scarr/iGen3_06-04.html	4
Figure 1-4 The HBP pathway in humans. Yuzwa, S. A. and D. J. Vocadlo, O-glcnac and neurodegeneration: Biochemical mechanisms and potential roles in alzheimer's disease and beyond. <i>Chemical Society Reviews</i> , 2014. 43 (19): p. 6839-6858.	7
Figure 1- 5 Enzymatic cycling of O-GlcNAc by OGT and OGA. Hart, G. W. , M. P. Housley, and C. Slawson, Cycling of o-linked beta-n-acetylglucosamine on nucleocytoplasmic proteins. <i>Nature</i> , 2007. 446 (7139): p. 1017-22.	8
Figure 1-6 Methods of O-GlcNAc detection. Banerjee, P. S. , G. W. Hart, and J. W. Cho, Chemical approaches to study o-glcnacetylation. <i>Chemical Society Reviews</i> , 2013. 42 (10): p. 4345-4357	13
Figure 1-7 Steps involved in molecular cloning. http://cnx.org/contents/185cbf87-c72e-48f5-b51e-f14f21b5eabd@9.45:88/Biology	17
Figure 1-8 CBCANH experiment. http://www.protein-nmr.org.uk/pictures/experiment_types/cbcanh.png	20
Figure 2-1The O-GlcNAc modification of CREB1 disrupts binding with TAF _{II} 130.	22
Figure 2-2 The Ni-affinity purified soluble lysate and pellet from co-expression of CREB1-Q2 (pET-28b-CREB1-Q2) and OGT (pET-43. 1-OGT) in E. coli BL21-Gold(DE3) analyzed with (a) Coomassie-stained SDS-PAGE and (b) FITC-WGA lectin blot detecting O-GlcNAc. The bands corresponding to CREB1-Q2 are indicated with asterisks.	34
Figure 2-3 The soluble lysate and pellet from co-expression of CREB1-Q2-bZIP (pET28b-CREB1-Q2-bZIP) and OGT (pET-43. 1-OGT) in E. coli BL21-Gold(DE3) analyzed with (a) Coomassie-stained SDS-PAGE and (b) FITC-WGA lectin blot detecting O-GlcNAc. The bands corresponding to CREB1-Q2-bZIP are indicated with asterisks.	35
Figure 2-4 ABL2-FABD (pET-28b-ABL2-FABD) and OGT (pET-43. 1-OGT) were co-expressed in E. coli BL21-Gold(DE3) with increasing concentrations of PUGNAc in the growth media (0, 1, 2, and 10 μ M PUGNAc). The soluble lysate (S) and pellet (P) from each culture was analyzed with (a) Coomassie-stained	

SDS-PAGE and (b) anti-O-GlcNAc MAb CTD 110. 6 western blot. HEWL was used as a positive control in the last lane. (c) A plot of the relative band density of the western blot band to the Coomassie-stained SDS-PAGE gel band (indicated with an asterisk). For each band in (a) and (b), the background density was measured thrice, subtracted from each band, and used to estimate the propagated measurement error (represented as error bars in (c)).36

Figure 2-5 The initial velocities of *E. coli* β -N-acetylglucosaminidase, partially purified with CXC from BL21-Gold(DE3) cleared lysate, at varying MU-GlcNAc concentrations and with different concentrations of PUGNAc. The points represent the observed velocities, and the lines represent the best fits of the data to Eqs. (1) and (2).39

Figure 2-6 The β -N-acetylglucosaminidase activity of cleared lysate with PUGNAc (0, 1, 2 μ M) was measured using the pNP-GlcNAc spectrophotometric assay. The lysate was from the co-expression of ABL2-FABD (pET-28b-ABL2-FABD) and OGT (pET43. 1-OGT) in *E. coli* BL21-Gold(DE3). Each sample and control were run in triplicate, and the error bars represent the standard deviation of these measurements. The 405 nm absorbance values were converted to activity (right vertical axis) using the extinction coefficient of nitrophenolate.41

Figure 2-7 Fractions from CXC of BL21-Gold(DE3) lysate with β -N-acetylglucosaminidase activity toward MU-GlcNAc were pooled and tested for β -N-acetylglucosaminidase activity toward the synthetic O-GlcNAc-modified peptide, SVES(β -O-GlcNAc)GSADAK-NH₂. (a) CXC fraction control, (b) CXC fraction control with 200 μ M GlcNAc added, (c) CXC fraction with 200 μ M glycopeptide, and (d) 200 μ M glycopeptide control in CXC buffer. The spectra shown are expanded regions from 5.1 to 5.3 ppm, with a gray dashed line indicating the α -GlcNAc chemical shift (center of the doublet). All spectra were processed and expanded identically for direct comparison. The full spectra are included in Supplementary Figure 4. The glycopeptide control sample (d) has extra noise due to truncation artifacts from the water peak; the additional noise is small and does not affect spectral interpretation. Based on the ¹H anomeric α -GlcNAc intensity in (b), approximately 40 μ M GlcNAc was hydrolyzed by the *E. coli* β -N-acetylglucosaminidase in 5 h at 37 °C (c).42

Figure 2-8 ¹H NMR spectra of (a) Δ nagZ K12 lysate, (b) lysate with 200 μ M GlcNAc, and (c) lysate with 200 μ M glycopeptide, SVES(β -O-GlcNAc)GSADAK-NH₂. All samples were incubated at 37 °C for 22.5 h to test for β -N-acetylglucosaminidase activity toward the synthetic O-GlcNAc-modified peptide. The spectra shown are expanded regions from 5.1 to 5.3 ppm, with a gray dashed line indicating the α -GlcNAc chemical shift (center of the doublet). All spectra were processed and expanded identically for direct comparison. The full spectra are included in Supplementary Figure 6. Based on the signal-to-noise

ratio of the 200 μ M GlcNAc peak in (b), the limit of quantitation (LOQ) in (c) is 10 μ M GlcNAc.	44
Figure 3-1 Glucose, glucosamine, and N-acetylglucosamine are used by <i>E. coli</i> 's hexosamine biosynthesis pathway to make UDP-GlcNAc.	48
Figure 3-2 NagZ cleaves the bond between GlcNAc and MurNAc, of the peptidoglycan layer [70]. Adapted from: Triassi, A. J. , M. S. Wheatley, M. Savka, H. M. Gan, R. C. J. Dobson, and A. O. B. Hudson, L,L-diaminopimelate aminotransferase (dapI): A putative target for the development of narrow-spectrum antibacterial compounds. <i>Frontiers in Microbiology</i> , 2014. 5	49
Figure 3-3 O-GlcNAc-modified ABL2 expressed in the Δ nagZ system was unstable and O-GlcNAc-ABL2 selectively precipitated. The soluble and insoluble samples both contain ABL2 as seen on the protein gel (a); however, the western blot (b) shows that O-GlcNAc-modified ABL2 seems to be exclusively in the pellet.	58
Figure 3-4 Analysis of sugar supplemented expression of O-GlcNAc modified ABL2. The cleared lysate from cultures supplemented with glucose have a small amount of ABL2 (a) that is not glycosylated (b). The pellets corresponding to these samples contain about the same amount of ABL2 (c); however, the samples supplemented with glucose and glucosamine have an abundance of O-GlcNAc-modified ABL2 (d). The pellets containing PUGNAc express more ABL2 (c), but less O-GlcNAc-modified ABL2 compared to the glucose and glucosamine samples (d).	61
Figure 3-5 Purification of resolubilized O-GlcNAc-modified ABL2. The pellets were resolubilized with 10% sarkosyl, diluted to 1%, and purified with nickel affinity chromatography. The results were very similar to previous data—there is not an abundance of ABL2 on the protein gel, (a) yet it is highly glycosylated (b).	63
Figure 3-6 Optimized expression of soluble O-GlcNAc-modified protein. A starter culture from a single colony was used to inoculate sugar-supplemented cultures in MM. The expression conditions were optimized to produce soluble O-GlcNAc-modified protein. The soluble lysate was purified and concentrated for SDS-PAGE and western blot analysis. The glucose sample was the only sample to produce a substantial amount of soluble O-GlcNAc-modified protein.	64
Figure 3-7 Comparison of human OGT and codon optimized OGT. Anti-O-GlcNAc transferase was used to detect the amount of OGT present in the in vitro reaction. Anti-O-GlcNAc was used to detect the amount of O-GlcNAc-modified protein produced from the reaction.	66
Figure 4-1 Developmental cycle of Chlamydia. Brunham, R. C. ; Rey-Ladino, J. , "Immunology of Chlamydia infection: Implications for a Chlamydia trachomatis vaccine," <i>Nature Reviews Immunology</i> 2005 , 5 (2), 149-161.	70

Figure 4-2 Diagram of animal pathogenic bacterial T3SS similar to that of Chlamydia [86]. Büttner D , and He S Y Type III Protein secretion in plant pathogenic bacteria. *Plant Physiology*. **2009**;150:1656-1664.71

Figure 4-3 The CopN:CT663:Scs1 structure is modeled using Yersinia pestis virulence factor YopN in complex with the heterodimeric chaperone SycN-YscB by Phyre2.82

Figure 4-4 SDS-PAGE and native PAGE analysis of CT663, Scs1/CT663, and CopN. The protein molecular weight ladder is shown in lane 1 in the (a) SDS-PAGE and (b) native PAGE gel images. The molecular weight of CT663, Scs1/Ct663 and CopN are 14. 7 kDa, 30. 9 kDa, and 45. 2 kDa respectively. ..83

Figure 4-5 CT663-Scs1 interaction. Purified CT663 and CT663:Scs1 complex are shown in lanes 2 and 3, respectively. The purified CT663-Scs1 complex was immobilized on 1 mL of Ni-affinity media via a 6XHis tag on Scs1 and washed with 20 mL of wash buffer with 6 M urea (lane 4), 8 M urea (lane 5), and 2% SDS (lane 6). Even under strong denaturing conditions of 8 M urea, some CT663 remained bound to 6XHis-tagged Scs1. SDS (2%) was sufficient to disrupt the CT663:Scs1 interaction, resulting in pure Scs1.84

Figure 4-6 Gel filtration analysis of CT663 and CT663:Scs1 at various buffer conditions. Two gel filtration column were used to analyze the samples: (A) a HiPrep Sephacryl S-100 HR column using a 1 mL sample volume and (B) an Enrich SEC 70 column using 50 Molecular weight standards were used to determine the linear relationship between the elution volume and the logarithm of the molecular weight for each column/buffer combination. The molecular weight of the samples were determined using the best-fit linear function based on regression analysis to determine the oligomeric state of the samples in each buffer. The sample concentrations were 81-152 μ M except for the CT663:Scs1 complex in sodium acetate, which was 27 μ M due to its instability in this buffer. Details about the sample conditions are described in Table 4-1. (A) In Tris/saline buffer at pH 8. 0, CT663 (solid black line) predominantly exists as higher order oligomers (>300 kDa) and as a homodimer while the CT663:Scs1 complex (dashed black line) predominantly forms a stable tetramer. In acetate buffer at pH 4. 5, CT663 and the CT663:Scs1 complex are unstable. CT663 readily precipitates, but the complex stability was better and showed that soluble CT663:Scs1 in acetate buffer (gray solid line) exists as a heterodimer. (B) In phosphate buffer at pH 7. 3 (solid black line), CT663 forms a homodimer and in Tris buffer at pH 7. 4 (dashed black line), CT663 forms homotrimers.85

Figure 4-7 Detergent effects on CT663 oligomerization. (A) CT663 samples were run in 20 mM Tris, pH 7. 4 buffer (solid black line) with 1 mM SDS (dotted black line), 10 mM SDS (dashed black line), and 0. 01% Triton X-100 (dash-dotted black line). The predicted elution volumes for CT663 oligomers (+) were

calculated using the calibration curve (Table 4. 1). The number of monomer units for each oligomer is shown above the marker. (B) CT663 samples were run in 50 mM potassium-sodium phosphate, pH 7. 3 buffer (solid gray line) with 0. 01% Triton X-100 (dotted gray line) and 1% PEG (dashed gray line). The predicted elution volumes for CT663 oligomers (+) are shown as in (A). Each sample was 50 μ l. Na phosphate, pH 7. 3 buffer, except for the sample run in 10 mM SDS/Tris buffer. This sample was diluted by a factor of 4 with the running buffer to avoid SDS-phosphate precipitation. The absorbance values for this sample were multiplied by 4 for direct comparison to the other samples.87

Figure 4-8 The SPR sensogram corresponding to the binding event between the Scc1/CT663 complex and CopN. The calculated $K_d=3. 02 \pm 0. 17$ nM.89

Figure 4-9 $^1\text{H}-^{15}\text{N}$ HSQC of ^{15}N labeled CT663. Each peak corresponds to an $^1\text{H}-^{15}\text{N}$ bond from the backbone or a side chain of the protein.90

Figure 4-10 $^1\text{H}-^{15}\text{N}$ HSQC of ^{15}N labeled CT663 in 20 mM sodium phosphate, 20 mM tris, and 20 mM MES.91

Figure 4-11 Scc1 perturbs the amide peaks of ^{15}N -CT663. An overlay of $^1\text{H}-^{15}\text{N}$ HSQC spectra of ^{15}N -CT663 (red) and ^{15}N -CT663:Scc1 (blue). The inset illustrates the selective shifts of the putative glycine amide peaks upon binding Scc1. Note: No peaks are covered by the inset.92

Figure 4-12 $^1\text{H}-^{15}\text{N}$ HSQC of ^{15}N labeled Scc1/CT663. The inset is a zoomed view to show the peak dispersion.93

List of Abbreviations

O-GlcNAc	O-linked beta-N-acetylglucosamine
OGT	O-GlcNAc transferase
OGA	O-GlcNAcase
Glu	glucose
GlcN	glucosamine
GlcNAc	N-acetylglucosamine
UDP-GlcNAc	Uridine diphosphate N-acetylglucosamine
WGA	Wheat germ agglutinin
HBP	Hexosamine biosynthesis pathway
NagZ	β -N-acetylglucosaminidase

Abstract

O-linked β -*N*-acetylglucosamine is a regulatory post translational modification. This modification occurs on nearly all functional classes of proteins, in the nucleus and cytoplasm. O-GlcNAc is added to serine or threonine by O-GlcNAc transferase and removed by O-GlcNAcase. Previous attempts to study O-GlcNAc-modified proteins have resulted in low yields, making 3-dimensional structure determination impossible. In this dissertation O-GlcNAc transferase will be co-expressed with domains of human cAMP responsive element-binding protein (CREB1) and Abelson tyrosine-kinase 2 (ABL2) in *E. coli*, to produce O-GlcNAc-modified protein. The O-GlcNAc-modified protein was expressed in a variety of *E. coli* cell lines at a variety of conditions, but only small quantities of insoluble protein were produced. A glycosidase was suspected due to the disappearance of the O-GlcNAc modification from the protein. O-(2-acetamido-2-dexoy-dglucopyranosylidene) amino-*N*-phenylcarbamate (PUGNAc), a β -*N*-acetylglucosaminidase inhibitor, was added to the culture media and increased the production of O-GlcNAc-modified protein. This was the first evidence that β -*N*-acetylglucosaminidase (NagZ), an *E. coli* enzyme, cleaves O-GlcNAc from proteins *in vivo*. NagZ was isolated and shown to cleave O-GlcNAc from a synthetic O-GlcNAc-peptide *in vitro*. In *E. coli*, NagZ cleaves the GlcNAc- β 1,4-*N*-acetylmuramic acid linkage to recycle peptidoglycan in the cytoplasm. A NagZ knockout showed no activity towards the O-GlcNAc-peptide, confirming NagZ as the enzyme responsible for cleaving O-GlcNAc from our glycoprotein expressed *in vivo*. O-GlcNAc-modified protein produced by the NagZ knockout (Δ NagZ) co-expression system is highly glycosylated and can be resolubilized

from the pellet. The Δ NagZ is a step closer to production of milligram quantities of O-GlcNAc-modified protein for structure determination.

Chapter 1: Background

1. 1 Proteins

Proteins are macromolecules that are present in all living systems. They which play essential roles in many cellular processes such as transport, storage, cellular signaling and immune response [1]. A major class of proteins, known as enzymes, act as catalyst for many chemical reactions in biological systems. They are composed of 20 amino acid building blocks. These amino acids, containing a carboxyl end, amino end and side chain, are linked together when a dehydration reaction occurs between two amino acids and forms a peptide bond (Figure 1-1).

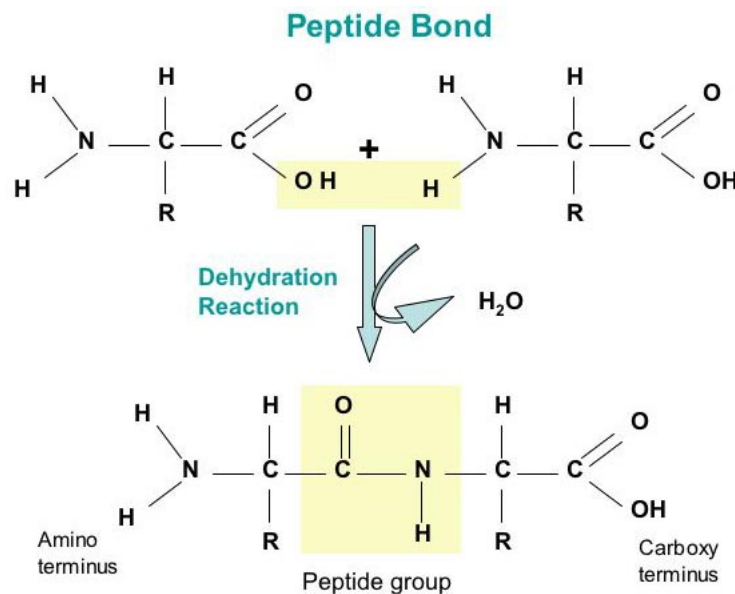


Figure 1-1 Synthesis and structure of the peptide backbone. <http://www.molecularsciences.org/book/export/html/253>

Proteins are synthesized through two processes called transcription and translation. During transcription, RNA polymerase transcribes single strands of DNA into complimentary strands of RNA called messenger RNA (mRNA). The mRNA is then translated into a polypeptide chain by ribosomes, composed of ribosomal proteins,

ribosomal RNA (rRNA), and transfer RNA (tRNA). tRNA is unlike mRNA in that it carries amino acids. One side of tRNA has nucleotide base pairs and the other has the amino acids that correspond to the nucleotide base pairs. The ribosome synthesizes a protein by reading mRNA, binding the appropriate tRNA, and creating peptide bonds between each amino acid (Figure 1-2) [2].

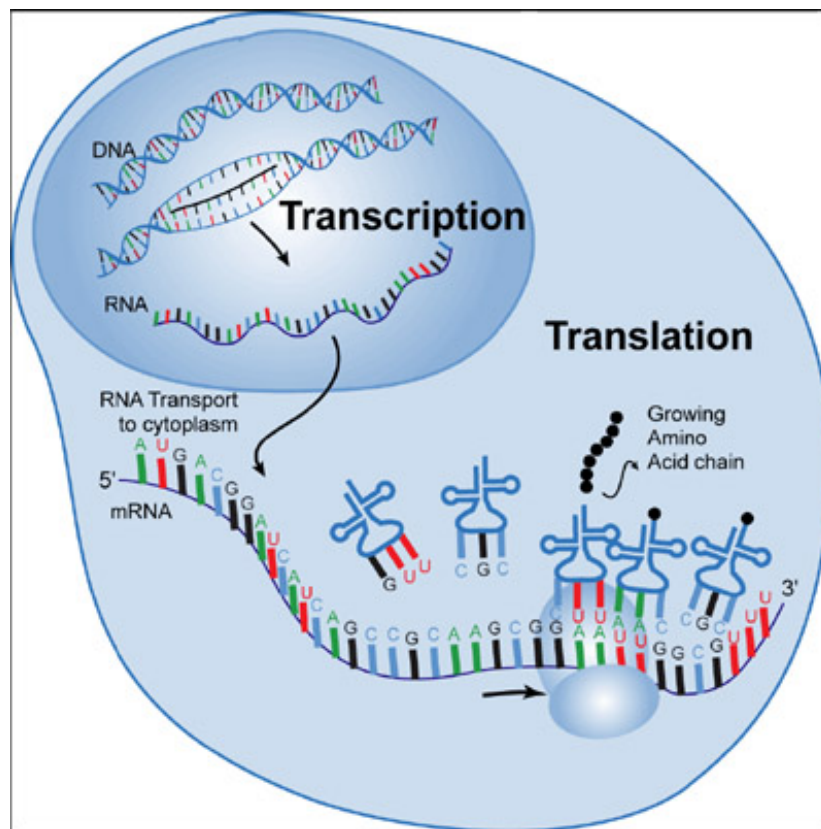


Figure 1-2 Transcription and translation of proteins. http://www.tokresource.org/tok_classes/biobiobio/biomenu/transcription_translation/

1. 2 Protein Structure

The overall 3-dimensional (3D) structure or tertiary structure of a protein is dependent upon its primary and secondary structure. The primary structure of a protein is its amino acid sequence. The secondary structure describes how this chain of amino acids behave in relation to their bond neighbor. Secondary structure consists of alpha

helices, beta sheets and non-structured regions called loops. Alpha helices form a spiral-like shape with 3.6 amino acids for every turn. The protein backbone forms the inside of the coil with the amino acid side chains at the exterior. The spiral or coil shape is stabilized by intramolecular hydrogen bonds formed between the amide group hydrogen and the oxygen of a carboxyl group 4 residues away (Figure 1-3). Beta sheets are flat or linear in shape and are also held together by hydrogen bonds. These bonds differ from those in the alpha helix in that they occur between separate chains. Also, the intramolecular hydrogen bond is between the hydrogen of an amide on one chain and the oxygen of a carboxyl of a neighboring chain. The alpha helices, beta sheet and loops fold to make subunits of the proteins. These subunits describe a protein's 3D tertiary structure. Finally, these tertiary structures of protein monomers fit together to make protein complexes or multimers. The structures resulting from these multimers is quaternary structure, which is the overall folded state of the entire complex (Figure 1-3). The tertiary and quaternary structures are held together by hydrogen bonds, ionic bonds, hydrophobic interactions and disulfide bonds between side chains [3].

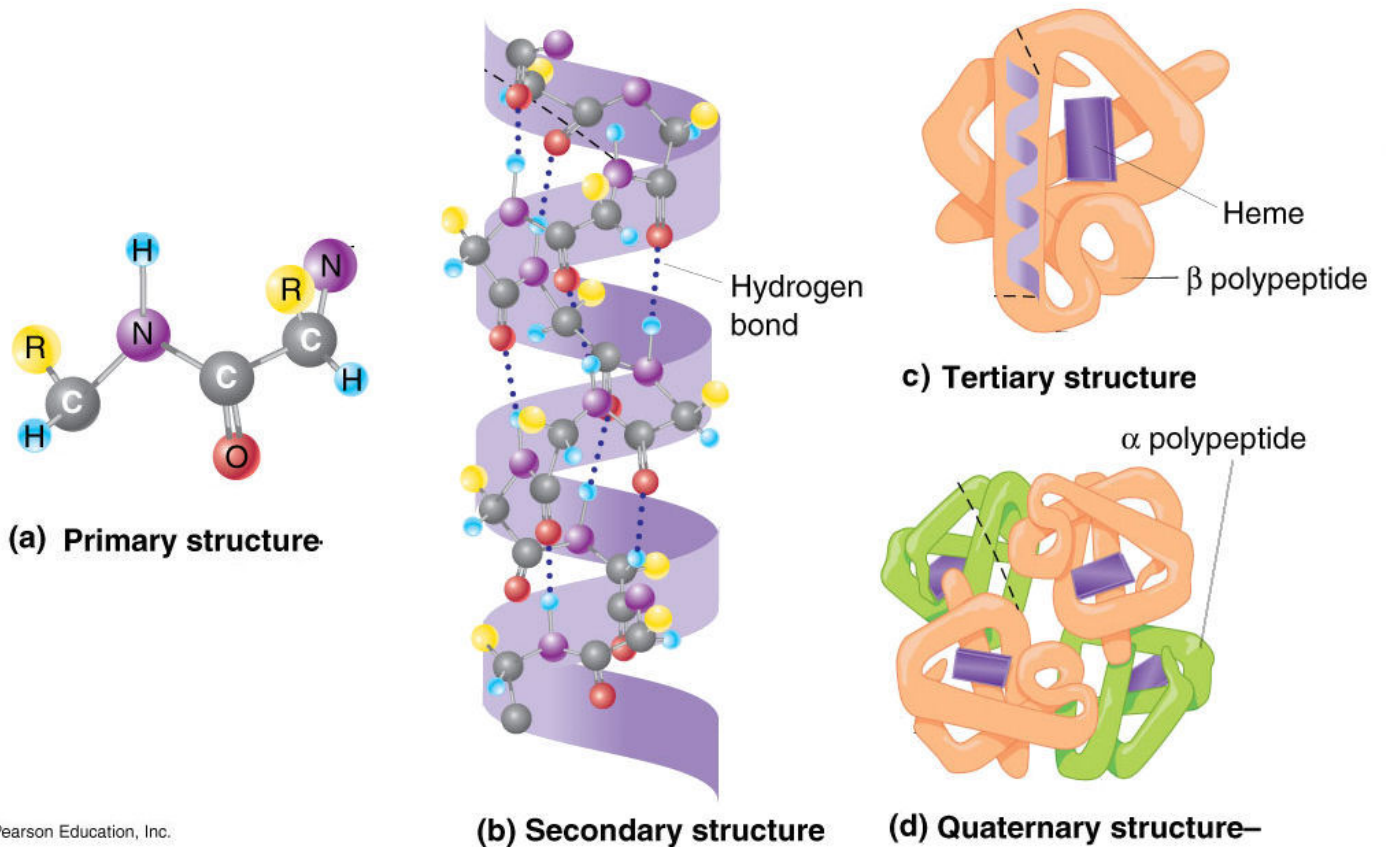


Figure 1-3 Protein structure. https://www.mun.ca/biology/scarr/iGen3_06-04.html

1. 3 Protein function as it relates to structure

Protein function depends on protein structure. In most cases, proteins have to be folded in a certain conformation to carry out their function. For example, in lock and key ligand binding, enzymes show specificity for substrates by forming an active site based on the enzyme's structure. This substrate binding event is often part of a cascade of events required for function to take place [3].

1. 4 Post translational modifications

A post translation modification (PTM) is a covalent chemical modification that occurs after the protein has been translated. There are many types of PTMs, such as methylation, ubiquitination, phosphorylation, and glycosylation. Methylation adds a

methyl group to a lysine or arginine. Ubiquitination is the addition of an entire ubiquitin protein molecule to a lysine residue. Phosphorylation and glycosylation are the two most common PTMs. Phosphorylation is the addition of a phosphate at a serine and/or threonine residue. Glycosylation can be N-linked and O-linked. N-linked glycosylation normally consists of the addition of a polysaccharide chain at an asparagine, as long as the asparagine is part of a tri-peptide motif consisting of the asparagine, any amino acid and then a serine or threonine (N, X, S/T). This consensus sequence is a tri-peptide within the sequence of the protein where the N-linked glycosylation may occur. Knowledge of this consensus sequence allows us to predict possible glycosylation sites. Most O-linked glycosylations are similar to N-linked glycosylations, in that they are long chain polysaccharides. The exception is O-linked *N*-acetylglucosamine (O-GlcNAc), which is the addition of a monosaccharide instead of a polysaccharide [4].

1. 5 O-GlcNAc in the hexosamine biosynthesis pathway

The hexosamine biosynthesis pathway (HBP) is a metabolic pathway that produces uridine diphosphate *N*-acetylglucosamine (UDP-GlcNAc), the substrate for O-GlcNAcylation, and other nucleotide hexosamines (Figure 1-4). Hexosamines, also called amino sugars, are sugars that have an amino group (-NH₂) in place of a hydroxyl group. The HBP varies based on the organism. In humans, glucose enters the cell and is phosphorylated by hexokinase. Only 2-5% of glucose that enters the cell enters the HBP, therefore the level of O-GlcNAcylation is dependent on nutrient availability [5]. Once glucose is phosphorylated and converted to fructose-6-phosphate, it can enter the HBP or the main glycolytic/glycogen pathway. The HBP is initiated by glutamine/fructose aminotransferase (GFAT). GFAT is an enzyme that transfers an

amino group to fructose-6-phosphate from the amino acid glutamine, producing glucosamine-6-phosphate [6]. It is important to add that glucosamine-6-phosphate can also be produced by glucosamine that enters the cell and is phosphorylated by hexokinase [7]. EMeg32, a glucosamine-6-phosphate *N*-acetyltransferase, is an enzyme that acetylates glucosamine-6-phosphate using acetyl-CoA as a substrate [8]. *N*-acetylglucosamine-6-phosphate, produced by the acetylation reaction, is interconverted to *N*-acetylglucosamine-1-phosphate by *N*-acetyl phospho-glucosamine mutase (AGM1). *N*-acetylglucosamine-1-phosphate is finally uridylated to UDP-GlcNAc by UDP-GlcNAc pyrophosphorylase (UAP1) [9].

1. 6 O-GlcNAc is not your typical glycosylation

O-GlcNAc is an O-linked glycosylation that occurs at a serine or a threonine. O-GlcNAc is cycled on and off of protein substrates by enzymes. It is added by the enzyme O-GlcNAc transferase (OGT) and removed by O-GlcNAcase (OGA) (Figure 1-5) [10]. This O-linked modification is unlike most glycosylation. Only one sugar is added, as opposed to a longer polysaccharide chain. There is no consensus sequence recognized by OGT, which makes O-GlcNAc-site prediction difficult [11]. O-GlcNAc has a reciprocal relationship with phosphorylation. They can occur on the same protein at the same time or at a specific residue interchangeably [12]. This relationship with phosphorylation is used by most prediction algorithms to identify possible O-GlcNAc modification sites [12]. In mammals, most protein glycosylation occurs in the extracellular matrix or the secretion pathway. However, O-GlcNAc can be found in the cytoplasm, nucleus and mitochondria.

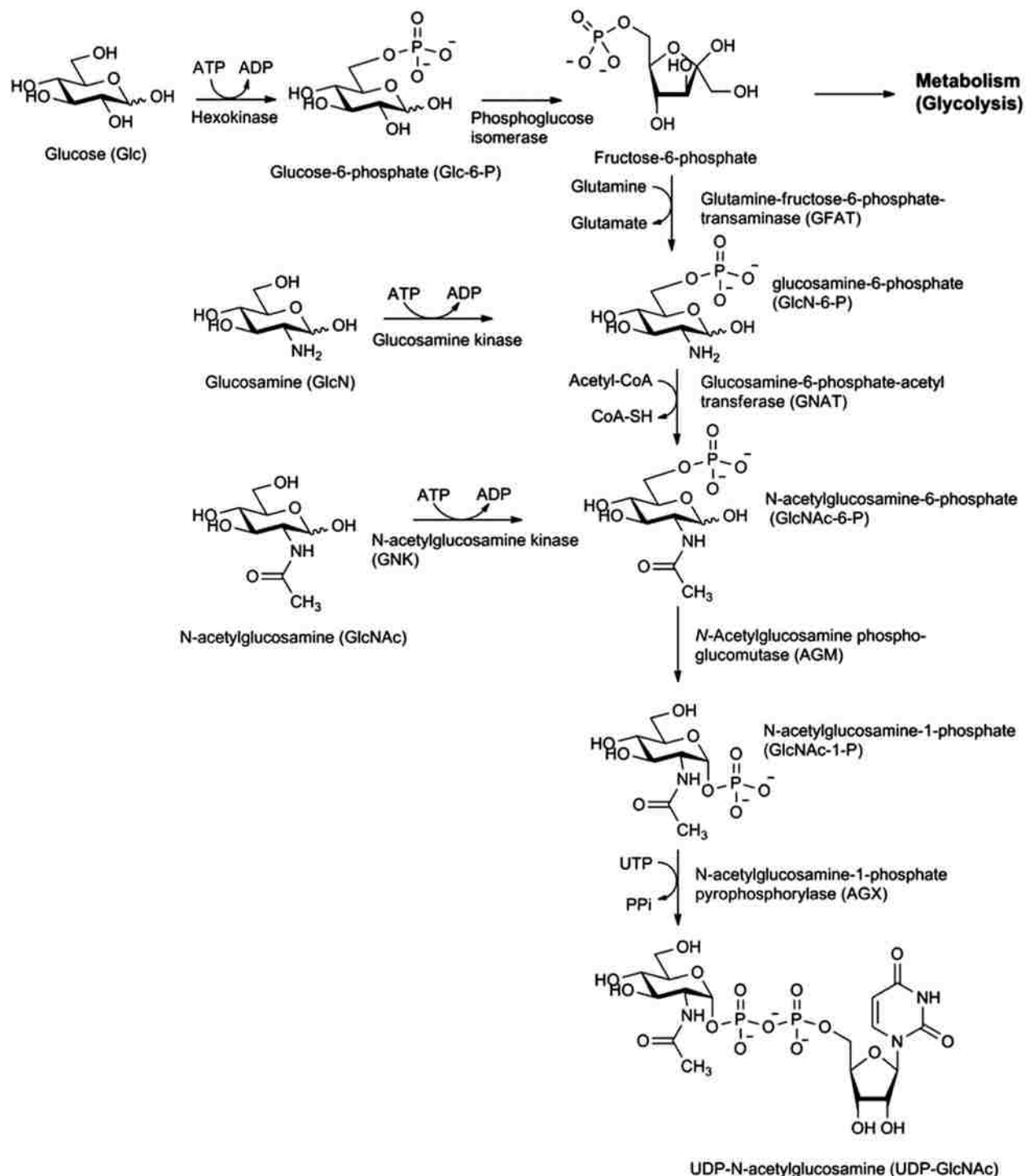


Figure 1-4 The HBP pathway in humans. Yuzwa, S. A. and D. J. Vocadlo, O-glcnac and neurodegeneration: Biochemical mechanisms and potential roles in alzheimer's disease and beyond. Chemical Society Reviews, 2014. **43**(19): p. 6839-6858.

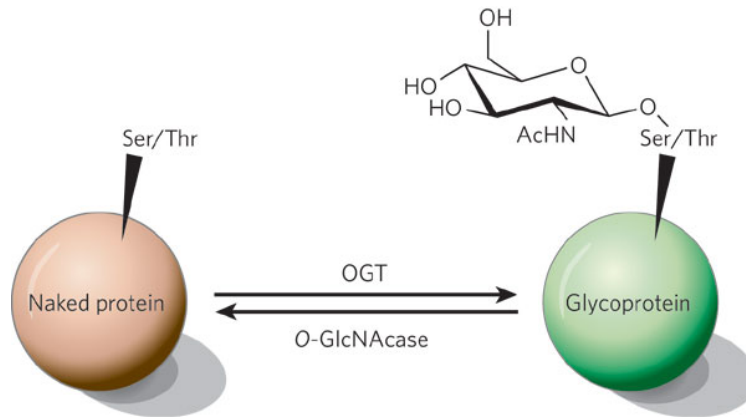


Figure 1- 5 Enzymatic cycling of O-GlcNAc by OGT and OGA. Hart, G. W. , M. P. Housley, and C. Slawson, Cycling of o-linked beta-n-acetylglucosamine on nucleocytoplasmic proteins. *Nature*, 2007. **446**(7139): p. 1017-22.

1. 7 O-GlcNAc is a regulatory enzyme that is implicated in disease when it is dysregulated

The O-GlcNAc modification was first discovered by Gerald Hart in 1984 and is dynamic and regulatory in nature [13]. O-GlcNAc-modified proteins regulate many biological processes, such as transcription, translation, protein degradation, cell signaling, cell trafficking, apoptosis and cell cycle control [14]. It also regulates some physiological events, such as the circadian rhythm, memory formation, and learning [15]. When the OGT gene is knocked out in mice, it is embryonic lethal, therefore this regulatory modification is essential for life [16].

Cyclic AMP-response element binding protein 1 (CREB1) is a transcription factor that binds to TATA-associated factor (Taf_{II}130), a coactivator, to induce transcription of the cyclic AMP response element (CRE) gene [17]. When O-GlcNAc is bound to CREB1, association with Taf_{II}130 is inhibited and transcription is repressed [17]. Repression of the CRE gene leads to reduced growth in axons and dendrites in mice brain cells [18]. Axon and dendrite growth and connection is how memories are formed.

Dysregulated O-GlcNAcylation of CREB1 leads to memory loss in mice and has been linked to Alzheimer's disease [20].

In the brain, neurons (brain cells) consist of a cell body, dendrites, and axons. The neuronal cell body is very similar to other cells and contains the nucleus, cytoplasm and cellular membrane. Dendrites carry information into the cell body through electric signals, while axons take information away from the cell body. Dendrites are connected to the cell body at one end while the axons are at the opposite end of the cell body [21]. Microtubules are the major structural component of axons. They play a key role in intracellular transport by carrying nutrients and other cellular cargo [21]. They are hollow and tubular in shape and are composed of polar filamentous polymers of α/β tubulin heterodimers. These microtubules are held together by MT binding proteins such as tau. Hyperphosphorylated tau separates tau from the microtubules, causing them to fall apart [22]. The hyperphosphorylated tau molecules aggregate within the neuron, causing tangles, inhibiting neurotransmission and destroying the cell [23]. The O-GlcNAc modification can compete with phosphorylation sites on tau. When OGA is inhibited in rodents, the O-GlcNAc modification increases and the phosphorylation decreases at pathologically relevant sites on Tau [23]. This increase in O-GlcNAcylation blocks hyperphosphorylation and hinders aggregation [24].

O-GlcNAc also plays a role in diabetes and cancer. Recent research has suggested that an increase of the O-GlcNAc modification in prostate cancer cells promotes the progression of tumor cells. This elevation of O-GlcNAc is associated with an increase in overall patient expiration [25]. Elevated O-GlcNAc levels have also been shown to enhance migration/invasion and metastasis in breast cancer cells [26]. Over

expression of OGT, the enzyme that adds O-GlcNAc; and GFAT, the rate limiting step of the HBP; in various tissues leads to insulin resistance in mice [27]. Insulin is a hormone that carries glucose from the bloodstream into cells. Insulin resistance refers to the decreased efficiency of insulin for glucose uptake, metabolism and storage. Insulin resistance is a major feature and cause of Type 2 diabetes [28].

1. 8 Inhibitors of hexosaminidase

O-GlcNAcase is a hexosaminidase that removes the O-GlcNAc modification from substrate proteins. Hexosaminidases are enzymes that remove hexosamines. There are several different hexosaminidases. When analyzing the O-GlcNAc modification in cells that contain a hexosaminidase, it is often necessary to inhibit the hexosaminidase in order to increase glycosylation and analyze the effect of hyperglycosylation. Streptozotocin (STZ) was one of the first OGA inhibitors used [29]. Recent research suggest that STZ acts as a chemical poison rather than an OGA inhibitor [5]. O-(2-Acetamido-2-deoxy-D-glucopyranosylideneamino) amino-*N*-phenylcarbamate (PUGNAc) is another inhibitor of OGA with a K_i value of 50 nM for OGT purified from rats [5]. Although PUGNAc inhibits OGA, it is not specific to OGA and inhibits other glycosidases [30]. PUGNAc inhibits lysosomal hexosaminidase with a K_i of 30-40 nM [5]. Thiamet-G has a K_i of 21 nM for OGA and is a stable bioactive inhibitor, leading to increased tau glycosylation in mice brain [31]. In an effort to create a strong and selective inhibitor, Dorfmuller and coworkers developed a class of small molecule OGA inhibitors called GlcNAcstatins [32]. This class of inhibitors is more selective towards human OGA and increases O-GlcNAcylation by 2 to 3 fold [5].

1. 9 Methods for O-GlcNAc detection

1. 9. 1 Galactosyltransferase-based methods

B-1,4-Galactosyltransferase is an enzyme that transfers a single galactose from UDP-galactose to a terminal GlcNAc molecule (Figure 1-6b). This enzymatic labeling method was used in 1984 by Gerald Hart when he first discovered the O-GlcNAc modification [13]. Radioactive ^3H -galactose was added to O-GlcNAcylated proteins and the radioactivity was detected. This technique can be used to transfer a galactose with a chemically reactive group to the O-GlcNAc residue [33]. This allows for biotin labeling, which can then be used for chemiluminescent and western blot detection, or enrichment for further analysis.

1. 9. 2 Lectin blot

A lectin blot is a technique in which a lectin is used to detect a glycoprotein (Figure 1-6a). The lectin, which is a protein that recognizes and binds a specific carbohydrate and binds carbohydrate based molecules, is conjugated to a fluorescent tag for detection. Lectins are found in most animals, plants and microorganisms. Wheat germ agglutinin (WGA) is the lectin that recognizes and binds to terminal GlcNAc or *N*-acetylneuraminic acid residues [34]. WGA is bound to agarose beads and used as a means of affinity purification in order to enrich the glycoprotein for further analysis. Although WGA binds O-GlcNAc, it is not always selective. For this reason, other means of detection are sometimes necessary.

1. 9. 3 Western blot

Western blots are very similar to lectin blots. Instead of using a lectin to bind and detect the protein of interest, primary and secondary antibodies are used (Figure 1-6a). Proteins are transferred from a sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) gel onto a nitrocellulose or PVDF membrane. The membrane is then blocked. The term blocked refers to the blocking of free space on the membrane to insure proper antibody binding. The membrane is probed with the primary antibody. This antibody binds to the epitope of interest. CTD 110. 8 and RL-2 are common antibodies for the detection of O-GlcNAc. CTD 110. 8 binds specifically to O-GlcNAc while RL2 less specific. A secondary antibody, which is conjugated to a fluorescent tag, binds the primary antibody. Although this method of detection is very similar to a lectin blot, it is more selective.

1. 9. 4 Chemical Derivatization – BEMAD

Mass spectrometry (MS) techniques that use high energy ionization methods tend to fragment the O-GlcNAc modification before it can be analyzed. In 2002 Lance Wells applied the beta elimination Michael addition (BEMAD) reaction, a two-step organic reaction, to O-GlcNAc-modified proteins [35]. The BEMAD reaction replaces the O-GlcNAc modification with another compound that is stable during MS analysis. The first step of the reaction is the beta elimination of O-GlcNAc, followed by the addition of a new compound via Michael addition of a nucleophile. Dithiothreitol (DTT) or biotin pentylamine (BAP) are the nucleophiles of that have been used [36]. At the time, this reaction was notable because it allowed detection of O-GlcNAc-modified protein and site mapping. Site mapping is a process where the exact amino acid modified is

identified. Now, there are lower energy fragmentation techniques that allow the O-GlcNAc modification to stay intact during site mapping [35].



Figure 1-6 Methods of O-GlcNAc detection. Banerjee, P. S. , G. W. Hart, and J. W. Cho, Chemical approaches to study o-glcNacylation. Chemical Society Reviews, 2013. **42**(10): p. 4345-4357

1. 10 Methods to study O-GlcNAc-modified proteins

There have been numerous strategies to study O-GlcNAc-modified proteins. O-GlcNAc-modified proteins have been extracted from animal tissues. OGT has been extracted and used for *in vitro* glycosylation. OGT has been expressed in *E. coli* and used for *in vitro* glycosylation. Target proteins have been expressed in human cells that have the OGT gene. The aim of methods is to identify O-GlcNAc-modified proteins. WGA affinity chromatography is often used to enrich protein samples. The sites of glycosylation are then mapped using MS. Mice and fly have been used to study to effects of the O-GlcNAc modification on cellular processes related to a specific protein [37, 38].

1. 11 Challenges

Although there have been major discoveries about the O-GlcNAc modification since it was first discovered, there are still many limitations. The greatest of these is the sample size. Proteins extracted from animal tissue are low in concentration. The same is true for modified proteins expressed in human cells. *In vitro* reactions are normally done on a small scale due to the low concentration of extracted OGT. Human OGT expressed in *E. coli* has also been used for *in vitro* glycosylation reactions. Limited sample concentration has made structural analysis of O-GlcNAc-modified proteins impossible. NMR and X-ray crystallography require milligram quantities of protein when the methods to date yield only micrograms of O-GlcNAc-modified protein.

1. 12 *In vivo* expression of O-GlcNAc-modified protein

Dysregulation of the O-GlcNAc modification, whether too much or too little, has been implicated in disease. Given that the modification is essential for life and has been

linked to several diseases, understanding the effect of the modification on protein structure is imperative. In this thesis, O-GlcNAc-modified proteins will be expressed *in vivo*. In this particular method, OGT and a target protein are co-expressed in *E. coli*. This method is especially interesting because it enables the expression of large amounts of modified target proteins. More importantly, this method allows isotopic labeling for NMR structural analysis. *E. coli* was chosen over mammalian or insect cells because it does not carry the OGT gene. *E. coli* is the most common expression system for milligram quantities of protein for NMR and X-ray crystallography structural analysis [39]. This method requires molecular cloning of the target proteins and OGT DNA into expression vectors, transformation of the vectors into expression cells, protein expression and purification. Once the protein has been successfully produced with and without the modification, their 3D structures will be determined and compared. *E. coli* requires sources of carbon and nitrogen to synthesize proteins. Growing *E. coli* with ^{15}N -ammonium chloride and ^{13}C -glucose produces proteins with isotopically labeled backbones and side chains. This method also allows the isotopic labeling of the O-GlcNAc modification.

1. 13 Molecular cloning and gene expression

Molecular cloning is a process where DNA encoding a target protein is inserted into a cloning vector. This process is usually achieved through polymerase chain reaction (PCR), restriction digestion, and ligation (Figure 1-7). PCR is an experiment where a small fragment of DNA is amplified, creating many copies of that fragment. PCR consist of 3 steps. First, the DNA is denatured separating it into single strands. Primers anneal, or bind, to their consensus sequences at the 3' end, and elongation

occurs. During elongation, RNA polymerase uses DNA base pair components to extend the primers, which makes a new double stranded DNA fragment. Every time this process occurs, the DNA fragments double. Once the fragment has been amplified, it is then digested with restriction enzymes. Restriction enzymes recognize specific DNA sequences and cleave them, leaving sticky or overlapping ends. This same digestion is done on the cloning vector. In the ligation step, the sticky ends are bound together by a ligase. The two pieces of DNA come together to form a plasmid, circular DNA, which is then transformed into competent *E. coli* cells. These cells make several copies of the plasmid that is used for protein expression.

1. 14 Recombinant protein expression and purification

Recombinant protein expression is a process in which DNA corresponding to a target protein is expressed from a vector. The expression vector is made of DNA that has all the necessary promoters and operators for protein expression in a bacterial host. The cloned and recombined DNA is transformed into *E. coli* and isopropyl β -D-1-thiogalactopyranoside (IPTG) is used to induce expression of the protein. The cells are harvested and lysed through mechanical or chemical force. Soluble proteins are separated from the cell debris by centrifugation. The protein of interest is isolated through a purification process such as affinity chromatography. Affinity chromatography is a process where a protein is tagged with the addition of some sequence of amino acids that has a moderate binding affinity for a compound immobilized on some type of resin. For example, nickel affinity chromatography uses the affinity between 6 histidine residues on the protein and Ni^{2+} bound to the resin to select for the protein of interest.

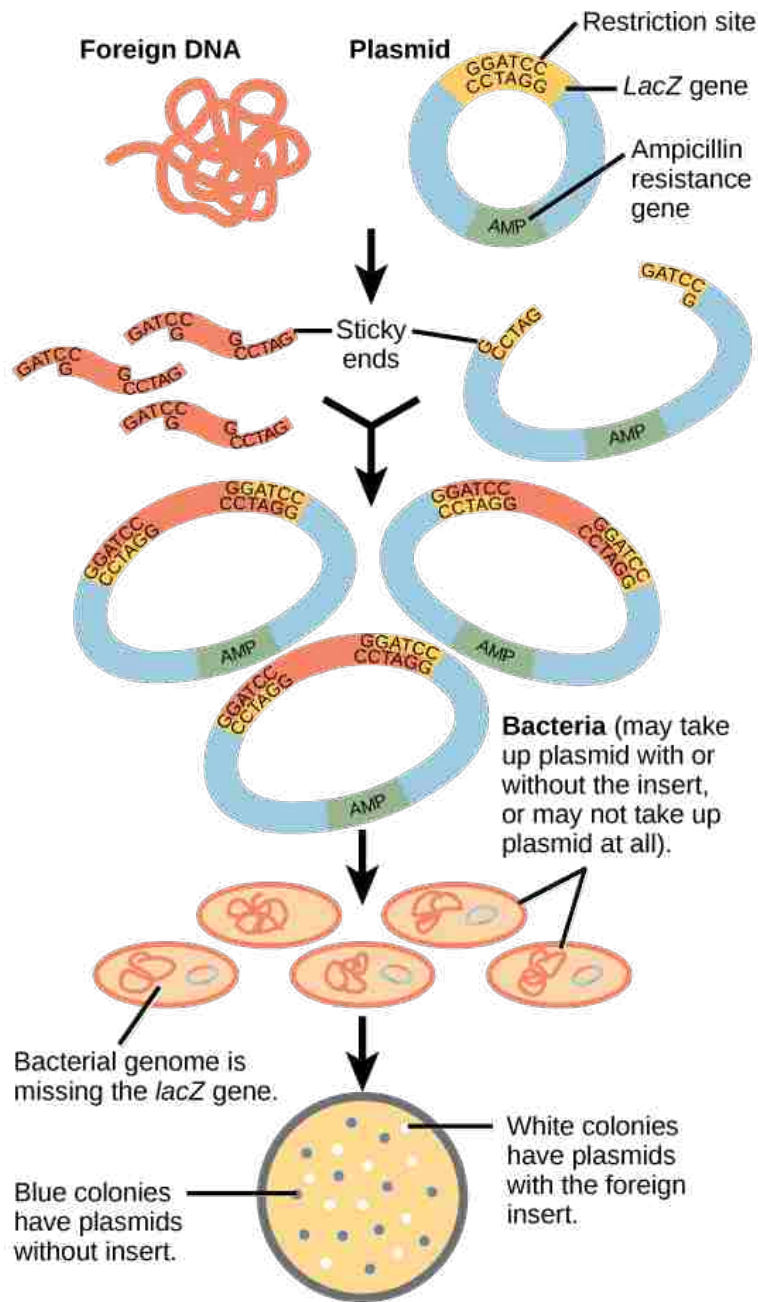


Figure 1-7 Steps involved in molecular cloning. <http://cnx.org/contents/185cbf87-c72e-48f5-b51e-f14f21b5eabd@9.45:88/Biology>

1. 15 Co-expression of target proteins with OGT

O-GlcNAc transferase (OGT) is the enzyme that adds O-GlcNAc to a target protein. Scientists began to question if a co-expression of OGT and a target protein would produce O-GlcNAc-modified protein. Although the mechanism of OGT is not fully

understood, we know that UDP-GlcNAc is the substrate for the addition of the O-GlcNAc modification. *E. coli* produces UDP-GlcNAc within the cells, making the expression of O-GlcNAc-modified proteins possible in a co-expression system. David Vocadlo has expressed a few O-GlcNAc-modified proteins using this method [24]. Expressing the modified protein in *E. coli* permits ^{15}N and ^{13}C isotopic labeling of the protein, as well as the modification, for NMR structural analysis.

1. 16 NMR

NMR is an analytical technique where the nuclear spins of a molecule are aligned by a constant magnetic field. An external oscillating magnetic field generated by a radiofrequency pulse is used to transfer a population of nuclei into a higher energy state with phase coherence. As the nuclei relax back to their lower energy state, the oscillating magnetic field generated by the coherent spins is detected. To date, there is no protein structure with the O-GlcNAc modification available by NMR or X-ray crystallography. One of the major setbacks of X-ray crystallography is that the technique depends on the protein's ability to be crystallized. It also only gives a static view of a protein's structure. NMR allows for the detection of conformational changes under physiological conditions. Although NMR is limited by the size of proteins that can be analyzed (about 30 kDa), it is not a problem for this research because all target proteins are within the applicable size range.

Methods for determination of protein structure by NMR have become standard, starting with backbone and side chain peak assignment, followed by distance and restraint measurements, and concluding with the final calculation of the structure. ^1H - ^{15}N heteronuclear single quantum coherence spectroscopy (HSQC) is the first NMR

experiment performed on the purified isotopically labeled protein. This experiment gives one peak for each amide group from each amino acid in the protein. 2D experiments show the correlation between two nuclei, in this case the ^1H attached to the ^{15}N of the amide group. Side chains containing amines also appear in ^1H - ^{15}N HSQC experiments as well. The dispersion, intensity and number of peaks gives valuable information about the protein sample. Spectra with narrow dispersion between 7-8 ppm indicate that the protein is unfolded or aggregated. Peak intensity is related to the concentration of the sample and the oligomerization state. Optimal sample conditions yield greater than 95% of expected backbone peaks. Once the conditions are optimized, additional experiments are done for further assignments. The next experiment collected is CBCANH. This 3D experiment shows which $\text{C}\alpha$'s and $\text{C}\beta$'s are connected to each amide group previously identified from the ^1H - ^{15}N HSQC experiment (Figure 1-8). In most cases, there are two $\text{C}\alpha$'s and two $\text{C}\beta$'s peaks for each amide, the $\text{C}\alpha/\text{C}\beta$ pair on the previous residue and the $\text{C}\alpha/\text{C}\beta$ pair directly connected to the amide. These $\text{C}\alpha/\text{C}\beta$ pairs are indicated by the residue number i and $i-1$. Because the $i-1$ $\text{C}\alpha/\text{C}\beta$ pair is not always visible, an additional experiment is done. CBCA(CO)NH only shows peaks corresponding to $\text{C}\alpha/\text{C}\beta$ $i-1$. These experiments allow scientists to identify the amino acid backbone peaks of the protein. Chemical shifts give information about the identity of the side chains and further experiments can be done if required.

The next set of experiments performed gives information about the secondary structure by measuring scalar coupling (J-coupling). J-coupling is a through bond measurement that gives information about bond angles. Nuclear overhauser effects (NOEs) are through space measurements and give information about protein

conformation. NOEs between neighboring residues also give secondary structural information. 2D Nuclear overhauser spectroscopy (NOESY) is used to measure distances between protons in close proximity. ^{15}N -NOESY experiments give NOE measurements for all hydrogens within 5 Å of an NH group. ^{13}C -NOESY measures all the hydrogens within 5 Å of a CH group. Once the NOESY experiments are completed, there are automated programs that assign the peaks. These assignments, along with the previous assignments and distance restraints, are used to calculate a protein structure [40].

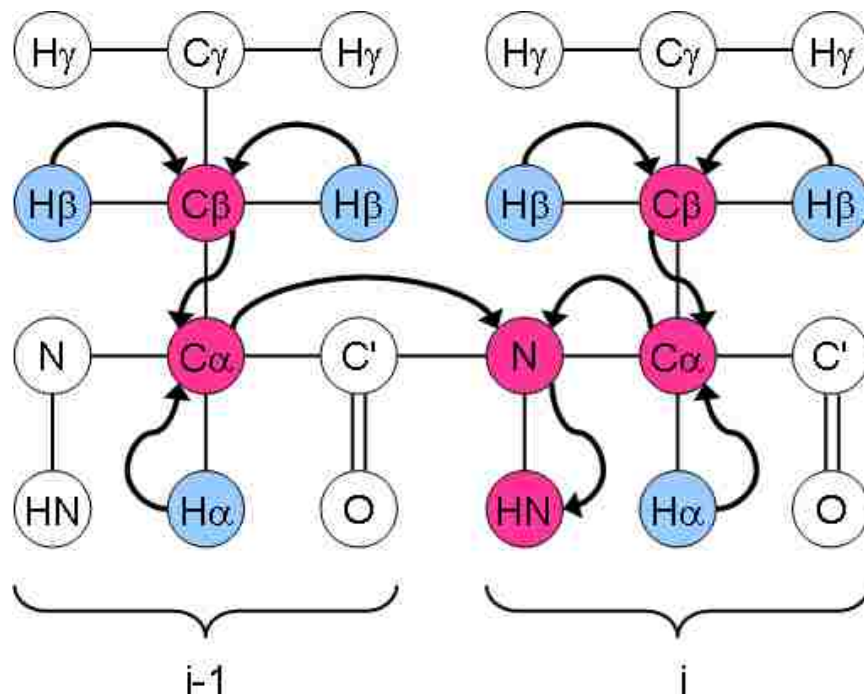


Figure 1-8 CBCANH experiment. http://www.protein-nmr.org.uk/pictures/experiment_types/cbcanh.png

Chapter 2: *E. coli* sabotages the *in vivo* production of O-linked-N-acetylglucosamine (O-GlcNAc) modified proteins

2. 1 Introduction

cAMP response element binding protein (CREB1) is a transcription factor, that plays a major role in the nervous system and brain. CREB1 is present in both neurons and glial cells, which are the major components of the nervous system. CREB1 is well known for its essential role in long-term memory and learning [41]. CREB1 is a human protein that contains 341 amino acids divided into 4 domains. It includes 2 hydrophobic and glutamine rich domains: Q1 at the *N*-terminus and Q2 in the middle of the sequence. A kinase-inducible domain, KID, separates the Q1 and Q2 domains. The bZip, basic leucine zipper, domain is at the *C*-terminus. CREB1 binds to DNA and activates expression of proteins that contribute to synaptic plasticity [42]. Synaptic plasticity refers to the process by which synapses are modified in structure and function in response to different stimuli or environmental cues. Activity-dependent synaptic plasticity is thought to be the underlying process of learning and memory [43]. Josselyn *et al.* showed the first direct evidence that increasing CREB1 levels in a defined mammalian brain region enhances long-term memory formation [44].

CREB1 is O-GlcNAcylated *in vivo*. The modification occurs at two residues between amino acids 256-261 in the Q2 domain at S¹er-260, Thr-256, Thr-259, or Thr-261 [41]. Transcriptional activity is repressed when CREB1 is modified with O-GlcNAc because O-GlcNAc it inhibits CREB1's association with the coactivator, TATA-

¹ This chapter previously appeared as: Octavia Y. Goodwin, Maggie S. Thomasson, Aaron J. Lin, Michelle M. Sweeney, Megan A. Macnaughtan. *E. coli* sabotage the *in vivo* production of O-linked- β -N-acetylglucosamine-modified proteins. *Journal of Biotechnology* 168 (2013) 315-323.

associated factor (TAF_{II}130) [41].

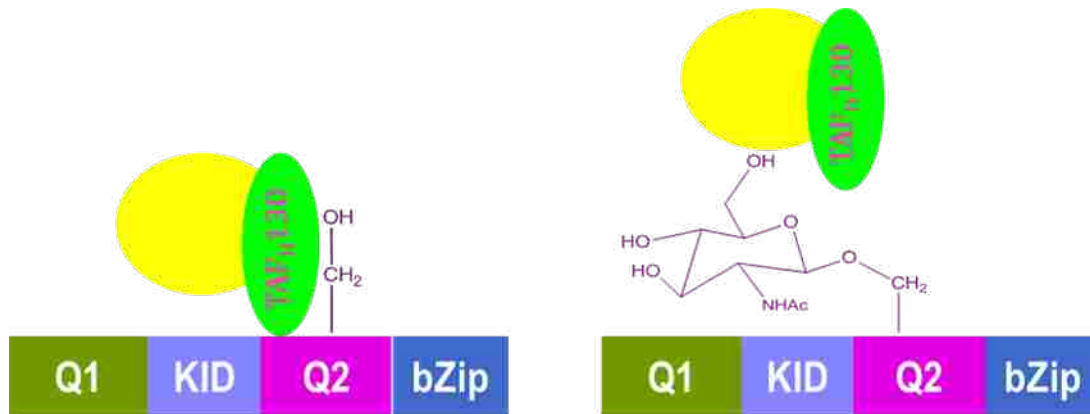


Figure 2-1 The O-GlcNAc modification of CREB1 disrupts binding with TAF_{II}130. Recent studies have shown that inhibition of O-GlcNAcylation leads to enhanced axonal and dendritic growth and long-term memory [18]. A study done in mice showed that mice expressing O-GlcNAcylated CREB1 had a decrease in long-term memory compared to mice that expressed non-glycosylated CREB1. In this chapter I describe how CREB1 was cloned, expressed, and glycosylated *in vivo* for structural analysis.

2. 2 Materials and methods

2. 2. 1 Materials

Ovalbumin, hen egg white lysozyme, *N*-acetylglucosaminidase from *Canavalia ensiformis* (Jackbean), fluorescein-conjugated wheat-germ agglutinin (FITC-WGA), glycerol, maleic acid, O-(2-acetamido-2-deoxyd-glucopyranosylidene)amino-*N*-phenylcarbamate (PUGNAc), 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS), sodium chloride (NaCl), sodium phosphate dibasic, sodium phosphate monobasic, bromophenol blue, dimethylformamide, 4-methylumbelliferyl-*N*-acetylglucosaminide (MU-GlcNAc), sodium carbonate, p-nitrophenyl- β -*N*-acetylglucosamine (pNP-GlcNAc), Coomassie brilliant blue R250, sodium dodecyl sulfate (SDS), glycine,

tris(hydroxymethyl)aminomethane (tris), 2-mercaptoethanol, His-Select Ni-affinity gel, imidazole, urea, acetic acid, methanol, and kanamycin were purchased from Sigma-Aldrich. pET-28b vector, Rosetta(DE3) cells, Tuner(DE3) cells, BugBuster protein extraction reagent, and ampicillin were purchased from EMD Millipore/Novagen. A UNO S1 cation exchange column, Precision Plus Protein Unstained Standards, 4–20% mini-PROTEAN TGX precast gels, immun-blot low fluorescence PVDF membrane, and Laemmli buffer were purchased from Bio-Rad. 2-log DNA ladder, 10-beta cells, Taq 5X master mix, T4 DNA ligase, NdeI restriction enzyme, and XhoI restriction enzyme were purchased from New England Biolabs. BioReagents EZ-Run Rec Protein Ladder, anti-O-GlcNAc MAb CTD 110. 6, and DyLight 550-conjugate Goat anti-Mouse IgG secondary antibody were purchased from Thermo Scientific/Pierce Antibodies. Luria broth (LB) and agar were purchased from BD Biosciences. Artic Express(DE3) cells and BL21-Gold(DE3) cells were purchased from Agilent Technologies, Strategene Products Division. OverExpress C41(DE3) cells and OverExpress C43(DE3) cells were purchased from Lucigen. Other chemicals and materials used include a glycopeptide (SVES(β -O-GlcNAc)GSADAK-NH₂, Sussex Research), deuterated water (D₂O, Cambridge Isotope Laboratories), isopropyl-D-1-thiogalactopyranoside (IPTG, Affymetrix Anatrace), primers (Integrated DNA Technologies), a TSKgel G3000 SWxl gel filtration column (Tosoh Bioscience LLC), full-length cDNA sequence of human cAMP responsive element-binding protein in pCMV-Sport 6 vector (CREB1, Invitrogen), QIAprep spin miniprep kit (Qiagen), Costar 96 well UV transparent flat bottom plates, 4 mm susceptibility matched NMR tube (D₂O-matched, Shigemi Inc.), K-12 *nagZ* knockout strain (JW1093-1) and K-12 parent strain (BW25113) (Keio Collection, *E. coli*

Genetic Resources at Yale, CGSC, The Coli Genetic Stock Center) [45], and clone HsCD00331895 of the F-actin binding domain of Abelson tyrosine-kinase 2 (ABL2-FABD) in pET15 vector (DNASU Plasmid Repository, PSI:Biological-Materials Repository) [46]. All water used was supplied from a Millipore Direct-Q 3 ultrapure water system.

2. 2. 2 Cloning and expression vectors

The human O-GlcNAc transferase (OGT) was expressed using the pET-43. 1 Ek/LIC vector with the full-length sequence provided by Dr. John Hanover [47]. The pET-43. 1 Ek/LIC-OGT includes an *N*-terminal NusA-, 6XHis-, and S-tags. The sequence of three substrate protein domains were cloned into the expression vector, pET-28b, with an *N*-terminal 6XHis tag: the second glutamine-rich domain, Q2, of human cAMP responsive element-binding protein (CREB1-Q2, amino acids 160–283), the Q2 and bZIP domains of CREB1 (CREB1-Q2-bZIP, amino acids 160–341), and the F-actin binding domain of Abelson tyrosine-kinase 2 (ABL2-FABD). CREB1-Q2 and CREB1-Q2-bZIP were cloned from cDNA of the full-length sequence using PCR primers and the restriction enzymes *Nde*I and *Xho*I. The forward primer used was 5'-TACGTCCATATGACTTCAGCACCTGCCATC-3', and the reverse primer for CREB1-Q2 was 5'-TAGCTGCTCGAGTCATGCTGCTTCTTCAGCAGG-3' and for CREB1-Q2-bZIP was 5'-TAGCTGGCGGCCGCTCAATCTGATTTGTGGCAGTA-3'. ABL2-FABD was cloned from a pET-15 vector (clone HsCD00331895, PSI:Biological-Materials Repository) using restriction enzymes, *Nde*I and *Xho*I. The pET-43. 1 Ek/LIC-OGT, pET-28b-CREB1-Q2, pET-28b-CREB1-Q2-bZIP, and pET-28b-ABL2FABD were each transformed into 10-beta *E. coli* cloning cells. Each plasmid was purified from a 5 mL

culture and sequenced (GeneLab DNA sequencing service at Louisiana State University).

2. 2. 3 *E. coli* cultures and protein co-expression

BL21-Gold(DE3) *E. coli* competent cells were transformed with pET-43. 1-OGT and a substrate vector (either pET-28b-CREB1-Q2, pET-28b-CREB1-Q2-bZIP, or pET-28b-ABL2-FABD), and selected on agar plates supplemented with ampicillin and kanamycin. The expression strains were each grown in LB media at 37 °C. At an optical density at 600 nm of 0. 6–0. 8, the culture was induced with 0. 5 mM IPTG and grown at 37 °C for 3 h, unless otherwise stated. The cells were harvested using BugBuster protein extraction reagent and French press and/or ultrasonication followed by centrifugation at 30,000 g for 30 min at 4 °C. For some experiments, OGT and the substrate proteins were purified using His-Select Ni-affinity gel media using gravity flow. Cleared lysate was loaded onto 1 mL of Ni-affinity media equilibrated with 50 mM sodium phosphate, 300 mM NaCl, pH 8. 0 buffer. The media was washed with 20 mL of buffer with 10 mM imidazole, and the protein was eluted with 10 mL of buffer with 300 mM imidazole. Having 6XHis-tags on both OGT and the substrate proteins allows for simultaneous purification and evaluation of protein expression levels using SDS-PAGE. OGT (~170 kDa) and the substrate proteins (<25 kDa) can be separated using gel filtration chromatography if needed, or alternatively, OGT can be cloned without the 6XHis tag. Separation of OGT and the substrate proteins was not necessary for the experiments presented in this chapter.

BL21-Gold(DE3) *E. coli* with empty pET28b vector, K-12 parent *E. coli*, and K-12 Δ naqZ *E. coli* were each grown in 250 mL of LB media at 37 °C. At an optical density at 600 nm of 0.6, 1.0 mM IPTG was added to the BL21-Gold(DE3) culture only and all cultures were incubated for 3 h at 37 °C. The cells from each culture were harvested at 3000 g for 10 min, suspended in 12 mL of Bugbuster protein extraction reagent, and lysed using French press. Cleared lysate was prepared by centrifugation at 30,000 g for 30 min.

2. 2. 4 O-GlcNAc detection and quantification

SDS-PAGE gels and blots were used to determine the relative amount of O-GlcNAc-modified protein compared to total protein. Samples of cleared lysate and insoluble pellet were prepared for SDS-PAGE in Laemmli buffer with 5% v/v 2-mercaptoethanol and incubated at 95 °C for 15 min. The samples were loaded into a 4–20% gel and run at 125 V for 60 min in 25 mM tris, 192 mM glycine, 0.1% SDS, pH 8.3 buffer. Protein ladder was run on each gel, and duplicate gels were run when blot analysis was needed. The gel was stained with Coomassie Brilliant Blue R250 solution (0.1% R250, 50% methanol, 40% water, 10% acetic acid), de-stained (50% water, 40% methanol, 10% acetic acid), and imaged using a GE Typhoon 9410 imager or Fotodyne Apprentice system. The contrast and brightness of the images were adjusted using Microsoft Powerpoint software or ImageJ software to improve the quality of the images for publication [48].

Blotting methods included both lectin and western blots. Fluorescein-conjugated wheat-germ agglutinin was used for the lectin blots. Anti-O-GlcNAc MAb CTD 110.6 and DyLight 550 conjugate Goat anti-Mouse IgG secondary antibody were used for the

western blots. Hen egg white lysozyme and ovalbumin were used as positive controls for the western blots and lectin blots, respectively. All blots were imaged using the GE Typhoon 9410 imager in the Louisiana State University AgCenter Biotechnology Laboratory, Protein Facility. ImageJ software was used to measure the band densities from the Coomassie-stained SDS-PAGE gels (relative total protein concentration) and from the blots (relative O-GlcNAc-modified protein concentration) [48]. The background signal was subtracted and the ratios of the blot to gel densities were calculated. The error calculated from multiple background measurements was propagated and is represented as error bars.

2. 2. 5 Partial purification of *E. coli* β -*N*-acetylglucosaminidase

E. coli β -*N*-acetylglucosaminidase was partially purified from the cleared lysate of BL21-Gold(DE3) expressing empty pET-28b, K-12 parent, and K-12 Δ *nagZ* cultures using cation exchange chromatography (CXC). A BioLogic DuoFlow 10 system with UV detection and a fraction collector was used for all chromatography experiments. Cleared lysate (12 mL) was separated using a UNO S1 cation exchange column into 1 mL fractions using a buffer gradient of 20 mM sodium phosphate pH 7. 0, from 0 to 0. 5 M NaCl. The fractions were analyzed for β -*N*-acetylglucosaminidase activity using the MU-GlcNAc fluorescent well-plate assay described in Section 2. 6. The active fractions from the cleared lysate of BL21-Gold(DE3) expressing empty pET-28b were combined and used to determine enzyme kinetics and activity toward a glycopeptide. The K-12 parent and K-12 Δ *nagZ* fractions were used to test for other β -*N*-acetylglucosaminidase activity toward MU-GlcNAc. The cleared lysate from the K-12 Δ *nagZ* culture was used to test for other β -*N*-acetylglucosaminidase activity toward a glycopeptide.

2. 2. 6 *E. coli* β -*N*-acetylglucosaminidase detection assays

Spectrophotometric and fluorescence well-plate methods were used to measure β -*N*-acetylglucosaminidase activity. A solution of a chromophoric, β -*N*-acetylglucosaminidase substrate, pNP-GlcNAc, was prepared in 50 mM tris-maleate, 20% glycerol, pH 7. 7 at 1. 0 mg/mL (2. 9 mM) and preheated to 37 °C. In a 96 wellplate, 20 μ L of sample (cleared lysate or chromatography fraction) was added to 90 μ L of the pNP-GlcNAc solution (final concentration was 2. 4 mM) and incubated at 37 °C for 30 min. The reaction was quenched with 200 μ L of 42 mg/mL (396 mM) sodium carbonate, which stopped the reaction and converted any cleaved p-nitrophenol to yellow nitrophenolate. The absorbance at 405 nm was measured with a Biotek PowerWave XS microplate spectrophotometer and converted to nmoles of nitrophenolate formed per mL of fraction/min using an extinction coefficient of 18,100 M⁻¹cm⁻¹ for nitrophenolate [49]. An assay blank was run with 20 μ L of buffer (lysis buffer or chromatography buffer) in place of the sample; and, a sample blank was run with 90 μ L of tris-maleate buffer in place of substrate. Each sample and blank were run in triplicate where indicated. For triplicate measurements, the average absorbance and standard deviation for each sample and blank were calculated. Detection of β -*N*-acetylglucosaminidase activity in the chromatography fractions were done by Maggie Thomasson.

For the fluorescent β -*N*-acetylglucosaminidase activity assay, a stock solution of 4-methylumbelliferyl- β -*N*-acetylglucosaminide, MU-GlcNAc, was prepared at 25 mM in dimethylformamide. A working solution of 1. 5 mM MU-GlcNAc for well plate assays was made by diluting the stock solution in 100 mM sodium phosphate buffer, pH 7. 4.

The working solution was preheated to 37 °C. For the 96 well-plate assay, 20 µL of sample (cleared lysate or chromatography fraction) was added to 90 µL of the MU-GlcNAc working solution (final MU-GlcNAc concentration was 1.2 mM) and incubated at 37 °C for 30 min. The fluorescence intensity was measured with an excitation wavelength of 372 nm and emission wavelength of 445nm using a Perkin Elmer Wallac Victor² V fluorescence plate reader, housed in the Louisiana State University AgCenter Biotechnology Laboratory, Protein Facility. A positive control was run using 1 µL of β-*N*-acetylglucosaminidase from *C. ensiformis* (Jack bean) in 19 µL of buffer (lysis buffer or chromatography buffer) with 90 µL of the working MU-GlcNAc solution. A negative control was run using 20 µL of buffer (lysis buffer or chromatography buffer) with 90 µL of the working MU-GlcNAc solution. Each sample and blank were run in triplicate where indicated. For triplicate measurements, the average fluorescence intensity and standard deviation for each sample and blank were calculated. This fluorescent assay was done by Maggie Thomasson.

2. 2. 7 Kinetic assays

The Michaelis–Menten constants (K_m) for MU-GlcNAc and pNP-GlcNAc and the inhibition constant (K_i) for PUGNAc were determined for the *E. coli* β-*N*-acetylglucosaminidase using the pooled β-*N*-acetylglucosaminidase-active CXC fractions (see Section 2. 5). For the pNP-GlcNAc reactions, three replicate samples of 0, 0.05, 0.5, 5 mM pNP-GlcNAc and 80 µL of active CXC fraction were mixed in 50 mM sodium phosphate, 500 mM sodium chloride, pH 7.1 to a final volume of 100 µL. In addition to the sample blank (0 mM pNP-GlcNAc), an assay blank was made in triplicate with 5 mM pNP-GlcNAc in buffer. The samples were incubated at 37 °C. One set of the

triplicate samples was quenched with 200 μL of 396 mM sodium carbonate at 10, 15, and 20 min. The formation of product was measured as described in Section 2. 6.

The K_m for MU-GlcNAc and the K_i for PUGNAc were determined using sets of three replicate samples at 0, 13. 6, 16. 4, 20. 4, 27. 3, 40. 9, and 81. 8 μM MU-GlcNAc with 0, 200, and 500 nM PUGNAc. Each sample volume was 110 μL with 20 μL of active CXC fraction and 100 mM sodium phosphate, pH 7. 4 buffer. The fluorescence intensities of the samples were measured at room temperature after 4, 9, and 14 min using a fluorescence plate reader as described in Section 2. 6. The kinetics assay was done by Maggie Thomasson.

2. 2. 8 Kinetic analysis

Initial velocities (V) at each substrate (S) concentration were determined using least squares fitting to a linear function, and the average values were calculated for triplicate measurements. The data were then analyzed by nonlinear regression analysis using the computer programs of Cleland (1979). Initial velocities were fit to Eq. (1) to determine the maximal velocity (V_{max}) and the Michaelis–Menten constant (K_m).

$$\text{Eq. (1)} \quad V = \frac{V_{\text{max}}S}{K_m + S}$$

Data for competitive inhibition were fit to Eq. (2), where I is the concentration of PUGNAc and K_i is the inhibition constant. The kinetics analysis was done by Maggie Thomasson.

$$v = \frac{V_{max}S}{K_m + (1 + I/K_i)S}$$

Eq. (2)

2. 2. 9 Inhibition of *E. coli* β -*N*-acetylglucosaminidase activity

Cleared lysate from a BL21-Gold(DE3) culture was tested for β -*N*-acetylglucosaminidase activity using the pNP-GlcNAc spectrophotometric assay described in Section 2. 6. PUGNAc was added to 20 μ L of lysate at varying concentrations (0, 1, and 2 μ M PUGNAc) and mixed with 90 μ L of 1. 0 mg/mL pNP-GlcNAc. Conditions, blanks, and measurements are the same as described in Section 2. 6.

Changes in the relative amount of O-GlcNAc-ABL2-FABD were measured for co-expression cultures grown with 0, 1, 2, and 10 μ M PUGNAc. PUGNAc was added to the growth media of BL21-Gold(DE3) cells co-expressing pET43. 1 Ek/LIC-OGT and pET28b-ABL2-FABD. The cultures were grown, harvested, and lysed as described in Section 2. 3. The soluble and insoluble proteins from each culture were analyzed with SDS-PAGE and western blot as described in Section 2. 4. The relative amount of O-GlcNAc was measured as described in Section 2. 4.

2. 2. 10 NMR experiments to detect *E. coli* β -*N*-acetylglucosaminidase hydrolysis of an O-GlcNAc-modified peptide

To evaluate the activity of the *E. coli* β -*N*-acetylglucosaminidase toward protein O-GlcNAc, a synthetic O-GlcNAc-modified peptide (SVES(β -O-GlcNAc)GSADAK-NH₂) was incubated with partially purified, *E. coli* β -*N*-acetylglucosaminidase and K-12 Δ nagZ lysate. Each reaction mixture was analyzed with ¹H NMR to observe the ¹H anomeric

signal of any released α -GlcNAc (5.2 ppm). This signal was chosen because it has relatively little background interference. The reaction mixtures contained 200 μ M glycopeptide, 100 μ L of β -*N*-acetylglucosaminidase-active CXC fraction or K-12 Δ *nagZ* lysate, 50 mM sodium phosphate, 500 mM NaCl, pH 7.1 buffer, 10% D₂O, and trace DSS (to lock and reference the NMR signal, respectively).

A glycopeptide control, sample control, and sample control with 200 μ M *N*-acetylglucosamine (GlcNAc) were prepared with 100 μ L of CXC buffer instead of active fraction or lysate, 3.0 μ L of water instead of glycopeptide stock solution, and 1.3 μ L of 20 mM GlcNAc plus 2.7 μ L of water instead of glycopeptide stock solution, respectively. The samples were either incubated at 37 °C for 5 h (CXC β -*N*-acetylglucosaminidase-active fraction) or 22.5 h (K-12 Δ *nagZ* lysate). The glycopeptide control and sample controls were treated with the same conditions. The reactions were performed one time. The reaction mixtures and controls were analyzed in 4 mm D₂O-susceptibility matched tubes. ¹H NMR spectra of the samples were collected on a 700 MHz Varian NMR equipped with a 5 mm HCN probe. The spectra were collected at 20 °C with presaturation of water (1 s saturation time, 12 dB saturation power), 1 s acquisition time, 1 s recycle delay, 7000 Hz spectral width, 30 dB gain, and 2048 scans. Data processing, line broadening (1 Hz exponential), chemical shift referencing, baseline correction, peak height measurements, and signal-to-noise calculations were performed with VnmrJ 3.2 software (Agilent). The amount of released GlcNAc was estimated by comparing the peak intensities at 5.198 ppm (left peak of the anomeric ¹H α -GlcNAc doublet) of the reaction mixture to the control sample with 200 μ M GlcNAc. For the glycopeptide reaction with K-12 Δ *nagZ* lysate, the limit of quantitation (S/N = 10) was

calculated using the S/N of the ^1H anomeric α -GlcNAc peak in the control sample with 200 μM GlcNAc.

2. 3 Results and discussion

2. 3. 1 Co-expression of OGT with substrate proteins using *E. coli*

The second glutamine-rich domain of human CREB1 (CREB1-Q2) and the Q2 and bZIP domains of CREB1 (CREB1-Q2-bZIP) were co-expressed with full-length human OGT in BL21-Gold(DE3) *E. coli*. Even though these vectors have the same origin of replication, the different antibiotic resistance prevents incompatibility [50]. The CREB1-Q2 and CREB1-Q2-bZIP domains were chosen as substrate proteins because the Q2 domain is natively modified with at least two O-GlcNAc molecules that suppress CREB1-activated transcription [17]. The co-expression cultures were lysed and separated into soluble lysate and insoluble pellet fractions. The soluble lysate and pellet were analyzed with Coomassie SDS-PAGE and lectin blot. The results for the co-expression of OGT with CREB1-Q2 and CREB1-Q2-bZIP are shown in Figures 2-2 and 2-3, respectively. Figures 2-2a and 2-3a show the SDS-PAGE gel images for Ni-affinity purified CREB1-Q2 and CREB1-Q2-bZIP soluble lysate. CREB1-Q2 is highly expressed and soluble, while CREB1-Q2-bZIP is mainly found in the pellet. The additional bands in the Ni-affinity purified lysate in Figure 2-2a are *E. coli* background proteins and 6XHis-tagged OGT (with multiple translational truncations due to rare codons). Figures 2-2b and 2-3b show the lectin blots corresponding to the SDS-PAGE gels. The dark bands indicate where FITC-WGA bound O-GlcNAc-modified proteins. The presence of O-GlcNAc was also confirmed with an anti-O-GlcNAc MAbs CTD 110. 6 western blot of Ni-affinity purified protein from the expression of CREB1-Q2-bZIP (pET-28b-CREB1-Q2-

bZIP) without and with co-expression of OGT (pET-43. 1-OGT) in *E. coli* BL21-Gold(DE3) (see Supplementary Figure 1). In both lectin blots (Figures 2-2 and 2-3), only insoluble CREB1-Q2 and CREB1-Q2-bZIP were modified with O-GlcNAc.

The absence of soluble, O-GlcNAc-modified protein from the co-expression system was unexpected. Others have shown that *E. coli* co-expression of OGT with substrate protein produce O-GlcNAc-modified proteins *in vivo*, but the solubility of the modified proteins was not specifically addressed in these studies [51-53].

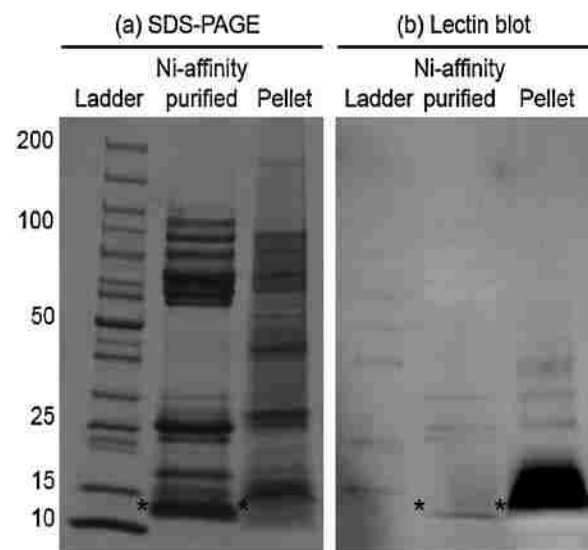


Figure 2-2 The Ni-affinity purified soluble lysate and pellet from co-expression of CREB1-Q2 (pET-28b-CREB1-Q2) and OGT (pET-43. 1-OGT) in *E. coli* BL21-Gold(DE3) analyzed with (a) Coomassie-stained SDS-PAGE and (b) FITC-WGA lectin blot detecting O-GlcNAc. The bands corresponding to CREB1-Q2 are indicated with asterisks.

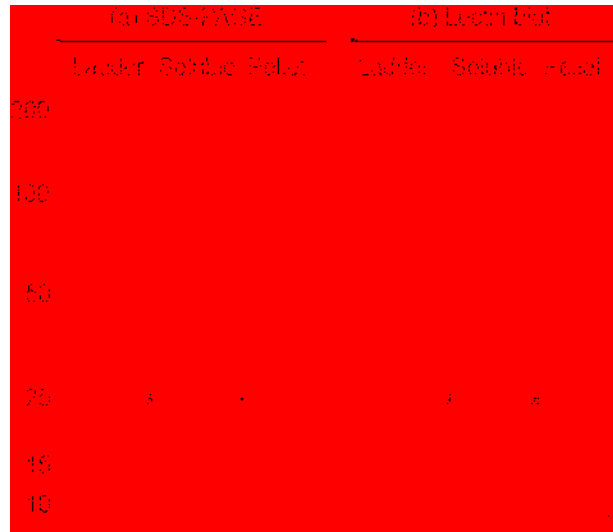


Figure 2-3 The soluble lysate and pellet from co-expression of CREB1-Q2-bZIP (pET28b-CREB1-Q2-bZIP) and OGT (pET-43. 1-OGT) in *E. coli* BL21-Gold(DE3) analyzed with (a) Coomassie-stained SDS-PAGE and (b) FITC-WGA lectin blot detecting O-GlcNAc. The bands corresponding to CREB1-Q2-bZIP are indicated with asterisks.

From Figure 2-3a, CREB1-Q2-bZIP shows little soluble protein in the SDS-PAGE image, so observing only O-GlcNAc-modified CREB1-Q2-bZIP in the pellet is not informative. However, the solubility of the CREB1-Q2 domain is good, but the O-GlcNAc-modified protein only occurs in the pellet. To improve the solubility of the O-GlcNAc-modified proteins, the co-expression of OGT with CREB1-Q2 and CREB1-Q2-bZIP was tested using additional *E. coli* expression strains at various conditions. Rosetta, C41, and C43 *E. coli* strains were induced with 0.5 mM IPTG and grown at 37 °C for 3 h and 18 °C for 16 h. The Tuner *E. coli* strain was grown at the same temperatures but with 0.25 mM–1 mM IPTG. Arctic express *E. coli* cells were induced with 0.5 mM IPTG and grown at 4 °C and 12 °C for 24 h. Regardless of the type of *E. coli* expression strain or conditions used, the O-GlcNAc-modified CREB1-Q2-bZIP and CREB1-Q2 proteins were only found in the pellet (data not shown).

To circumvent the possibility that O-GlcNAc-modified CREB1-Q2 and CREB1-Q2-bZIP are inherently insoluble, the F-actin binding domain of Abelson tyrosine-kinase 2 (ABL2-FABD, also known as ARG, Abelson-related gene protein) was selected as a third test protein. ABL2-FABD was selected because it expresses in high yield with high solubility in *E. coli*, actin-binding proteins have been shown to be regulated by O-GlcNAc, and ABL2-FABD is predicted to have two sites of O-GlcNAc modification [54-57].

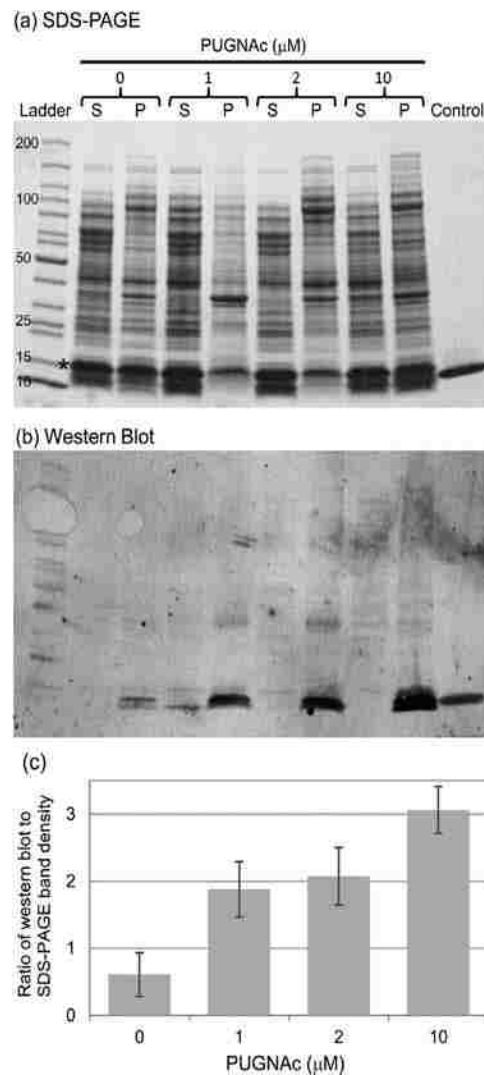


Figure 2-4 ABL2-FABD (pET-28b-ABL2-FABD) and OGT (pET-43. 1-OGT) were co-expressed in *E. coli* BL21-Gold(DE3) with increasing concentrations of PUGNAc in the

growth media (0, 1, 2, and 10 μ M PUGNAc). The soluble lysate (S) and pellet (P) from each culture was analyzed with (a) Coomassie-stained SDS-PAGE and (b) anti-O-GlcNAc MAb CTD 110. 6 western blot. HEWL was used as a positive control in the last lane. (c) A plot of the relative band density of the western blot band to the Coomassie-stained SDS-PAGE gel band (indicated with an asterisk). For each band in (a) and (b), the background density was measured thrice, subtracted from each band, and used to estimate the propagated measurement error (represented as error bars in (c)).

Figure 2-4a (Lanes 2 and 3) shows an SDS-PAGE image of the soluble lysate and pellet from ABL2-FABD co-expressed with OGT. The ABL2-FABD protein is abundant in both the lysate and pellet, but the western blot image in Figure 2-4b (Lanes 2 and 3) shows small amounts of O-GlcNAc-modified ABL2-FABD only in the pellet. The absence of soluble O-GlcNAc-modified protein from the co-expression system suggests two possibilities: (1) the modification causes the proteins to precipitate and/or (2) the O-GlcNAc modification is being cleaved by an endogenous *E. coli* enzyme in the cytoplasm. The second possibility assumes that the insoluble, O-GlcNAc-modified proteins are protected from the suspected *E. coli* β -*N*-acetylglucosaminidase activity.

2. 3. 2 Inhibiting NagZ activity in *E. coli* increases *in vivo* O-GlcNAc modification

NagZ, the only *E. coli* β -*N*-acetylglucosaminidase identified, is suspected to be the enzyme that deglycosylates soluble O-GlcNAc-modified protein from the co-expression system. The existence of an O-GlcNAc transferase in *E. coli* has been ruled out but the *E. coli* β -*N*-acetylglucosaminidase, NagZ, has not been tested for activity toward O-GlcNAc-modified peptides or proteins [58]. NagZ was first purified from K-12 *E. coli* and characterized using pNP-GlcNAc in a spectrophotometric assay [59]. The soluble, cytoplasmic enzyme was found to specifically cleave O- β -linked GlcNAc. Later characterization of K-12 NagZ demonstrated its role in peptidoglycan recycling by cleaving the GlcNAc-(1,4)-MurNAc bond (MurNAc, *N*-acetylmuramic acid) with a K_m of

35 μM for the anhydro-disaccharide [60, 61]. A null mutant showed that NagZ was not essential to *E. coli* growth and morphology and was the only *E. coli* β -*N*-acetylglucosaminidase [59, 60]. It is possible that NagZ has activity toward O-GlcNAc-modified proteins because it is similar to human O-GlcNAcase (OGA) in substrate (pNP-GlcNAc) and inhibitor (PUGNAc). Human OGA cleaves pNP-GlcNAc and is inhibited by PUGNAc [62, 63]. K-12 NagZ, which has 99% amino acid identity to BL21-Gold(DE3) NagZ, has a pNP-GlcNAc K_m of 0.31–0.43 mM [59, 61]. PUGNAc is a potent inhibitor of *Vibrio cholerae* NagZ ($K_i = 48$ nM), a homolog of *E. coli* NagZ with 56% amino acid sequence identity [64]. Following the hypothesis that NagZ cleaves O-GlcNAc from modified proteins in the co-expression system, PUGNAc was tested as an inhibitor of the enzyme *in vitro* and *in vivo*.

The MU-GlcNAc K_m and the PUGNAc K_i of BL21-Gold(DE3) *E. coli* β -*N*-acetylglucosaminidase were determined using partially purified β -*N*-acetylglucosaminidase from BL21-Gold(DE3) lysate. β -*N*-Acetylglucosaminidase-active CXC fractions were identified with the MU-GlcNAc well-plate assay and pooled together (Supplementary Figure 2). The MU-GlcNAc K_m and the PUGNAc K_i were determined by measuring the formation rate of 4-methylumbelliferone with a fluorescence plate reader at varying MU-GlcNAc and PUGNAc concentrations. The data shown in Figure 2-5 were globally fit to competitive, noncompetitive, and uncompetitive inhibition models using nonlinear regression analysis, and the competitive model gave the best fit. This result is expected since PUGNAc is a competitive inhibitor of other β -*N*-acetylglucosaminidases and has been shown to bind at *V. cholerae* NagZ's active site [62, 64, 65]. The best fit to Eqs. (1) and (2) give a K_m of 5.6 ± 1.2 μM MU-GlcNAc and a K_i of 25.5 ± 4.6 nM

PUGNAc. These results are similar to the K_m of K-12 NagZ (35 μM anhydro-disaccharide) and the K_i of *V. cholerae* NagZ ($K_i = 48$ nM PUGNAc), respectively [61, 64].

Even though PUGNAc is a strong inhibitor of partially purified *E. coli* β -*N*-acetylglucosaminidase, the effectiveness of the inhibitor is reduced in cleared lysate. The *E. coli* β -*N*-acetylglucosaminidase activity in cleared lysate was tested with the addition of 0, 1, and 2 μM PUGNAc.

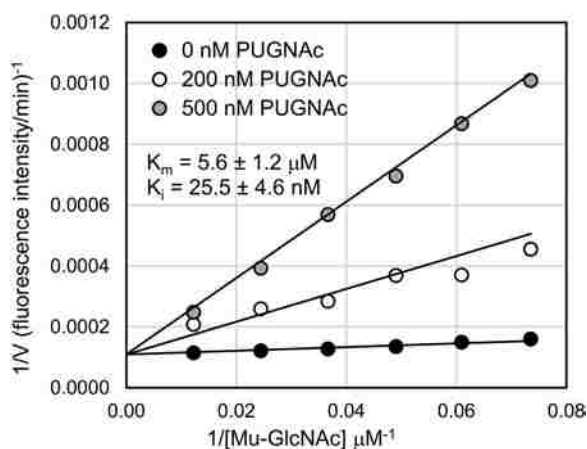


Figure 2-5 The initial velocities of *E. coli* β -*N*-acetylglucosaminidase, partially purified with CXC from BL21-Gold(DE3) cleared lysate, at varying MU-GlcNAc concentrations and with different concentrations of PUGNAc. The points represent the observed velocities, and the lines represent the best fits of the data to Eqs. (1) and (2).

The results of this assay (Figure 2-6) show that PUGNAc inhibits the β -*N*-acetylglucosaminidase activity, but not as well as expected for a K_i of 25.5 nM. To rule out the possibility that *E. coli* β -*N*-acetylglucosaminidase has a low pNP-GlcNAc K_m , the K_m was measured using partially purified β -*N*-acetylglucosaminidase from BL21-Gold(DE3) lysate. The best fit of the data (Supplementary Figure 3) to Eq. (1) results in a K_m of 0.35 ± 0.05 mM, which is within the range of reported values for K-12 NagZ (0.31–0.43 mM) and does not explain the reduced effectiveness of PUGNAc in the

cleared lysate [59, 61]. Another *E. coli* β -*N*-acetylglucosaminidase with activity for pNP-GlcNAc could also explain these results, but this hypothesis is ruled-out by the lack of activity of the null mutant (Section 3. 4) and by work published previously [60]. Another possibility is that the effective concentration of PUGNAc is reduced in the lysate compared to the CXC active fraction; other enzymes or cellular material in the lysate may sequester or degrade PUGNAc.

To assess whether inhibition of the *E. coli* β -*N*-acetylglucosaminidase increases the amount of O-GlcNAc-modified ABL2-FABD produced with the co-expression system, PUGNAc was added to the growth media at 0, 1, 2, and 10 μ M. The relative amount of O-GlcNAc-modified ABL2-FABD produced was determined by normalizing the western blot band density to total ABL2-FABD protein using the band density of a duplicate SDS-PAGE gel image (Figure 2-4a and b). As shown Figure 2-4c, the amount of insoluble, O-GlcNAc-modified ABL2-FABD increases with increasing PUGNAc, consistent with an *E. coli* β -*N*-acetylglucosaminidase cleaving O-GlcNAc *in vivo*. Even with 10 μ M PUGNAc, soluble O-GlcNAc-modified ABL2-FABD is absent (Figure 2-4b). This result is consistent with poor inhibition of *E. coli* β -*N*-acetylglucosaminidase in the cleared lysate with PUGNAc. The possible explanations are the same as discussed in the previous paragraph with the additional possibility of inefficient transport into the cell. Alternatively, the results suggest that soluble, O-GlcNAc-modified ABL2-FABD is unstable. However, the latter is not supported by the fact that the relative levels of soluble to insoluble ABL2-FABD (as assessed by SDS-PAGE (Figure 2-4a)) are not significantly different even when O-GlcNAc is increased. These results taken together indicate that an *E. coli* enzyme, likely NagZ, cleaves O-GlcNAc from the soluble protein,

the enzyme is inhibited by PUGNAc, the effectiveness of PUGNAc inhibition in cleared lysate and *in vivo* is reduced, and the insoluble O-GlcNAc-modified ABL2-FABD is protected from the β -N-acetylglucosaminidase activity.

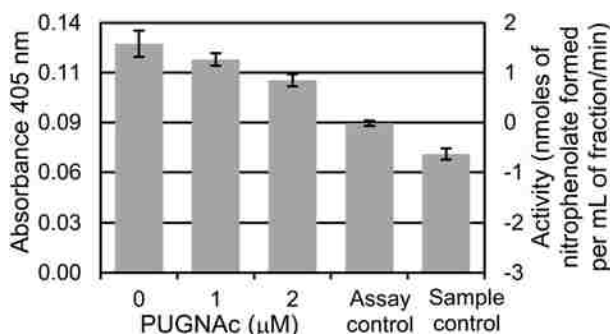


Figure 2-6 The β -N-acetylglucosaminidase activity of cleared lysate with PUGNAc (0, 1, 2 μ M) was measured using the pNP-GlcNAc spectrophotometric assay. The lysate was from the co-expression of ABL2-FABD (pET-28b-ABL2-FABD) and OGT (pET43. 1-OGT) in *E. coli* BL21-Gold(DE3). Each sample and control were run in triplicate, and the error bars represent the standard deviation of these measurements. The 405 nm absorbance values were converted to activity (right vertical axis) using the extinction coefficient of nitrophenolate.

2. 3. 3 *E. coli* NagZ cleaves O-GlcNAc from a peptide

To determine if *E. coli* β -N-acetylglucosaminidase can deglycosylate an O-GlcNAc-modified peptide, the release of GlcNAc from a synthetic O-GlcNAc-modified peptide incubated with partially purified *E. coli* β -N-acetylglucosaminidase was monitored with ^1H NMR. The α -GlcNAc anomeric ^1H signal at 5.2 ppm (Figure 2-7, full spectra in Supplementary Figure 4) was monitored for the glycopeptide control, the glycopeptide and active fraction mixture, the active fraction control, and the active fraction control with 200 μ M GlcNAc after incubation at 37 $^\circ\text{C}$ for 5 h. No α -GlcNAc was detected in the glycopeptide or active fraction control samples (Figure 2-7a and d). α -GlcNAc was detected in the mixture of glycopeptide and partially purified β -N-acetylglucosaminidase sample (Figure 2-7c) at a concentration of approximately 40 μ M

GlcNAc. These results demonstrate that the *E. coli* β -*N*-acetylglucosaminidase has activity toward this O-GlcNAc-modified peptide, and likely has activity toward O-GlcNAc-modified proteins *in vivo*.

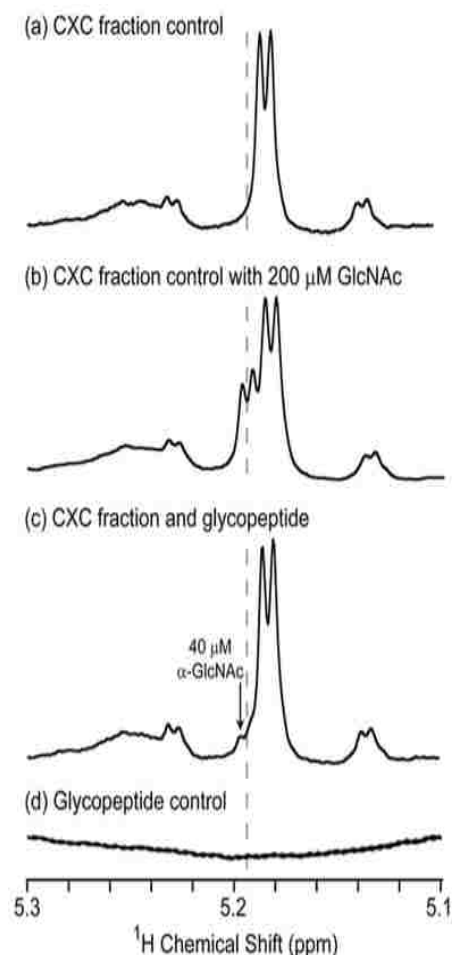


Figure 2-7 Fractions from CXC of BL21-Gold(DE3) lysate with β -*N*-acetylglucosaminidase activity toward MU-GlcNAc were pooled and tested for β -*N*-acetylglucosaminidase activity toward the synthetic O-GlcNAc-modified peptide, SVES(β -O-GlcNAc)GSADAK-NH₂. (a) CXC fraction control, (b) CXC fraction control with 200 μ M GlcNAc added, (c) CXC fraction with 200 μ M glycopeptide, and (d) 200 μ M glycopeptide control in CXC buffer. The spectra shown are expanded regions from 5.1 to 5.3 ppm, with a gray dashed line indicating the α -GlcNAc chemical shift (center of the doublet). All spectra were processed and expanded identically for direct comparison. The full spectra are included in Supplementary Figure 4. The glycopeptide control sample (d) has extra noise due to truncation artifacts from the water peak; the additional noise is small and does not affect spectral interpretation. Based on the ¹H

anomeric α -GlcNAc intensity in (b), approximately 40 μ M GlcNAc was hydrolyzed by the *E. coli* β -N-acetylglucosaminidase in 5 h at 37 °C (c).

2. 3. 4 Identifying the *E. coli* β -N-acetylglucosaminidase as NagZ

To verify that NagZ is the only β -N-acetylglucosaminidase in *E. coli* and the enzyme purified and characterized in these studies, a K-12 Δ nagZ strain and parent strain were tested for β -N-acetylglucosaminidase activity. Cleared lysate from cultures were tested for activity using the MU-GlcNAc fluorescence assay (Supplementary Figure 5). As expected, activity was detected in the K-12 parent strain. No significant activity was found in any of the fractions or the cleared lysate of the K-12 Δ nagZ strain, indicating that NagZ is the only β -N-acetylglucosaminidase in *E. coli*, as Yem and Wu also observed [59]. These results confirm that the β -N-acetylglucosaminidase-active fractions used in the kinetic assays and NMR experiments must have contained NagZ.

To determine if NagZ is the only *E. coli* enzyme with activity toward O-GlcNAc-modified proteins, lysate from K-12 Δ nagZ strain was incubated with 200 μ M O-GlcNAc-modified peptide and monitored with 1 H NMR. Even after 22.5 h of incubation at 37 °C, the 1 H spectrum (Figure 2-8c, full spectra in Supplementary Figure 6) did not show any detectable α -GlcNAc greater than the limit of quantitation (10 μ M). The lack of activity toward the glycopeptide indicates that NagZ is the one and only enzyme that sabotages the *in vivo* production of O-linked β -N-acetylglucosamine-modified proteins with *E. coli*.

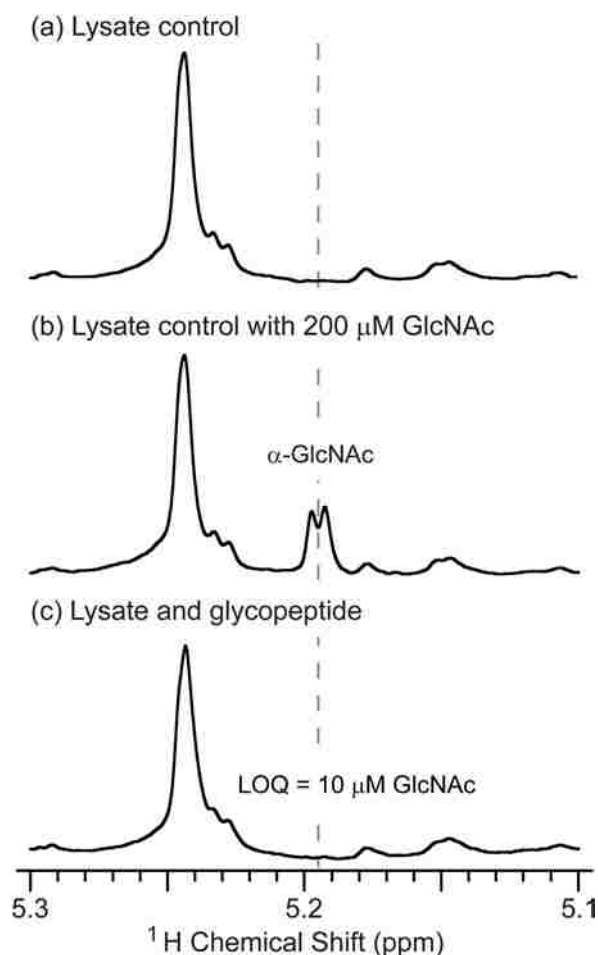


Figure 2-8 ^1H NMR spectra of (a) ΔnagZ K12 lysate, (b) lysate with 200 μM GlcNAc, and (c) lysate with 200 μM glycopeptide, SVES(β -O-GlcNAc)GSADAK-NH₂. All samples were incubated at 37 °C for 22.5 h to test for β -N-acetylglucosaminidase activity toward the synthetic O-GlcNAc-modified peptide. The spectra shown are expanded regions from 5.1 to 5.3 ppm, with a gray dashed line indicating the α -GlcNAc chemical shift (center of the doublet). All spectra were processed and expanded identically for direct comparison. The full spectra are included in Supplementary Figure 6. Based on the signal-to-noise ratio of the 200 μM GlcNAc peak in (b), the limit of quantitation (LOQ) in (c) is 10 μM GlcNAc.

2.4 Conclusions

E. coli has not been used to produce milligram quantities of soluble O-GlcNAc-modified proteins because an endogenous *E. coli* enzyme, NagZ, removes O-GlcNAc from modified proteins *in vivo* and *in vitro*. An *E. coli* co-expression system was developed to modify proteins with O-GlcNAc *in vivo* using human OGT. *E. coli* is a good

host because it lacks endogenous OGT activity, can produce milligram quantities of protein, and allows for metabolic labeling with isotopes for NMR structure determination. Three substrate proteins were successfully O-GlcNAc-modified with the co-expression system: CREB-Q2, CREB-Q2-bZIP, and ABL2-FABD. CREB-Q2 and ABL2-FABD, in particular, are soluble proteins, but the O-GlcNAc-modified proteins were only found in the insoluble pellet. The O-GlcNAc-modification may cause the substrate proteins to become unstable and precipitate; particularly, if the modification is added co-translationally before the protein has folded. Independent of whether O-GlcNAc causes insolubility, several pieces of evidence suggest that *E. coli* NagZ cleaves O-GlcNAc from the soluble substrate protein *in vivo* and *in vitro*. The amount of O-GlcNAc-modified ABL2-FABD produced with the co-expression system increased when a β -*N*-acetylglucosaminidase inhibitor, PUGNAc, was added to the growth media. NMR was used to show that an *E. coli* enzyme cleaves O-GlcNAc from an O-GlcNAc-modified peptide, a foreign substrate. Even though the *E. coli* β -*N*-acetylglucosaminidase was not directly identified with mass spectrometry, the lack of β -*N*-acetylglucosaminidase activity in lysate from a K-12 Δ *nagZ* strain both toward MU-GlcNAc and an O-GlcNAc-modified peptide demonstrates that NagZ is the only *E. coli* β -*N*-acetylglucosaminidase. Moreover, the lack of activity in the knockout confirms that NagZ was present in the samples shown to cleave O-GlcNAc and is inhibited by PUGNAc. The K_m 's of BL21Gold(DE3) NagZ determined, 0.35 mM pNP-GlcNAc and 5.6 μ M MU-GlcNAc, are consistent with values published for K-12 NagZ [59, 61]. The inhibition constant of BL21-Gold(DE3) NagZ was determined to be 25.5 nM and is consistent with *V. cholerae* NagZ, but the mechanism of reduced PUGNAc effectiveness in cleared lysate

and in *in vivo* experiments remains unclear [64]. Efforts are currently underway to test the co-expression system in $\Delta nagZ$ *E. coli* strains for increased yields and soluble O-GlcNAc-modified proteins.

2. 5 Acknowledgments

We thank Dr. John Hanover for his gift of the pET-43. 1 Ek/LIC-OGT vector. A. L. thanks the Louisiana State University HHMI Summer Undergraduate Research Program for support. M. M. thanks Prof. Aaron Smith for proof-reading the manuscript, Prof. Grover Waldrop for advice on enzyme kinetics, Dr. Tyler Broussard for assistance with the global kinetic fitting software, and the Louisiana State University Structural Biology Group for helpful suggestions regarding experiments. This research was funded by Louisiana State University.

Chapter 3: Is NagZ the saboteur?

Chapter 3. 1 Introduction

3. 1. 1 The Hexosamine biosynthesis pathway in *E. coli*

UDP-GlcNAc is an important component in *E. coli*. It is a major precursor of the bacterial cell wall. (3) The enzymes of the hexosamine biosynthesis pathway in *E. coli* are different from those in humans, but the steps of the pathway are similar. Glucose is transported into the *E. coli* cell and phosphorylated by a glucose specific phosphotransferase (PstG), and phospho-glucose isomerase (Pgi) interconverts glucose-6-phosphate to fructose-6-phosphate. Glucosamine-6-phosphate synthase (GlmS) transfers an amino group from glutamine to make glucosamine-6-phosphate, which is then converted to glucosamine-1-phosphate by phosphoglucosamine mutase (GlmM). *N*-acetylglucosamine-1-phosphate is formed through the acetylation of glucosamine-1-phosphate by glucosamine-1-phosphate acetyltransferase (GlmU). Finally, *N*-acetylglucosamine-1-phosphate is uridylylated by *N*-acetylglucosamine-1-phosphate uridylyltransferase (GlmU) to form uridine diphosphate *N*-acetylglucosamine (UDP-GlcNAc). Glucosamine and *N*-acetylglucosamine can enter the HBP as well. Glucosamine is transported into the cell and phosphorylated by mannose transporter (ManXYZ), and *N*-acetylglucosamine is transported and phosphorylated by NagE, a *N*-acetylglucosamine specific transporter. *N*-acetylglucosamine is deacetylated by *N*-acetylglucosamine-6-phosphate deacetylase to glucosamine-6-phosphate where it continues along the HBP [66, 67]. *N*-acetylglucosamine that is cleaved from the cell wall by NagZ is phosphorylated by NagK and used to make more UDP-GlcNAc [68].

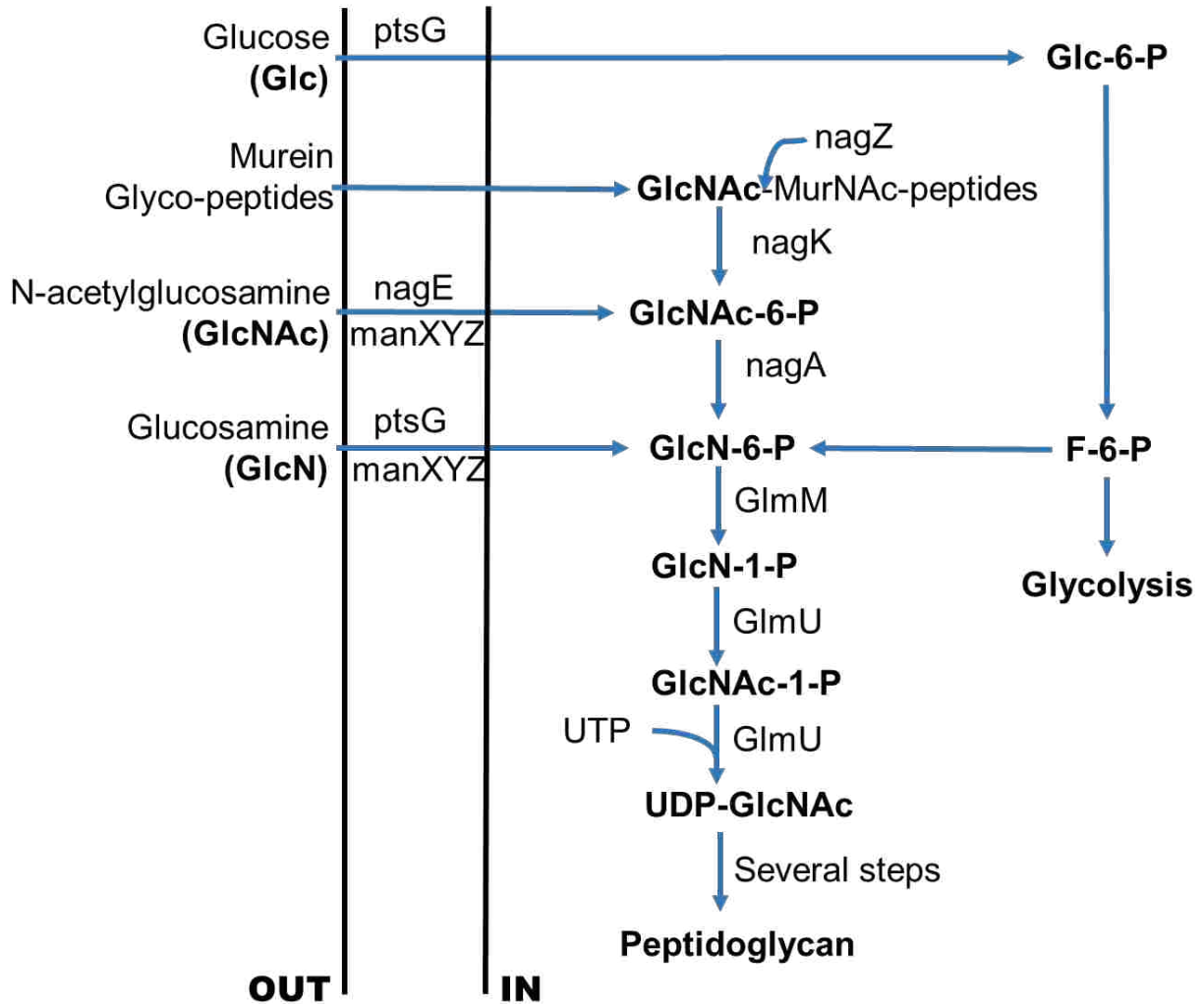


Figure 3-1 Glucose, glucosamine, and N-acetylglucosamine are used by *E. coli*'s hexosamine biosynthesis pathway to make UDP-GlcNAc.

3. 1. 2 The role of β -N-acetylglucosaminidase (NagZ) in *E. coli*

E. coli is a prokaryotic single cell organism with a cellular membrane made of murein also known as peptidoglycan. Peptidoglycan is composed of cross-linked alternating GlcNAc- β 1,4-MurNAc-peptide monomers [68]. The main function of the peptidoglycan layer is to support the cell wall structure and balance the cytoplasm's

osmotic pressure. There are several enzymes that participate in the degradation of the peptidoglycan layer. NagZ is the only β -*N*-acetylglucosaminidase whose function is to cleave *N*-acetylglucosamine (GlcNAc) from *N*-acetylmuramic acid (MurNAc) peptides for peptidoglycan recycling [69].

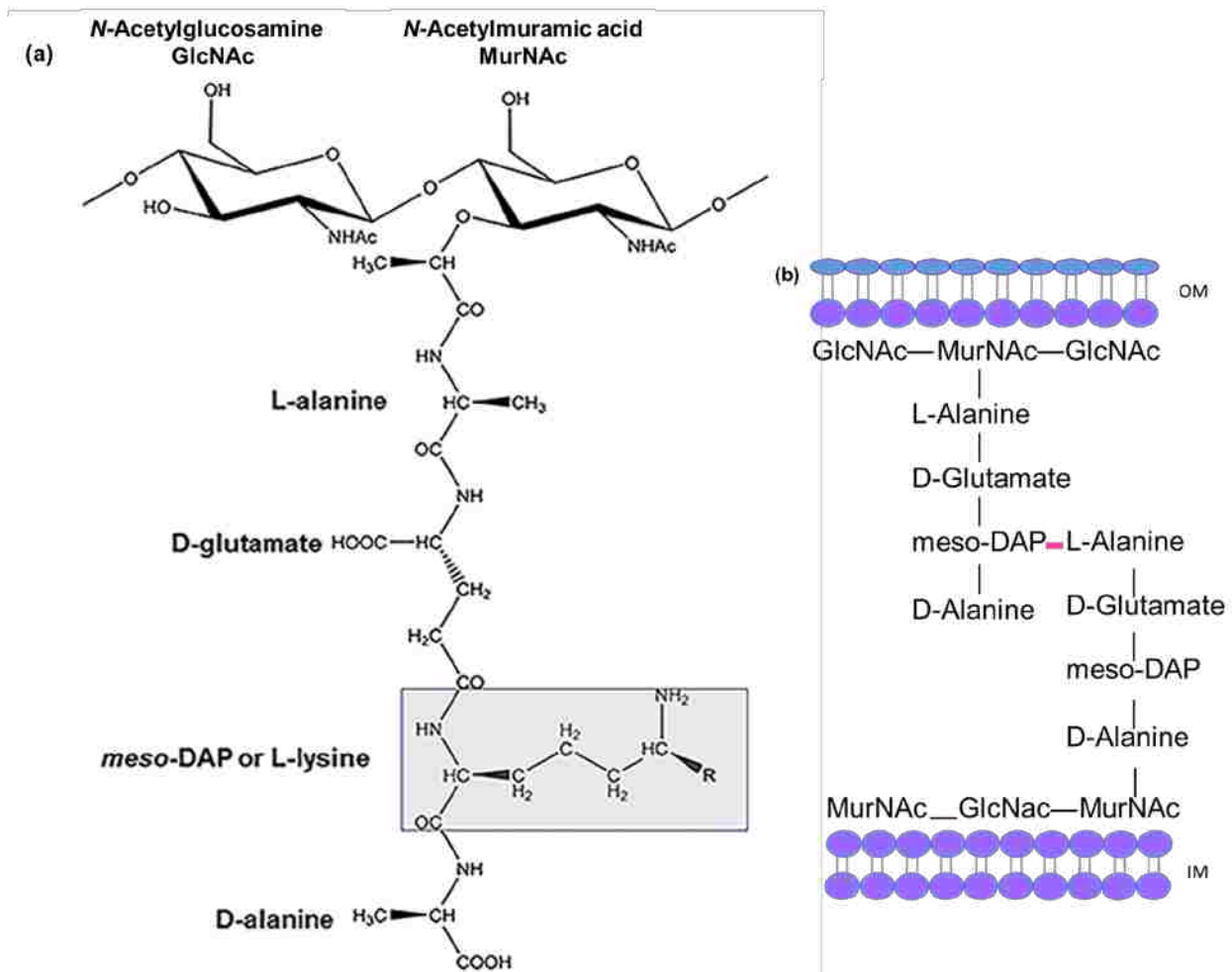


Figure 3-2 NagZ cleaves the bond between GlcNAc and MurNAc, of the peptidoglycan layer [70]. Adapted from: Triassi, A. J., M. S. Wheatley, M. Savka, H. M. Gan, R. C. J. Dobson, and A. O. B. Hudson, *L, I-diaminopimelate aminotransferase (dapI): A putative target for the development of narrow-spectrum antibacterial compounds*. *Frontiers in Microbiology*, 2014. **5**.

GlcNAc, cleaved from the cell, is recycled to make more UDP-GlcNAc, which is a precursor used to make the peptidoglycan. Knocking out *nagZ* inhibits the cell's ability

to cleave GlcNAc from MurNAc leading to a build-up of the disaccharide in the cytoplasm. *nagZ* knockout (Δ *nagZ*) cells have less GlcNAc and UDP-GlcNAc in the cytoplasm compared to the wild type [69]. The level of O-GlcNAcylation is dependent upon the amount of UDP-GlcNAc [71]. Although a Δ *nagZ* system for expression of O-GlcNAc-modified proteins may prevent glycosidase activity, it may be limited in the amount of glycoprotein it can produce because of the lower level of UDP-GlcNAc. In this chapter we will analyze the effects of Δ *nagZ* on the expression of O-GlcNAc-modified proteins.

3. 2 Materials and Methods

3. 2. 1 Materials

Hen egg white lysozyme, glycerol, glucose (Glc), glucosamine (GlcN), *N*-acetyl-D-glucosamine (GlcNAc), *O*-(2-acetamido-2-deoxy-d-glucopyranosylidene)amino-*N*-phenylcarbamate (PUGNAc), sodium lauroyl sarcosinate (sarkosyl), sodium chloride (NaCl), sodium phosphate dibasic, sodium phosphate monobasic, 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), bromophenol blue, Coomassie brilliant blue R250, sodium dodecyl sulfate (SDS), glycine, tris(hydroxymethyl)aminomethane (tris), 2-mercaptoethanol, imidazole, urea, acetic acid, methanol, and kanamycin were purchased from Sigma-Aldrich. Rosetta(DE3)pLysS cells, BugBuster protein extraction reagent, and ampicillin were purchased from EMD Millipore/Novagen. A UNO S1 cation exchange column, Precision Plus Protein Unstained Standards, 4–20% mini-PROTEAN TGX precast gels, immun-blot low fluorescence PVDF membrane, and Laemmli buffer were purchased from Bio-Rad. 2-log DNA ladder, 10-beta cells, T4 DNA ligase, *Nde*I restriction enzyme, and *Xho*I restriction enzyme were purchased from New England

Biolabs. BioReagents EZ-Run Rec Protein Ladder, Coomassie (Bradford) Protein Assay kit, O-GlcNAc-modified BSA, anti-O-GlcNAc MAb CTD 110. 6, and DyLight 550-conjugate Goat anti-Mouse IgG secondary antibody were purchased from Thermo Scientific/Pierce Antibodies. Luria broth (LB) and agar were purchased from BD Biosciences. Wheat germ agglutinin (WGA) agarose and O-GlcNAc-modified BSA were purchased from Vector Labs. Other chemicals and materials used include isopropyl β -D-1-thiogalactopyranoside (IPTG, Affymetrix Anatrace), Ni Sepharose 6 Fast flow (GE Healthcare), chloramphenicol (Calbiochem), QIAprep spin miniprep kit (Qiagen), Costar 96 well UV transparent flat bottom plates, K-12 Δ nagZ knockout strain (JW1093-1) and K-12 parent strain (BW25113) (Keio Collection, *E. coli* Genetic Resources at Yale, CGSC, The Coli Genetic Stock Center) [72], and clone HsCD00331895 of the F-actin binding domain of Abelson tyrosine-kinase 2 (ABL2-FABD) in pET15 vector (DNASU Plasmid Repository, PSI:Biological-Materials Repository) [46]. All water used was supplied from a Millipore Direct-Q 3 ultrapure water system.

3. 2. 2 Knocking out the *nagZ* gene in the Rosetta(DE3) pLysS strain

The Δ nagZ Rosetta(DE3)pLysS strain was made by Dr. William Doerrler and Lisa Boughner, a graduate student.

3. 2. 3 Making the Δ nagZ Rosetta(DE3) pLysS cells competent

Four single Δ nagZ Rosetta(DE3)pLysS colonies were cultured in 5 mL of LB with kanamycin (50 μ g/ml) and chloramphenicol (35 μ g/ml) at 37 °C overnight. These cultures were added to 50 mL of fresh LB with the appropriate antibiotics and cultured for 3 hr. The cells were cooled on ice for 20 min and spun down at 3,000 g for 15 min at

4 °C. The cells were resuspended in 10 mL of 0.1 M CaCl₂ at 4 °C and spun down again. The supernatant was removed and the cells were allowed to dry for 1 min. The cells were reconstituted in 2 mL of 0.1 M CaCl₂ at 4 °C and stored as 100 µl aliquots for transformation.

3. 2. 4 Transformation

pCDF-ABL2 and pET43. 1-OGT were transformed into the *ΔnagZ* Rosetta(DE3)pLysS competent cells using a heat shock method. A mixture of 1 µl of each plasmid and 100 µl of the competent cells was heat shocked at 42 °C for 45 s. The cells were immediately chilled on ice for 2 min and incubated with shaking for 1 hr at 37 °C. *ΔnagZ* Rosetta(DE3)pLysS colonies with pCDF-ABL2 and pET43. 1-OGT were grown on LB agar with 50 µg/mL spectinomycin, 100 µg/mL ampicillin, 50 µg/mL kanamycin, and 35 µg/mL chloramphenicol.

3. 2. 5 Expression conditions

ABL2 and OGT were co-expressed in *ΔnagZ* Rosetta(DE3)pLysS cells. Three colonies of *ΔnagZ* Rosetta(DE3)pLysS with pCDF-ABL2 and pET43. 1-OGT were cultured overnight at 37 °C in 150 mL of LB with the appropriate antibiotics. 1 mL of the overnight culture was used to inoculate fresh LB or minimal media (see Table 3-1 for minimal media expression recipe) with the appropriate antibodies at 37 °C. Once the cultures reached an OD₆₀₀ of 0.6-0.8, they were divided into three 50 mL aliquots and induced with 50 µL of 0.5 M IPTG. Each 50 mL aliquot was incubated at 16 °C, 23 °C, and 37 °C for 16 hr, 12 hr, and 3 hr, respectively. The cells were harvested by centrifugation at 3000 g for 15 min at 4 °C. The supernatant was discarded, and the

pellet was reconstituted with 1 mL of BugBuster. The suspended cells were sonicated using a Sonics Vibra-cell VCX130 at 30% power for 120 s (3 cycles of 30 s on/10 s off). The lysed cells were spun down at 30,000 g for 30 min at 4 °C. The cleared lysate was separated from the pellet and saved for nickel affinity chromatography. The expression procedure was performed for pCDF-ABL and pETDuet-1-codon optimized OGT co-expression in *ΔnagZ* Rosetta(DE3)pLysS.

Table 3-1 The recipe for 1 L of minimal growth media, 1 L of 10X M9 salts, and 100 mL of 1000X trace mineral.

10X M9 salts	100. 0 ml
D-glucose stock (20 g/100 ml) (0. 2 μm filter sterilized)	20. 0 ml
1000X Trace minerals	1. 0 ml
1 M MgCl ₂ (autoclaved)	2. 0 ml
50 mM CaCl ₂ (autoclaved)	2. 0 ml
Vitamin Solution (1 tablet/20 ml) (0. 2 μm filter sterilized)	2 ml

10X M9 salts (autoclaved)

Na ₂ HPO ₄ •7H ₂ O	128 g
KH ₂ PO ₄	30 g
NaCl	5 g
NH ₄ Cl	10g

1000X Trace Minerals (10 mM each) (autoclaved)

FeCl ₃ •6H ₂ O	0. 2703 g
CuSO ₄ •5H ₂ O	0. 2497 g
MnSO ₄ •6H ₂ O	0. 1690 g
ZnSO ₄ •6H ₂ O	0. 2875 g

3. 2. 6 Ni-affinity purification conditions

Ni Sepharose 6 Fast Flow with a binding capacity of 40 mg/mL was used in 0.1 mL aliquots. The nickel media was placed in microspin columns and centrifuged at 300 g to remove the storage buffer. Water was used to rinse the media, followed by 1 mL of wash buffer (50 mM phosphate, 150 mM NaCl, and 20 mM imidazole at pH 8) to equilibrate the media. The cleared lysate was added to the equilibrated media and allowed to flow through by gravity. The media was washed with 2 mL of wash buffer. The protein of interest was eluted using 200 μ L of elution buffer (50 mM phosphate, 150 mM NaCl, and 300 mM imidazole at pH 8).

3. 2. 7 SDS-PAGE and western blot analysis

The cleared lysate and insoluble pellet were analyzed with SDS-PAGE and western blot analysis. The cleared lysate was diluted in a 1:5 ratio with BugBuster or SDS-PAGE buffer. 20 μ L of the dilute cleared lysate was added to 20 μ L of Laemmli sample buffer. The samples were divided in half to run on two different gels. Purified samples were prepared as 30 μ L aliquots in a 1:1 ratio with Laemmli buffer. These samples were also divided for two different gels: one gel was stained with Coomassie brilliant blue R-250 stain, and the other gel was prepared for O-GlcNAc detection by western blot analysis. All samples were boiled for 15 min before loaded onto a 4-20% tris protein gel. The protein gels were run in buffer (25 mM tris, 192 mM glycine, and 0.1% SDS) at 120 V for 1 hr.

Western blot detection of O-GlcNAc was performed by transferring the proteins from the gel to a PVDF-membrane in transfer buffer (25 mM tris, 192 mM glycine, and 10% methanol) using a Bio-Rad transfer electrophoresis system at 120 V for 3 hr. Anti-

O-GlcNAc MAb CTD 110. 6, an O-GlcNAc specific antibody, in combination with DyLight 550-conjugate Goat anti-Mouse IgG, a secondary antibody conjugated to a fluorescent tag, was used to probe the blot for detection of O-GlcNAc. The membranes were first blocked using 1X BSA blocking buffer (25 mM sodium phosphate, 150 mM NaCl, 0. 1% tween and 1% BSA at pH 7. 4) for 1 hr at room temperature. Anti-O-GlcNAc antibody and fresh BSA blocking buffer in a 1:5000 dilution were added to the blot and incubated at room temperature for 1 hr. The blot was rinsed with PBST buffer (25 mM sodium phosphate, 150 mM NaCl, and 0. 1% tween at pH 7. 4) three times for 5 min each. The secondary antibody was added to fresh 1X BSA blocking buffer and incubated at room temperature for an additional hour. The blot was rinsed a final time and imaged using a GE Typhoon 9410 imager. The fluorescent secondary antibody was excited at 532 nm and detected between 565-595 nm using a 580 nm filter.

3. 2. 8 WGA agarose affinity chromatography

The nickel purified samples were exchanged into WGA wash buffer (20 mM HEPES, 150 mM NaCl at pH 8) using Amicon centrifugal filter units with a molecular weight cut-off of 10 kDa to remove the imidazole. The WGA agarose was divided into 100 μ L aliquots and washed with 1 mL of wash buffer. The sample was incubated with the WGA agarose for 20 hr at 4 °C. The mixture was then transferred to a microspin column and centrifuged at 300 g for 1 min. The agarose was washed with 1 mL of wash buffer, incubated with 200 μ L of WGA elution buffer (20 mM HEPES, 150 mM NaCl, 500 mM GlcNAc at pH 8) for 30 min at 4 °C, and centrifuged again to collect the elution.

3. 2. 9 Coomassie protein concentration assay

Nickel-purified and WGA-separated samples were exchanged into WGA wash buffer using Amicon centrifugal filter units with a molecular weight cut-off of 10 kDa. Aliquots (10 μ L) aliquots were saved for Coomassie analysis. BSA standards were prepared in WGA wash buffer at 0, 10, 20, 40, 50 μ g/mL. Standards and samples were prepared in 50 μ L aliquots. Samples were tested and prepared within the linear range based on visual inspection. The flow-through and eluents from the WGA purification were analyzed using a Biotek PowerWave XS microplate spectrophotometer at 595 nm.

3. 2. 10 Supplementing Δ nagZ *E. coli* cultures with sugar

pCDF-ABL2 and pET43. 1-OGT were transformed into Δ nagZ Rosetta(DE3)pLysS competent cells and selected on LB agar with spectinomycin, ampicillin, kanamycin, and chloramphenicol, as previously described. A colony known to express ABL2 and OGT well was cultured in MM in the presence 0. 5% glucose, glucosamine, or *N*-acetylglucosamine in addition to the 0. 4% glucose already in the MM. At an optical density at 600 nm of 0. 6-0. 8, the cultures were induced with 0. 5 mM IPTG and grown for 16 hr at 22-25 $^{\circ}$ C. The cells were harvested using BugBuster protein extraction reagent and lysed with 3 freeze/thaw cycles and ultrasonication. Cleared lysate was prepared by centrifugation at 30,000 g for 30 min at 4 $^{\circ}$ C. Ni Sepharose 6 Fast Flow, a nickel affinity media, was used to purify the His-tagged ABL2 protein. Gravity flow was used to ensure maximum binding. The nickel affinity media was equilibrated with wash buffer. The cleared lysate was loaded onto 200 μ L of media. The media was then washed with 2 mL of buffer and the protein was eluted with 1 mL of 300 mM imidazole elution buffer.

3. 2. 11 Resolubilized O-GlcNAc-modified ABL2

The insoluble portion of the cell lysates were dissolved in 1 mL of 10% sodium lauroyl sarcosinate (sarkosyl), followed by dilution to 1% sarkosyl with water and ultrasonication at 30% for 120 s (3 cycles of 30 s on/10 s off). The resolubilized protein was centrifuged at 30,000 g for 30 min to pellet any remaining insoluble material and purified using nickel affinity chromatography as previously described.

3. 2. 12 OGT activity assay

ΔnagZ Rosetta(DE3)pLysS with pCDF-ABL2 and pET43. 1-OGT and *ΔnagZ* Rosetta(DE3)pLysS with pCDF-ABL2 with pETDuet-1-codon optimized OGT were expressed for 3 hr at 37 °C. The cells were spun down at 3000 g for 10 min at 4 °C. The cells were resuspended in 50 mM tris and ultrasonicated at 30% for 120 s (3 cycles of 30 s on/10 s off). The lysate was separated from the cells by centrifuging at 30,000 g for 10 min. 1 mM UDP-GlcNAc, 2 μM phenylmethylsulfonyl fluoride (PMSF), 75 units of shrimp alkaline phosphatase, and 12.5 mM magnesium chloride were added to the lysates and incubated at 20 °C for 1 hr. The mixtures were analyzed by SDS-PAGE and western blot to determine the relative activity of OGT and codon-optimized OGT. In addition to O-GlcNAc detection, OGT was detected using Anti-O-GlcNAc Transferase (DM-17) produced in rabbits and DyLight 650-conjugate Goat anti-Rabbit IgG.

3. 3 Results

3. 3. 1 Expression of O-GlcNAc-modified ABL2 in *ΔnagZ E. coli*

pET43. 1-OGT and pCDF-ABL2 were co-expressed in *ΔnagZ* Rosetta(DE3)pLysS using LB media at 16 °C. Figure 3-3a shows the SDS-PAGE of the soluble purified ABL2 at a molecular weight of about 14 kDa. Once purified proteins

began to precipitate. The precipitated proteins were also analyzed by SDS-PAGE, Figure 3-3a shows the precipitated protein, referred to as insoluble. ABL2 purified at pH 6 did not produce precipitate or ABL2. Figure 3-3b shows the western blot corresponding to these samples. The O-GlcNAc-modified ABL2 was only in the insoluble sample.

ABL2 was expressed in abundance, but the purification was poor and once purified, it was unstable. The O-GlcNAc-modified ABL2 selectively precipitated with many other proteins. The instability is attributed to purification at high pH (> than 8.5). Even though glycosylated ABL2 was temporarily soluble, the amount of O-GlcNAc-modified protein produced was similar to the wild type expressions. ABL2 was purified at a lower pH in an attempt to stabilize the glycoprotein. However, lowering the pH to 6 hindered the lysis and purification of ABL2.

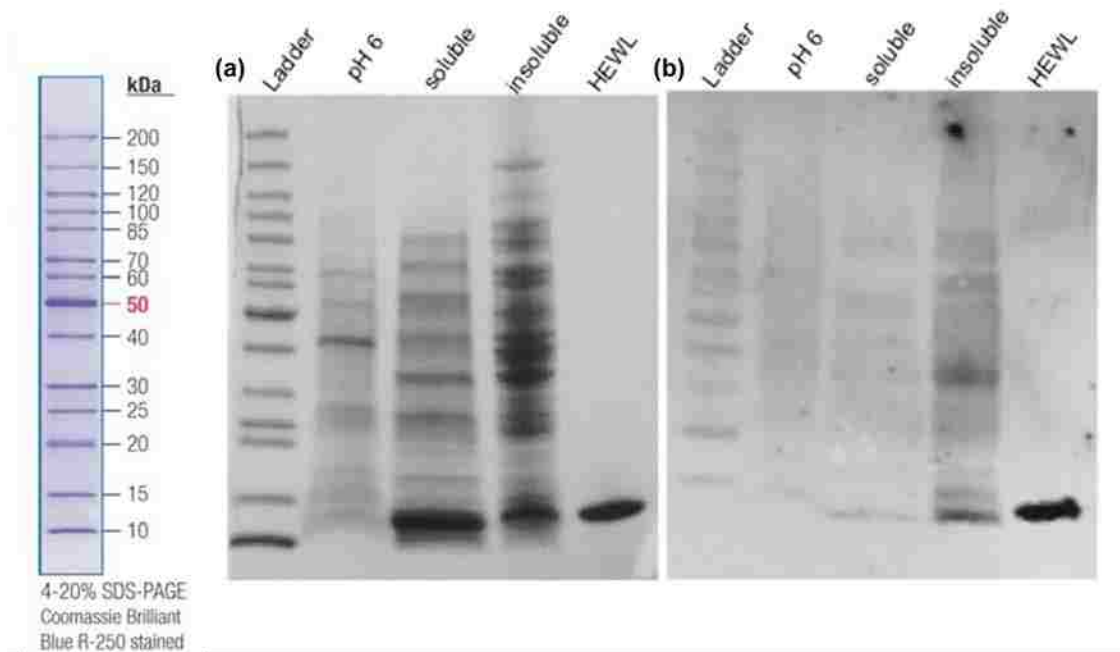


Figure 3-3 O-GlcNAc-modified ABL2 expressed in the Δ nagZ system was unstable and O-GlcNAc-ABL2 selectively precipitated. The soluble and insoluble samples both

contain ABL2 as seen on the protein gel (a); however, the western blot (b) shows that O-GlcNAc-modified ABL2 seems to be exclusively in the pellet.

3. 3. 2 *ΔnagZ* Rosetta(DE3)pLysS sugar trials

pET43. 1-OGT and pCDF-ABL2 were co-expressed in *ΔnagZ* Rosetta(DE3)pLysS using minimal media (MM) at 22 °C. Glucose, glucosamine, *N*-acetylglucosamine or PUGNAc were added to the MM cultures. The soluble lysate and the insoluble pellets were analyzed by SDS-PAGE and western blot. Figure 3-4a shows the SDS-PAGE analysis of cleared lysate, where the asterisk indicates the expression of ABL2. The corresponding western blot (Figure 3-4b) does not show any glycosylation for the samples with additional sugar. However, there is a small amount of glycoprotein in the control sample. Figure 3-4c shows the SDS-PAGE analysis of the insoluble pellet. The western blot (Figure 3-4d) shows an increased amount of glycoprotein in the glucose and glucosamine supplemented expressions, compared to the control. The addition of PUGNAc also increased the amount of glycoprotein expressed. However, ABL2 expressed in the presence of *N*-acetylglucosamine produced glycoprotein comparable to the control.

Glucose, glucosamine, and *N*-acetylglucosamine were added to the culture media in an attempt to compensate for the possible decrease of UDP-GlcNAc caused by knocking out *nagZ*. These monosaccharaides were selected because they participate in the hexosamine biosynthesis pathway. The cultures were grown in MM to control the sugar composition and concentration in the media. The soluble lysate and the insoluble pellets were analyzed by SDS-PAGE and western blot. As previously observed, little or no glycoprotein was present in the lysate of the cultures (Figure 3-4a). MM cultures with extra glucose and glucosamine had the highest amount of

glycosylation observed thus far. In contrast, the addition of *N*-acetylglucosamine had no effect on the expression of glycoprotein. The cells with *N*-acetylglucosamine grew four times faster than the cultures with the other sugars and minimal media alone. The accelerated growth and insubstantial change in glycosylation may be attributed to *N*-acetylglucosamine's ability to up-regulate *nagA* and *nagB*, which promote glycolysis [73]. Glycolysis is a process where *E. coli* uses sugars (mainly glucose) to make ATP, which is required for maintenance and reproduction of the cells [74]. The addition of PUGNAc also increased the amount of glycosylation, suggesting the presence of a second glycosidase with activity toward O-GlcNAc-modified proteins in *E. coli*.

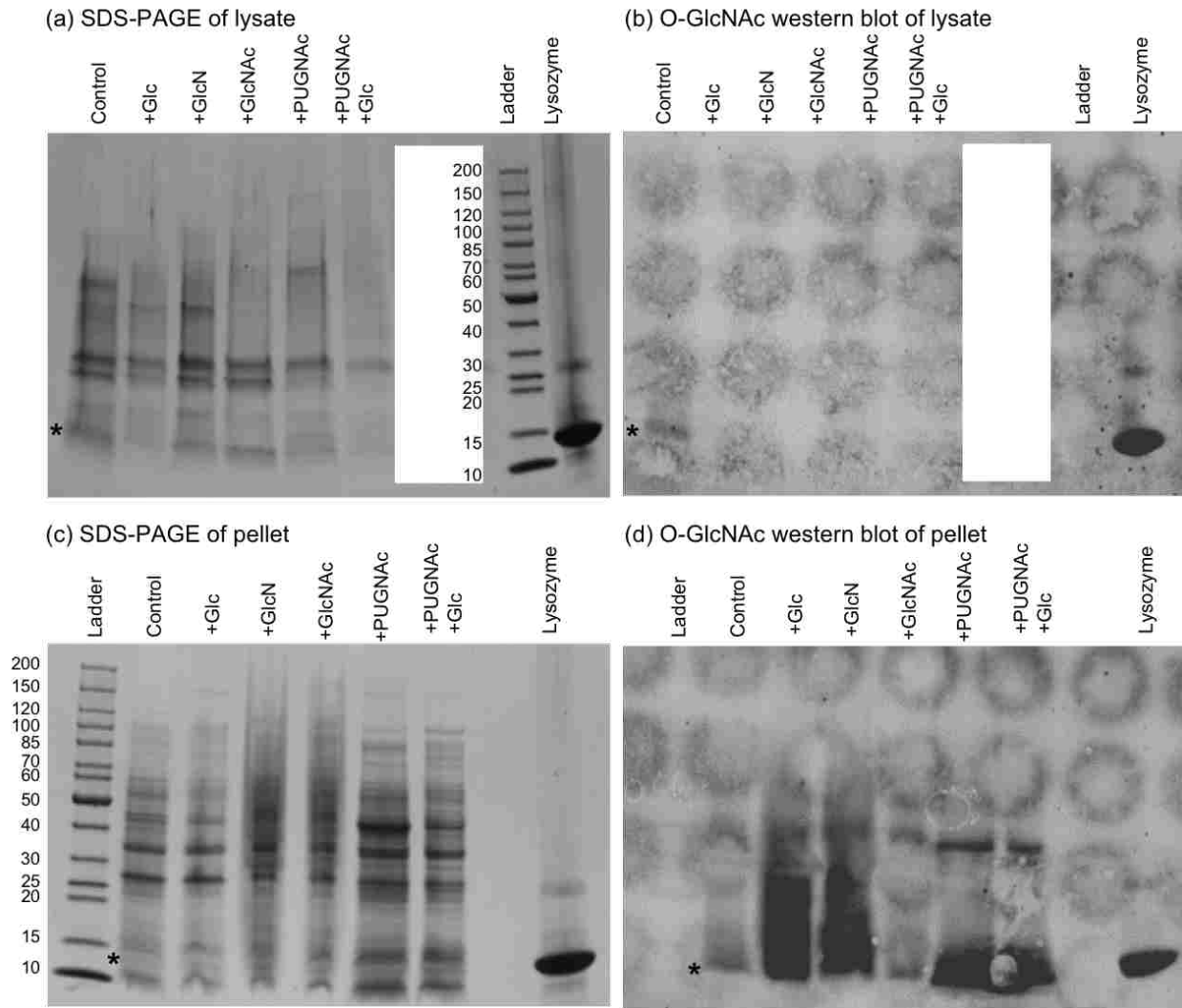


Figure 3-4 Analysis of sugar supplemented expression of O-GlcNAc modified ABL2. The cleared lysate from cultures supplemented with glucose have a small amount of ABL2 (a) that is not glycosylated (b). The pellets corresponding to these samples contain about the same amount of ABL2 (c); however, the samples supplemented with glucose and glucosamine have an abundance of O-GlcNAc-modified ABL2 (d). The pellets containing PUGNAc express more ABL2 (c), but less O-GlcNAc-modified ABL2 compared to the glucose and glucosamine samples (d).

3. 3. 3 Resolubilized ABL2

The insoluble glycoprotein from Figure 3-4d was solubilized using detergent and isolated from insoluble content and cell debris using nickel affinity chromatography.

Figure 3-5a shows the SDS-PAGE analysis of purified resolubilized ABL2. There were very small amounts of ABL2 and some other proteins observed. The western blot (Figure 3-5b) shows that these results were similar to previous results. The glucose, glucosamine, PUGNAc, and PUGNAc+glucose samples had higher levels of glycosylation than the control and N-acetylglucosamine sample.

The overall goal of this project is to express milligram quantities of soluble O-GlcNAc-modified protein for NMR analysis. In an effort to solubilize O-GlcNAc-modified ABL2 without denaturing the protein, sarkosyl was used to solubilize any protein from the inclusion bodies. The results were similar to the previous experiments: glucose and glucosamine appeared to have more glycoprotein than both the control and the *N*-acetylglucosamine sample. The glucose in combination with PUGNAc appeared to have the highest level of purified O-GlcNAc-modified ABL2. The asterisks in Figure 3-5b indicate multiple O-GlcNAc-modified ABL2 bands that migrate differently in the SDS-PAGE gel. More than one O-GlcNAc modification can occur on ABL2, so the bands may correspond to different numbers of O-GlcNAc residues. There are a few proteins that consistently co-purify with ABL2, suggesting that O-GlcNAc-modified ABL2 may be bound to or tightly associated with these proteins.

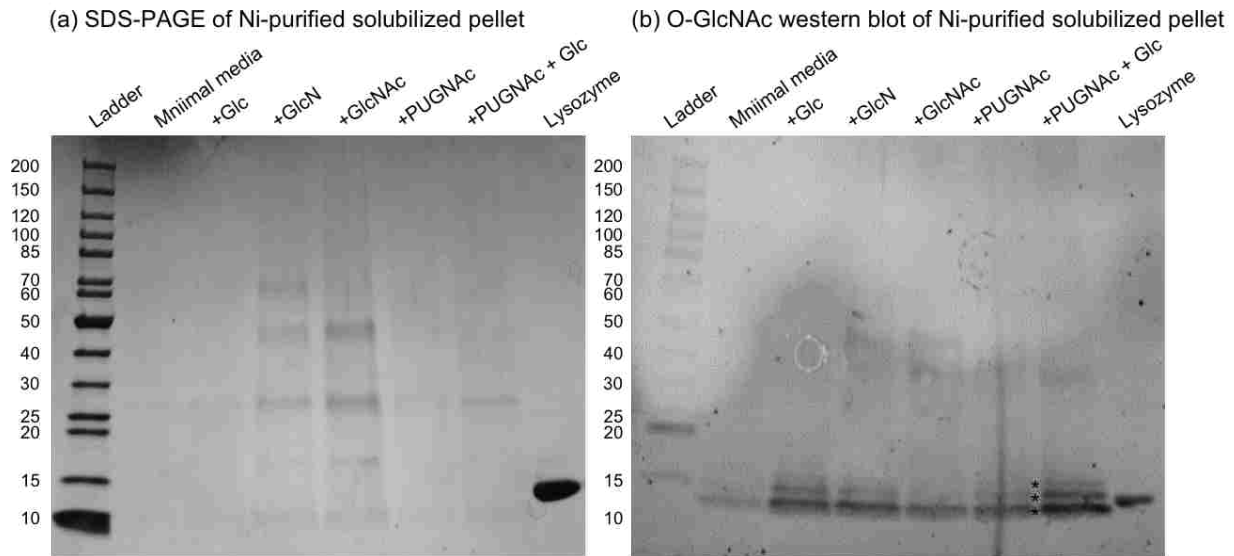


Figure 3-5 Purification of resolubilized O-GlcNAc-modified ABL2. The pellets were resolubilized with 10% sarkosyl, diluted to 1%, and purified with nickel affinity chromatography. The results were very similar to previous data—there is not an abundance of ABL2 on the protein gel, (a) yet it is highly glycosylated (b).

3. 3. 4 Optimized expression of O-GlcNAc-modified ABL2 in $\Delta nagZ$

Previous experiments produced highly glycosylated protein (Figure 3-4d) that was insoluble. This glycoprotein can be isolated from the pellet using a sarkosyl. Although the detergent experiments worked, it is not clear whether sarkosyl will affect the protein's structure. In an effort to optimize the expression, for the production of soluble O-GlcNAc-modified ABL2, the cells were cultured in different conditions. There are many variables that can greatly affect the expression of soluble target protein. The cells were cultured in a tissue culture flask at slower revolutions per minute (rpm). The flat tissue flask doubled the surface area of aeration. These expressions were done at room temperature (22-25 °C) instead of a controlled temperature of 22 °C for 18 hrs. The expressions were induced with IPTG at an OD_{600} of 1.0, instead of 0.6. Most importantly, 10 μ M PUGNAc was added to all of the purification buffers. The samples

were purified and concentrated for analysis. The sample with 0.5% glucose was the only sample that showed significant amounts of pure, soluble O-GlcNAc-modified protein (Figure 3-6). Although the overall expression of ABL2 appears to be low, the ratio of glycoprotein to naked protein was high. In contrast to expression in wild type *E. coli*, O-GlcNAc-modified ABL2 expressed in $\Delta nagZ$ *E. coli* remained glycosylated in the presence of PUGNAc after several days. O-GlcNAc-modified ABL2 expressed in the wild type, in the presence of 10 μ M PUGNAc lost the modification after 2 days. It is not clear whether the decreased amount of ABL2 is real or if the glycosylation disrupts the binding of the coomassie G-250 dye, used for detection of protein. Coomassie is also the reagent used in protein concentration assays, and is not compatible with high levels of GlcNAc.

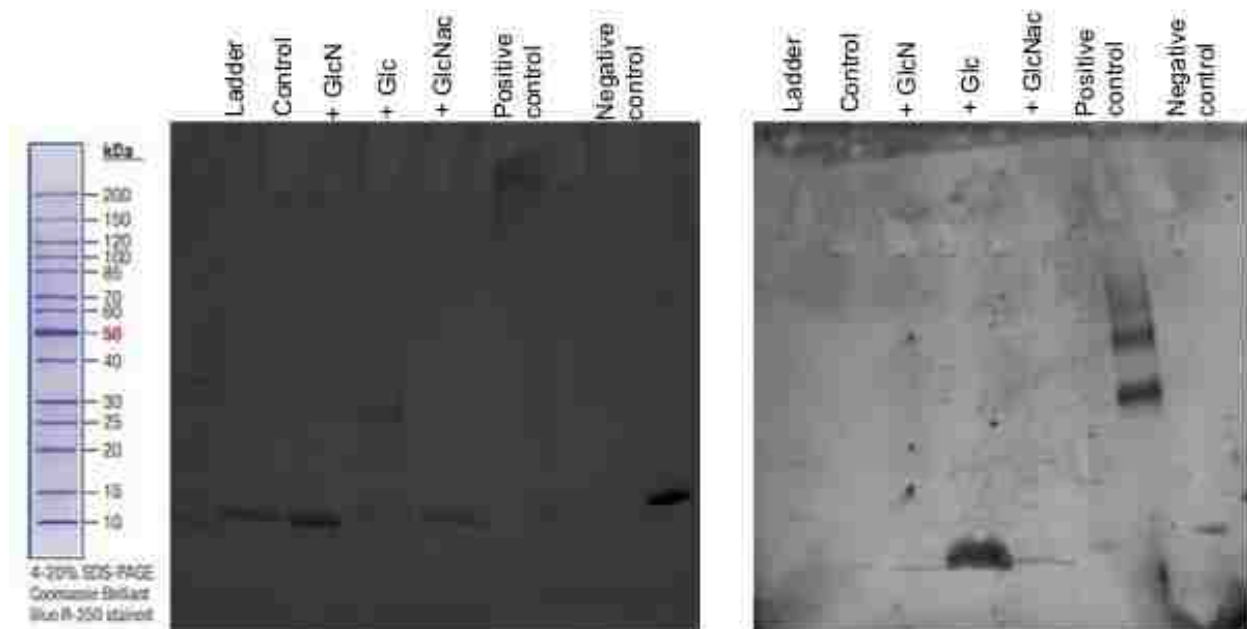


Figure 3-6 Optimized expression of soluble O-GlcNAc-modified protein. A starter culture from a single colony was used to inoculate sugar-supplemented cultures in MM. The expression conditions were optimized to produce soluble O-GlcNAc-modified protein. The soluble lysate was purified and concentrated for SDS-PAGE and western

blot analysis. The glucose sample was the only sample to produce a substantial amount of soluble O-GlcNAc-modified protein.

3. 3. 5 Comparison of OGT activity produced from human OGT (hOGT) and codon-optimized OGT (coOGT)

The human *OGT* sequence contains 52 rare codons. Rare codons are DNA codons that have a low amount of corresponding tRNA, required for protein synthesis, in *E. coli*. The *OGT* sequence was codon optimized, cloned into pETDuet-1 by the protein facility at LSU, and overexpress in *E. coli*. Although codon optimized OGT overexpressed in $\Delta nagZ$ *E. coli*, it did not show any glycosylation, even when glucose and glucosamine were added as supplements to the media. To verify that the coOGT was active, cleared lysate containing coOGT was used to glycosylate ABL2 *in vitro*, and the relative amount of glycosylation was compared to hOGT *in vitro* glycosylation. Figure 3-7 shows the SDS-PAGE and western blot of the *in vitro* glycosylation reaction. There is more coOGT than hOGT and there is more O-GlcNAc-modified protein produced from the hOGT reaction. A comparison of the relative band densities of the western blot bands to the corresponding SDS-PAGE band for the *in vitro* coOGT and hOGT reactions show that the extent of glycosylation is similar for both enzyme. The same cleared lysate was analyzed to assess the relative concentrations of OGT. coOGT was more abundant than hOGT. In theory more coOGT should produce more O-GlcNAc-modified protein, but it does not. These results suggest that coOGT is not as active as hOGT, likely because it does not fold properly during translation.

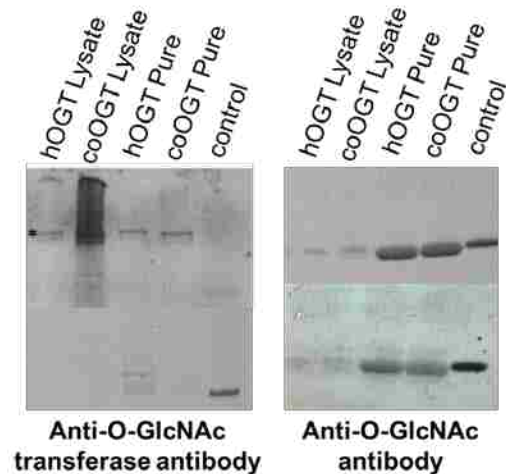


Figure 3-7 Comparison of human OGT and codon optimized OGT. Anti-O-GlcNAc transferase was used to detect the amount of OGT present in the *in vitro* reaction. Anti-O-GlcNAc was used to detect the amount of O-GlcNAc-modified protein produced from the reaction.

3. 4 Conclusion

The Δ nagZ Rosetta(DE3)pLysS *E. coli* co-expression system produced more O-GlcNAc-modified ABL2 compared to wild type Rosetta(DE3)pLysS *E. coli*, but only when the media was supplemented with extra glucose or glucosamine. The insolubility of O-GlcNAc-modified ABL2 was a problem for both *E. coli* strains, providing more evidence that co-translational modification with O-GlcNAc disrupts ABL2 folding. Another possibility is that a second glycosidase is removing O-GlcNAc from soluble ABL2. Evidence for this hypothesis includes increased levels of O-GlcNAc-modified ABL2 when PUGNAc is added to cultures of Δ nagZ Rosetta(DE3)pLysS *E. coli* co-expressing OGT and ABL2. Since O-GlcNAc-modified ABL2 is available in the pellet, resolubilized O-GlcNAc-modified ABL2 may be available in sufficient quantities for NMR analysis. Another possible solution to producing O-GlcNAc-modified protein in sufficient quantities for NMR analysis is to switch the target protein to one that has been modified

with *in vivo* methods. David Vocadlo has had great success with tau and it is biologically relevant. [24]

In vitro glycosylation could be a viable method to produce mg quantities of O-GlcNAc-modified protein. In the past, *in vitro* glycosylation has not been efficient, but NagZ may have contributed to this apparent inefficiency. The Δ nagZ *E. coli* may allow for *in vitro* glycosylation without the need for β -N-acetylglucosaminidase inhibitors. Optimizing the expression of human OGT is also very important. Increasing the amount of hOGT expressed, while maintaining its activity, should increase the amount of protein being glycosylated in the lysate. Partial codon optimization may lead a properly folded, active, and abundant OGT. Rosetta cells contain tRNA for 6 common rare codons. There are 52 rare codons in the hOGT sequence, Rosetta compensates for all but 10. These 10 codons are all CGA. Mutating these codons, without changing the amino acid may lead to more protein expression.

Wheat germ agglutinin is often used to enrich O-GlcNAc-modified proteins, but it is expensive and is not always efficient because of its weak affinity for terminal GlcNAc. In a situation where there is only a very small percent of O-GlcNAc-modified protein, enrichment is important for down-stream analysis like mass spectrometry and NMR. Galactosyltransferase adds a galactose to O-GlcNAc. This enzymatic process has been used to detect O-GlcNAc. While the addition of functionalized galactose provides other ways to analyze the modification. The functionalized galactose has a reactive group that can be used to bind biotin for enrichment with streptavidin [75]. The addition of a mass tag, such as polyethylene glycol (PEG) has been used to visualize multiple-O-GlcNAc-

modified proteins by SDS-PAGE gel and could be used to give valuable information such as how much of the protein is O-GlcNAc-modified and to what degree [33].

While there is still some optimization required for the co-expression system to produce mg quantities of soluble O-GlcNAc-modified protein, there are some promising opportunities that need to be explored. The resolubilized samples are highly glycosylated and the structures of these proteins should be analyzed. If they are properly folded, they may lead to an NMR structure. Even with small amounts of resolubilized O-GlcNAc-modified protein, mass spectrometry can be used to map the O-GlcNAc sites to specific residues. This information alone can provide valuable insight into the role of O-GlcNAc in the structure and function of proteins.

Chapter 4: *Chlamydia trachomatis* CT663: Structure and Ligand Interaction Model

4. 1 Introduction

4. 1. 1 *Chlamydia trachomatis*

Chlamydia is the most common sexually transmitted bacterial disease in the United States [76]. There are 2. 8 million new infections annually [77]. *Chlamydia* can cause reproductive health complications such as pelvic inflammatory disease, infertility and ectopic pregnancy. It can also cause neonatal conjunctivitis and pneumonia [78]. The disease is asymptomatic in 50-80% of patients, so they are often unaware of the infection [79]. The rate of re-infection is alarmingly high [80]. Although *Chlamydia* is both curable and preventable, there is a need for better therapies. In 2000, three chlamydial strains in clinical patients demonstrated multi-drug resistance to doxycycline, azithromycin, and ofloxacin [79].

Chlamydia is an obligate intracellular bacteria, which means it requires a host cell for replication. It infects columnar epithelial cells, which are found in the uterus, the respiratory tract, and the eye. The disease manifests in these areas as a urogenital infection, pneumonia, and conjunctivitis, respectively. The developmental cycle is unique and consists of two alternating cellular forms. The infectious form, elementary bodies (EB), attach to host cells and initiate the developmental cycle, after endocytosis [81]. The EBs develop into the noninfectious replication form called reticulate bodies (RB). These RBs reorganize in an asynchronous manner and then revert back to EBs. The now multiplied EBs exit the host cells, lysing the cells in the process [82]. The newly produced EBs infect new cells and the cycle starts again.

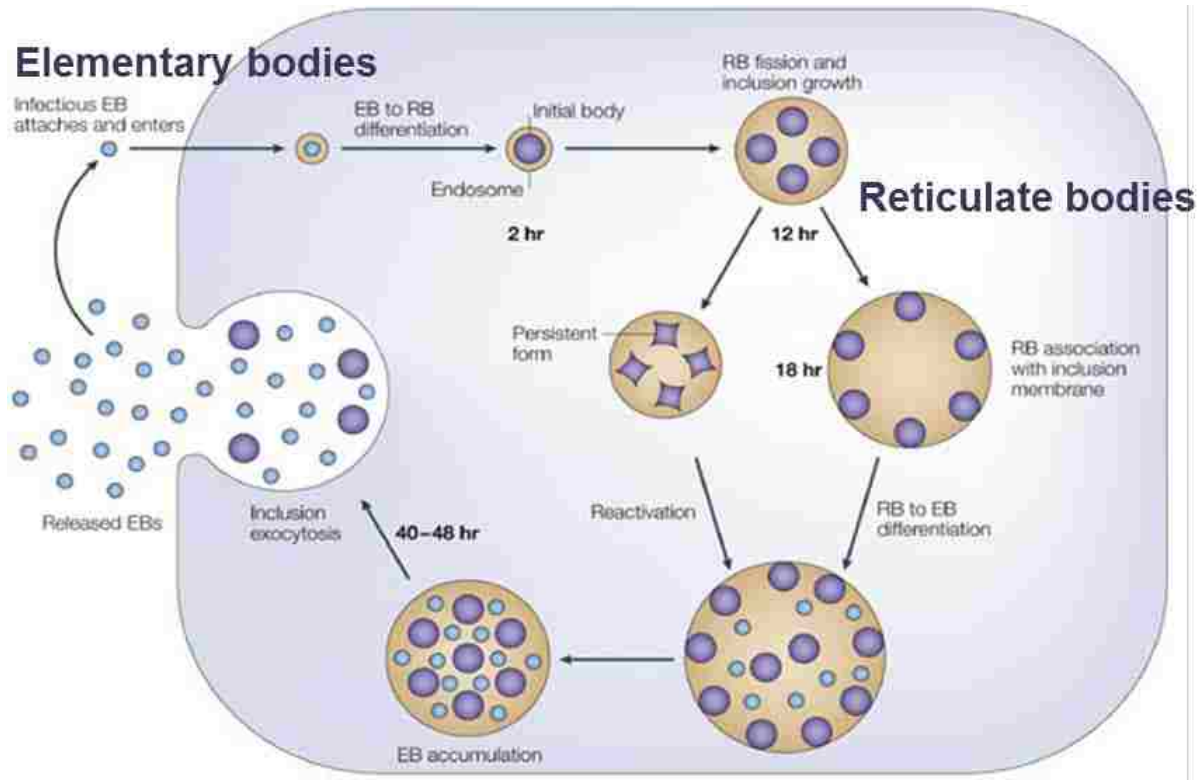


Figure 4-1 Developmental cycle of Chlamydia. Brunham, R. C. ; Rey-Ladino, J. , "Immunology of Chlamydia infection: Implications for a Chlamydia trachomatis vaccine," Nature Reviews Immunology **2005**, 5 (2), 149-161.

4. 1. 2 Type III secretion

The elementary bodies use the virulence associated type III secretion system (T3SS) to infect the host cells. The T3SS is used by gram-negative bacterial pathogens to deliver virulence factors, molecules essential for the induction of infection, into the cytosol of host cells [3]. Virulence factors can be a wide variety of molecules, and are usually on the bacterial surface or secreted from the pathogen. Because the molecules are critical for the initiation and replication of the disease, they are sensible targets for drug development [83]. The T3SS is composed of an injectisome, an ATPase, translocator molecules, effector molecules, and chaperones [84]. The injectisome is the major machinery of the virulence T3SS. It is a multi-protein complex that consists of a

basal body embedded within the bacterial membrane topped by a needle-like structure [85]. It serves as a channel through which proteins reach the host cell. Proteins are secreted through the injectisome by an ATP consuming process, using ATPases located in the bacterial cytosol [84]. Translocator proteins are secreted to form the translocon, which sits at the needle-like end of the injectisome. The translocon physically links the injectisome to the host cell, enabling secretion of virulence factors into the host cell.

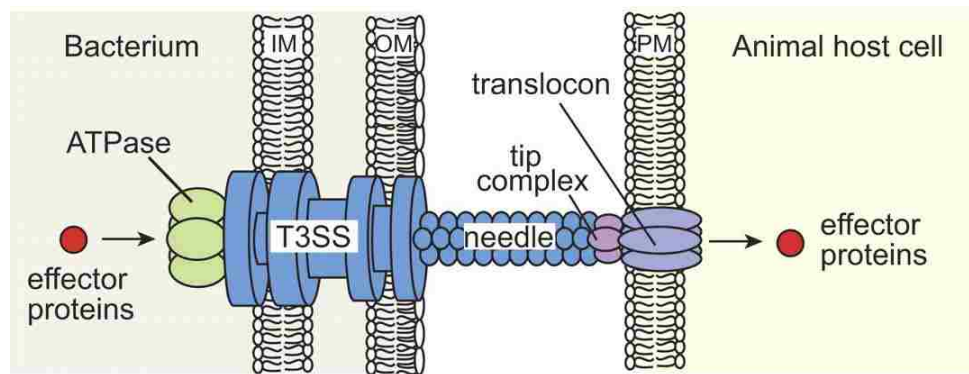


Figure 4-2 Diagram of animal pathogenic bacterial T3SS similar to that of Chlamydia [86]. Büttner D , and He S Y Type III Protein secretion in plant pathogenic bacteria. *Plant Physiology*. **2009**;150:1656-1664.

The secreted virulence factors consist of effector and chaperone proteins. Effectors are eukaryotic-like proteins that modulate the host cell throughout the infectious cycle. Bacterial pathogens secrete effectors that manipulate the actin cytoskeleton of the host cell, which enables remodeling of the membrane in order to promote uptake or adhesion of the bacteria [84]. Effectors and translocators are often bound to chaperones in the bacterial cytosol. Chaperones are proteins that bind to virulence factors in order to keep them in secretion-competent state as they travel from the bacteria and through the injectisome. The chaperones also prevent aggregation and premature interaction of virulence factors, as well as regulate transcription of T3SS

genes [87]. T3SS chaperones can be categorized into three classes based on structure and substrate similarity. Class I chaperones are approximately 130 residues. Class IA chaperones associate with only one effector, while those in class IB associate with multiple effectors. Both sub-classes of class I adopt homo-dimeric structural folds. Class II chaperones are approximately 160 residues and associate with translocator proteins. They are structurally distinct, typically containing three tetratricopeptide repeats (TPR) [84, 88]. Class III chaperones are extra-cellular filament proteins that have extended alpha helical structure.

4. 1. 3 CT663/Scc1-CopN complex

CT663 is a class 1A chaperone found in *Chlamydia trachomatis*. It is a dual function enzyme; it functions as a chaperone and as a transcription factor that regulates the global transition from RB to EB in the Chlamydial developmental cycle. CT663 carries out this function by associating with the β subunit (RpoB) and the primary sigma subunit σ^{66} (RpoD) of Chlamydial RNA polymerase (RNAP). This association inhibits σ^{66} dependent transcription [89]. Analysis of CT663 as a chaperone has shown that it interacts with secretion substrates as well as another class IA T3SS chaperone, Scc1 [84]. This interaction suggests that CT663 and Scc1 form a heterodimeric chaperone complex. The Ct663/Scc1 complex binds Chlamydial outer protein N (CopN), a T3SS effector that is essential for virulence. CopN destroys microtubules in cytoskeleton networks, resulting in cell cycle arrest. CopN is also a T3SS plug protein [90]. Plug proteins typically block the secretion pore, allowing protein export only in the presence of a currently unknown extracellular signal [84]. Although the formation of a CT663,

Scc1, and CopN complex has been confirmed, the order of binding events and arrangement are unknown.

4. 1. 4 Research goal

The goal of this research is to determine the structure and analyze the ligand interactions of CT663. Analysis of CT663's interactions with σ^{66} , β -subunit, Scc1, and CopN will provide information about how CT663 functions as a dual enzyme. Structural analysis of CT663 and its binding partners will provide a rational drug model. Knowledge of how these proteins interact will facilitate the design of a small molecule that inhibits key interactions necessary for infection. An x-ray crystallographic structure of CT663 has been attempted by David WorthyLake, a collaborator at LSU Health Science Center School of Medicine, but crystal formation has not been successful. Nuclear magnetic resonance (NMR) spectroscopy will be used to determine the structure of the chaperones CT663 and Scc1, as well as the chaperone complex CT663/Scc1. Structural determination of any of the chaperones bound to the effector CopN is not feasible, due to size constraints. NMR can still provide information about these interactions, such as which residues participate in binding and how binding affects the structure of the chaperones. Further binding analysis will be done using surface plasmon resonance (SPR) spectroscopy to determine the binding kinetics and arrangement of CT663, Scc1, and CopN.

4. 2 Materials and Methods

4. 2. 1 Materials

Glycerol, glucose, sodium chloride (NaCl), sodium phosphate dibasic, sodium phosphate monobasic, 2-(N-morpholino)ethanesulfonic acid (MES), ammonium acetate, dithiothreitol (DTT), ethylenediaminetetraacetic acid (EDTA), bromophenol blue, Coomassie brilliant blue R250, sodium dodecyl sulfate (SDS), glycine, tris(hydroxymethyl)aminomethane (tris), 2-mercaptoethanol, imidazole, urea, acetic acid, methanol, kanamycin, ampicillin, and chloramphenicol were purchased from Sigma-Aldrich. Rosetta(DE3)pLysS cells, BugBuster protein extraction reagent, 3K and 10K Amicon centrifugal filters, and ampicillin were purchased from EMD Millipore/Novagen. Precision Plus Protein Unstained Standards, 4–20% mini-PROTEAN TGX precast gels, and Laemmli buffer were purchased from Bio-Rad. BioReagents glutathione agarose, HRV 3C protease, EZ-Run Rec protein ladder, and Coomassie (Bradford) protein assay kit were purchased from Thermo Scientific/Pierce Antibodies. Luria broth (LB) and agar were purchased from BD Biosciences. Hiprep 16/60 Sephacryl-S-100 HR and Ni Sepharose 6 Fast flow were purchased from GE Healthcare. Other chemicals and materials used include isopropyl β -D-1-thiogalactopyranoside (IPTG, Affymetrix Anatrace), deuterium oxide (D₂O, Cambridge Isotopes), tween 20 (Fisher), chloramphenicol (calbiochem), QIAprep spin miniprep kit (Qiagen), Costar 96 well UV transparent flat bottom plates were also purchased.

4. 2. 2 Structure predictions

The amino acid sequence of CT663, Scc1, and CopN were submitted to an online protein homology/analogy recognition engine version 2. 0 (Phyre2) for structure prediction.

4. 2. 3 CT663 expression and purification

pET21-CT663 was transformed into Rosetta (DE3)pLysS competent *E. coli* cells and plated on Luria broth (LB) agar with 100 µg/mL ampicillin and 35 µg/mL chloramphenicol. Single colonies were selected for inoculating 5 mL of LB as starter cultures and grown overnight at 37 °C with a shaking speed of 250 revolutions per minute (rpm). The starter culture was added to 0. 5 L of LB media and grown at 37 °C with 250 rpm. Once the culture reached an OD₆₀₀ of 0. 6-0. 8, it was induced with 0. 5 mM IPTG. The culture grew at 37 °C for 3 hr and the cells were harvested at 4,000 g for 10 minutes at 4 °C. The cells were reconstituted in 10 mL of BugBuster followed by ultrasonication at 30% power for 120 s (3 cycles of 30 s on/10 s off) using a Sonics Vibra-cell VCX130. Cleared lysate was prepared by centrifugation at 30,000 g for 30 min at 4 °C. CT663 was purified from the cleared lysate using 1 mL of Ni Sepharose 6 fast flow media by gravity flow. CT663 was washed with 20 mL of 20 mM sodium phosphate, 300 mM NaCl, 20 mM imidazole, pH 8 buffer and eluted in 10 mL of buffer with 300 mM imidazole.

4. 2. 4 CT663 and Scc1 co-expression and co-purification:

E. coli cells containing pACYC184-Scc1-6XHis and pET28-CT663 were sent to us by our collaborator, Dr. Li Shen, at the LSU Health Science Center School of Medicine. The cells were plated on LB agar with 35 µg/mL chloramphenicol and 50

µg/mL kanamycin. 5 mL starter cultures were grown from single colonies. The starter culture was added to 0.5 L of LB media at 37 °C. Once the culture reached an optical density of 0.6-0.8 at 600 nm, they were induced with 0.5 mM IPTG and 0.2% arabinose. The cells were harvested and the Scc1:CT663 complex was co-purified following the same procedures used for CT663 (see section 4.2.1).

4.2.5 CopN expression and purification

E. coli cells containing pGEX098-CopN were sent to us by our collaborator, Dr. Li Shen, at the LSU Health Science Center School of Medicine. The cells were plated on LB agar with 100 µg/mL ampicillin. The cells were grown and harvested following the same procedures used for CT663 (see section 4.2.2). CopN was purified using 1 mL of glutathione agarose media by gravity flow. The media was equilibrated with 10 mL of 50 mM Tris, 150 mM NaCl, pH 8 buffer (TBS). The cleared lysate was added to the media and gently mixed at 4 °C for 30 min. The media was washed with 20 mL of TBS. The GST-fusion tag was cleaved from CopN using 20 units of HRV 3C protease in 2 mL of TBS with gently mixing at 4 °C for 16 hr. The flow-through containing CopN and HRV 3C was collected, and HRV 3C was removed by passing the sample over 0.5 mL of nickel affinity media.

4.2.6 Separation of CT663 and Scc1

Previous research indicates that Scc1 is insoluble when expressed alone. However, when it is co-expressed with CT663, it forms a soluble complex. The CT663/Scc1 complex was expressed and purified as previously described. 6M urea, 8M urea, 2% SDS and 1% Sarkosyl were used to separate the complex. Fast protein liquid chromatography (FPLC) was used to bind the complex to nickel media. The media was

washed with wash buffer containing one of the separating agents. Any remaining bound protein was eluted with 20 mM sodium phosphate, 300 mM NaCl, and 300 mM imidazole at pH 8 buffer. The UV absorbance was monitored for the release of protein during the separation wash. SDS-PAGE was used to analyze the proteins that eluted after separation.

4. 2. 7 Gel electrophoresis analysis

Purified CT663, Scc1/CT663 complex, and CopN were run in 4-20% tris-HCl gels under denatured and native conditions. The denatured samples were prepared in a 1:1 ratio with Laemmli buffer containing 5% v/v 2-mercaptoethanol and heated at 95°C for 15 minutes. The samples were run at 125 V for 60 min in 25 mM Tris, 192 mM glycine, 1.0% SDS, pH 8.3 running buffer. Prior to running the native samples the 4-20% tris-HCl gel was run in 25 mM tris, 192 mM glycine, pH 8.3 native running buffer for 60 min to remove the SDS from the gel. Samples prepared for the native gel were buffer exchanged into 20 mM tris, 100 mM NaCl, pH 8 buffer and concentrated to 25 µM in 20 µL using Amicon centrifugal filter units with a molecular weight cut-off of 10 kDa. Each sample was run individually along with 1:1 mixtures of CopN:CT663 and CopN:Scc1/CT663. These mixtures were incubated at 25°C for 1.5 hr. 10% glycerol was added to the native sample in a 1:2 ratio and the native gel was run for 75 min at 125 V.

4. 2. 8 Gel filtration analysis

CT663 and the Scc1/CT663 complex were analyzed with quantitative gel filtration analysis using either a HiPrep Sephacryl S-100 HR with a 1 mL sample injection volume or an ENrich 70 SEC column with a 50 µL sample injection volume. Two stock solutions

of CT663 and two stock solutions of the Scc1/CT663 complex were prepared for analysis using Amicon centrifugal filter units with a molecular weight cut-off of 10 kDa. The concentrations and buffer conditions of each sample are listed in Table 4-1. Protein concentrations were determined using the Bradford colorimetric assay with Coomassie Plus reagent and a Biotek PowerWave XS microplate spectrophotometer. Standard curves were generated using molecular weight standards (2 mg/mL BSA, 1 mg/mL β -lactoglobulin, 1 mg/mL cytochrome C, and 50 mM cytidine) using each running buffer: (1) 20 mM tris, 150 mM NaCl, pH 8.0, (2) 20 mM tris, pH 7.4, (3) 50 mM potassium phosphate monobasic/sodium phosphate dibasic (K-Na-phosphate), pH 7.3, and (4) 20 mM sodium acetate, pH 4.5. Regression analysis was used to determine the linear relationship between the elution volume and the logarithm of the molecular weight for each column/buffer combination. The molecular weights of the samples were determined using the linear function to determine the oligomeric state of the samples in each buffer.

4.2.9 SPR Analysis

Binding experiments were conducted using a kinetic titration method called single-cycle kinetics. This method consists of serial injections of analyte at increasing concentrations over the surface of a ligand-functionalized sensor chip surface without regeneration between sample injections [91]. SPR studies were performed at 25 °C on a nitrilotriacetic acid (NTA) chip (Xantec) on a BIAcore T100 system (GE Healthcare). Both cells of the uncharged NTA chip were conditioned with 0.5 M EDTA at a flow rate of 10 μ L/min with 180 s of contact time, followed by 0.1 M NaOH at a flow rate of 10 μ L/min with 180 s of contact time. Two capture cycles were performed on the second

cell of the sensor chip to create a capture baseline. The capture cycle included three steps: charging the surface with Ni^{2+} , capturing the ligand, a series of washes, and surface regeneration. The second cell of the chip was charged with 5 mM NiCl_2 at a flow rate of 10 $\mu\text{L}/\text{min}$ with a contact time of 120 s. Charging the surface with nickel allows the binding of histidine tagged ligands, such as the His-Scc1/Ct663 complex. The His-Scc1/CT663 complex was prepared in running buffer (50 mM potassium phosphate, 150 mM NaCl, 5 mM EDTA, 0.005 % Tween at pH 7.3) and immobilized onto the surface at a flow rate of 10 $\mu\text{L}/\text{min}$ with a contact time of 150 s. The chip was regenerated with 0.5 M NaOH at a flow rate of 30 $\mu\text{L}/\text{min}$ for 2 min, followed by 350 mM EDTA at 30 $\mu\text{L}/\text{min}$ for 2 min. The capture cycle was repeated. The analyte binding step was a single cycle kinetic method with 5 increasing concentrations. This cycle consisted of a charging step, a ligand capture step, followed by 5 analyte binding and dissociation steps. CopN, the analyte, was flown at a rate of 10 $\mu\text{L}/\text{min}$ with 150 s of contact time across the second cell with a dissociation time of 240 s. The parameters of the other steps were the same. CopN was passed across the surface of the first flow cell as a control or blank.

4. 2. 10 NMR buffer optimization for CT663

CT663 was expressed in minimal media as previously described (Chapter 3). Isotopic labeling was incorporated by adding $^{15}\text{NH}_4\text{Cl}$ to the MM. The proteins were purified with Ni-affinity chromatography as previously described. Purified CT663 was purified with size exclusion chromatography using a Hiprep 16/60 Sephacryl-S-100 HR column in 20 mM PBS at pH 7. ^{15}N -CT663 was concentrated to 500 μM and exchanged into 20 mM sodium phosphate, 20 mM tris, and 20 mM MES buffers using Amicon

centrifugal filters units with a molecular weight cut-off of 3 kDa. The samples were concentrated to a final volume of 250 μ L and transferred to 5 mm Shigemi tubes with 10% D₂O. An ¹H-¹⁵N HSQC spectrum of each sample was collected on a Varian 700 MHz NMR.

4. 2. 11 ¹⁵N-CT663/Scs1 NMR analysis

¹⁵N-CT663 was expressed, purified, exchanged into 20 mM sodium acetate at pH 6 buffer, and concentrated to 500 μ M as previously described. pACYC-184-Scs1 was transformed into BL21 cells and selected on LB agar with 35 μ g/mL chloramphenicol. A single colony was added to 10 mL of LB and grown overnight. This starter culture was added to a 1L of LB and cultured at 37 °C until it reached an OD₆₀₀ of 0.6-0.8. The cells were induced with 0.5 mM IPTG and cultured at 37 °C for 3 hr. The cells were harvested by centrifugation at 4000 g for 10 minutes at 4 °C and reconstituted in 10 mL of 8 M urea in BugBuster. Sonication at 30% power pulsed for 120 s (3 cycles of 30 s on/10 s off) was used to lyse the cells, and cleared lysate was prepared by centrifugation at 30,000 g for 30 min at 4 °C. After filtering the cleared lysate with a 0.22 μ m syringe filter. The denatured Scs1 was bound to Sepharose 6 fast flow media using the FLPC. The media was washed with 10 mL of 20 mM sodium phosphate, 300 mM NaCl, 20 mM imidazole, and 6 M urea at pH 8 buffer. A 30 mL linear gradient at 0.1 mL/min from 6 M urea to 0 M urea was used to refold the protein on the column. Refolded Scs1 was eluted with 20 mM sodium phosphate, 300 mM NaCl, and 300 mM imidazole at pH 8 buffer. Scs1 was exchanged into 20 mM sodium acetate at pH 6 buffer and added to the ¹⁵N-CT663 sample. The ¹H-¹⁵N HSQC spectra of 250 μ L of 500

μM ^{15}N -CT663 before and after the addition of Scc1 were collected on a Varian 700 MHz NMR.

4. 3 Results

4. 3. 1 Structural predictions of the *Chlamydia trachomatis* T3SS effectors-chaperone complex CT663:SCC1:CopN

The T3SS is not exclusive to *Chlamydia trachomatis* and is used by other pathogens as well. The T3SS apparatus itself has been shown to display notable similarities among different bacterial species [92]. *E. coli*, *Yersinia*, *Salmonella*, and *Shigella* are a few examples of pathogens that use the T3SS. The macromolecules of the T3SS are well characterized for these pathogens. High-resolution crystal structures of all classes of T3SS chaperones, crystal structures of the proteins involved in basal body formation, and NMR and FTIR analyses of needle proteins have advanced the structural knowledge of the T3SS [92]. Using the structures of chaperones and effectors from other bacterial species as a basis has allowed for a structural prediction of CT663/Scc1 in complex with CopN. Protein homology/analogy recognition engine version 2. 0 (Phyre2) is a structure prediction server that uses homology/analogy of known structures to predict the structure of other proteins. Structural homology is based on two principles: “first, the structure of a protein is determined by its amino acid sequence; second, proteins with similar sequences are practically identical in structure, and even proteins that have distantly related sequences still fold into similar structures” [93]. Protein analogy refers to proteins with similar functions despite their different sequences. The structure of CT663 was predicted with 97% sequence coverage from 4 template structures of T3SS chaperones from other bacteria [94-97]. The structure of

Scs1 was predicted based on three proteins with 92% coverage [94, 97, 98]. The structural orientation of CopN in complex with CT663/Scs1 was predicted based on a *Yersinia pestis* effector chaperone complex (PDB 1XKP) [99]. The *Yersinia pestis* complex as well as a protein from *Shigella flexneri* was used to predict the structure of CopN with 92% coverage [99, 100]. The structure prediction was done to better understand how CT663 forms a complex with SCC1 and CopN.

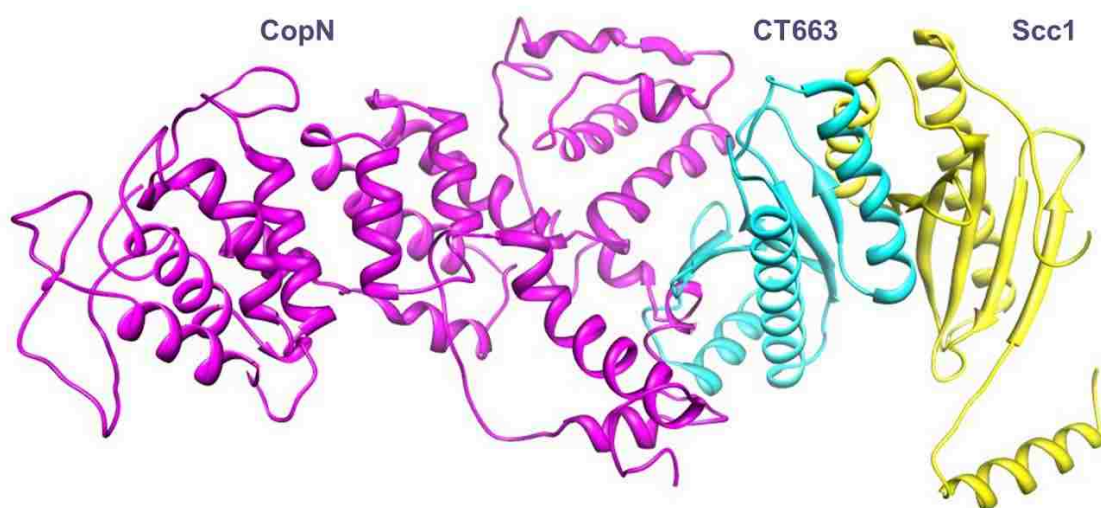


Figure 4-3 The CopN:CT663:Scs1 structure is modeled using *Yersinia pestis* virulence factor YopN in complex with the heterodimeric chaperone SycN-YscB by Phyre2.

4. 3. 2 Expression and Purification

The expression and purification conditions were optimized by Dr. Li Shen, our collaborator at the LSU Health Science Center School of Medicine. CT663, Scs1/Ct663, and CopN express with yields of ~5 mg, ~66 mg, and ~2 mg of protein per liter of culture, respectively. The purification of these proteins is relatively clean. Figure 4-4 shows the SDS-PAGE and native gel analysis of the purified proteins. The SDS-PAGE gels separate denatured proteins based on their size and charge. While native gels separate the proteins in the same manner, the proteins are in their native state and SDS

is not available to bind and add negative charge. Without SDS to denature and add uniform negative charge, it is hard to determine the exact mass of the native proteins. However, it appears that CT663 and CT663/Scc1 form dimers under native conditions. CopN was cloned and expressed with a glutathione S-transferase (GST) fusion tag for purification. This GST tag forms dimers, so the CopN dimer was expected and not informative. While Figure 4-4a shows CopN cleaved from GST, the CopN used for the native gel is not cleaved. A quick and dirty binding assay, mixing proteins in 1:1 ratio, showed no binding of the chaperones with CopN. This result is inconsistent with gel filtration analysis and is likely an artifact from interactions with the gel and/or interference from the GST fusion tag.

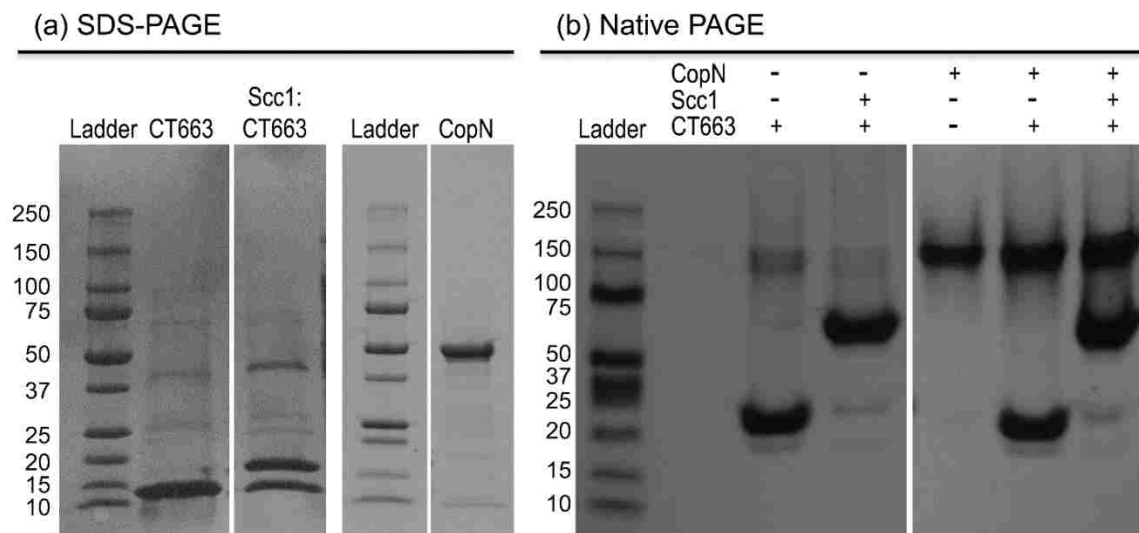


Figure 4-4 SDS-PAGE and native PAGE analysis of CT663, Scc1/CT663, and CopN. The protein molecular weight ladder is shown in lane 1 in the (a) SDS-PAGE and (b) native PAGE gel images. The molecular weight of CT663, Scc1/CT663 and CopN are 14.7 kDa, 30.9 kDa, and 45.2 kDa respectively.

4.3.3 Separation of CT663/Scc1 complex

The CT663/Scc1 complex is expressed within *E. coli* as a complex and also co-purifies. The fact that CT663 and Scc1 purify together suggests a very low dissociation

constant. Since Scc1 is difficult to produce when expressed in *E. coli* alone, we tried to separate CT663 and Scc1 and recover each protein in a soluble, folded state. The complex was bound to nickel affinity media through the His-tag on Scc1. Urea was used to try to separate and elute CT663 from Scc1. 6 and 8 M urea were not sufficient to destabilize the interaction between the CT663 and Scc1. Urea is a very common compound used to denature proteins. The exact mechanism of how urea denatures proteins is not known, but studies have shown that urea disrupts hydrophobic interactions. Urea also interacts with polar residues and the peptide backbone to stabilize the denatured form of the protein [101]. The fact that 8 M urea is not enough to separate CT663 from Scc1 suggests that interactions other than hydrophobic, like electrostatic and hydrogen bonds, contribute to the stability of the complex.

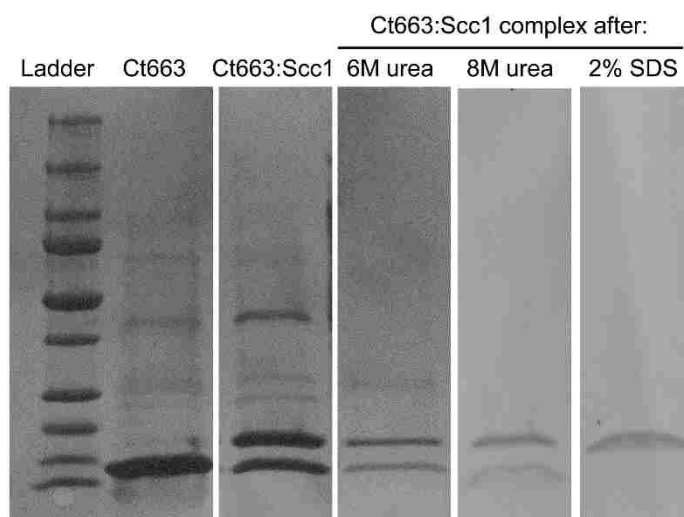


Figure 4-5 CT663-Scc1 interaction. Purified CT663 and CT663:Scc1 complex are shown in lanes 2 and 3, respectively. The purified CT663-Scc1 complex was immobilized on 1 mL of Ni-affinity media via a 6XHis tag on Scc1 and washed with 20 mL of wash buffer with 6 M urea (lane 4), 8 M urea (lane 5), and 2% SDS (lane 6). Even under strong denaturing conditions of 8 M urea, some CT663 remained bound to 6XHis-tagged Scc1. SDS (2%) was sufficient to disrupt the CT663:Scc1 interaction, resulting in pure Scc1.

4. 3. 4 Buffer dependent gel filtration analysis

Gel filtration was done in an effort to characterize the oligomeric state of the chaperones CT663 and CT663/Scs1. The goal of this study was to find conditions where the chaperones were in their monomeric state. CT663 and Scs1 formed a heterodimer in 20 mM sodium acetate at pH 4. 5, while CT663 formed high order oligomers regardless of buffer conditions.

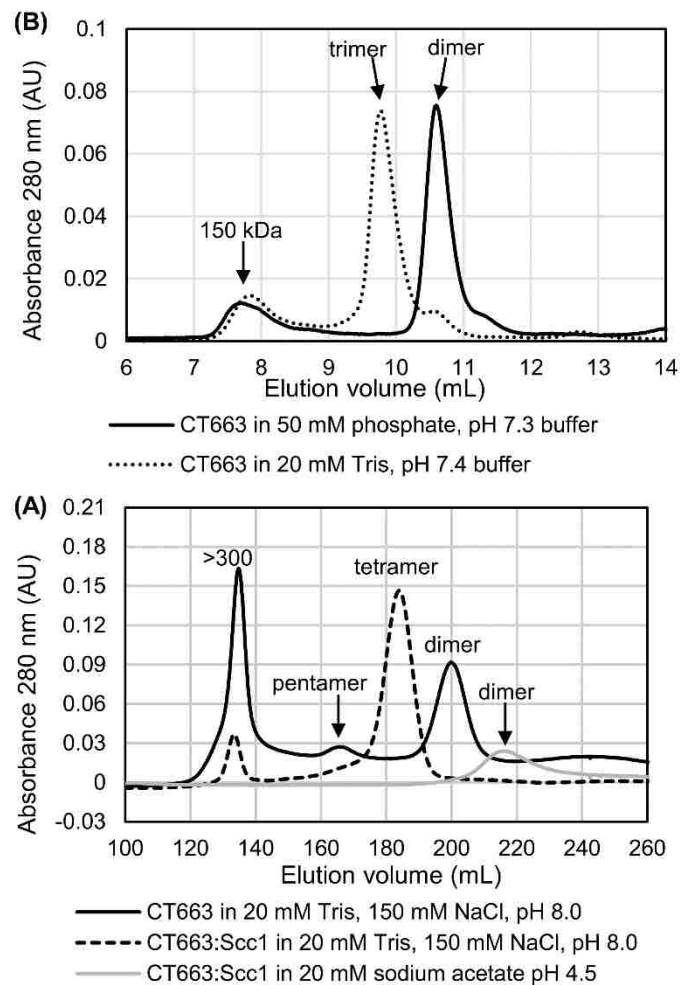
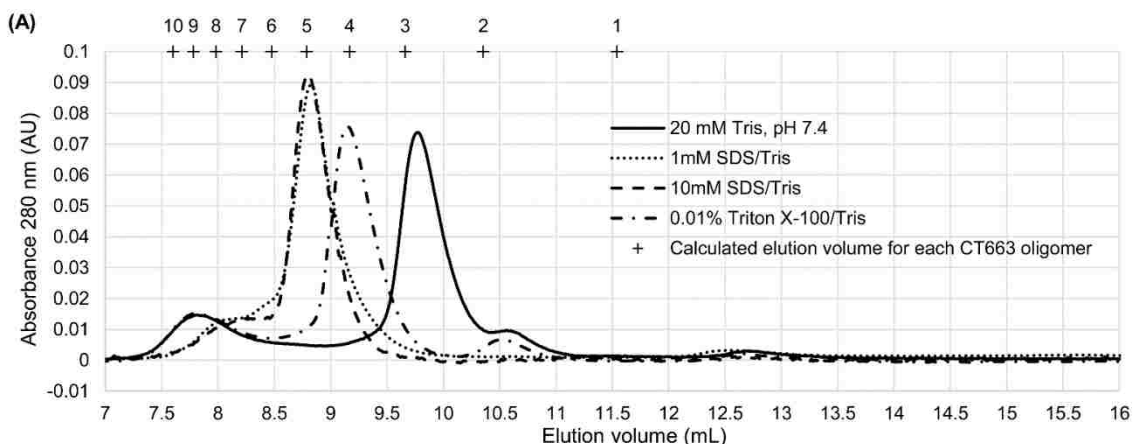


Figure 4-6 Gel filtration analysis of CT663 and CT663:Scs1 at various buffer conditions. Two gel filtration column were used to analyze the samples: (A) a HiPrep Sephacryl S-100 HR column using a 1 mL sample volume and (B) an Enrich SEC 70 column using 50 μ L sample volumes. Molecular weight standards were used to determine the linear relationship between the elution volume and the logarithm of the molecular weight for each column/buffer combination. The molecular weight of the samples were determined

using the best-fit linear function based on regression analysis to determine the oligomeric state of the samples in each buffer. The sample concentrations were 81-152 μM except for the CT663:Scc1 complex in sodium acetate, which was 27 μM due to its instability in this buffer. Details about the sample conditions are described in Table 4-1. (A) In Tris/saline buffer at pH 8.0, CT663 (solid black line) predominantly exists as higher order oligomers (>300 kDa) and as a homodimer while the CT663:Scc1 complex (dashed black line) predominantly forms a stable tetramer. In acetate buffer at pH 4.5, CT663 and the CT663:Scc1 complex are unstable. CT663 readily precipitates, but the complex stability was better and showed that soluble CT663:Scc1 in acetate buffer (gray solid line) exists as a heterodimer. (B) In phosphate buffer at pH 7.3 (solid black line), CT663 forms a homodimer and in Tris buffer at pH 7.4 (dashed black line), CT663 forms homotrimers.

4.3.5 Detergent assisted Gel filtration

Since different buffers and salt did not separate CT663 oligomers, detergents were added in an attempt to stabilize the monomeric form of CT663. When the detergents were added to Tris buffer, it increased the aggregation of the protein (Figure 4-6a). The detergents had little effect on CT663 in phosphate buffer, as shown in Figure 4-6b.



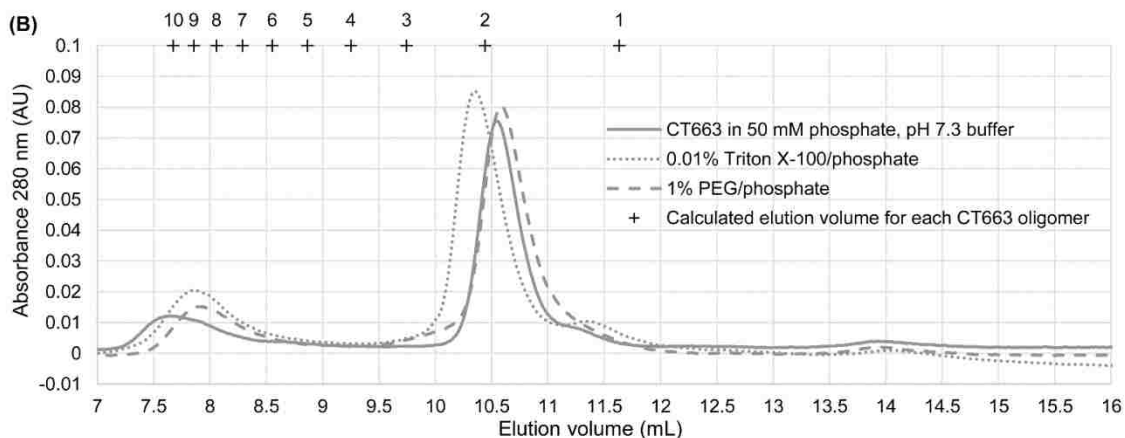


Figure 4-7 Detergent effects on CT663 oligomerization. (A) CT663 samples were run in 20 mM Tris, pH 7.4 buffer (solid black line) with 1 mM SDS (dotted black line), 10 mM SDS (dashed black line), and 0.01% Triton X-100 (dash-dotted black line). The predicted elution volumes for CT663 oligomers (+) were calculated using the calibration curve (Table 4.1). The number of monomer units for each oligomer is shown above the marker. (B) CT663 samples were run in 50 mM potassium-sodium phosphate, pH 7.3 buffer (solid gray line) with 0.01% Triton X-100 (dotted gray line) and 1% PEG (dashed gray line). The predicted elution volumes for CT663 oligomers (+) are shown as in (A). Each sample was 50 μ L of 81 μ M CT663 in 50 mM K-Na phosphate, pH 7.3 buffer, except for the sample run in 10 mM SDS/Tris buffer. This sample was diluted by a factor of 4 with the running buffer to avoid SDS-phosphate precipitation. The absorbance values for this sample were multiplied by 4 for direct comparison to the other samples.

Table 4-1 This table summarized the resulting oligomeric state of CT663 and Scc1/CT663 complex in different buffer conditions.

Sample (μ M)	Sample Buffer	Running Buffer	Column/Sample volume	Results
CT663 (152)	20 mM sodium phosphate, 300 mM imidazole, 300 mM NaCl, pH 8.0	20 mM Tris, 150 mM NaCl, pH 8.0	HiPrep Sephacryl S-100 HR/ 1 mL	hexamer, dimer
Scc1:CT663 (133)	20 mM sodium phosphate, 300 mM imidazole, 300 mM NaCl, pH 8.0	20 mM Tris, 150 mM NaCl, pH 8.0	HiPrep Sephacryl S-100 HR/ 1 mL	hetero-tetramer
Scc1:CT663 (27)	20 mM sodium acetate, pH 4.5	20 mM sodium acetate, pH 4.5	HiPrep Sephacryl S-100 HR/ 1 mL	hetero-dimer
CT663 (81)	50 mM K-Na-phosphate, pH 7.3	20 mM Tris, pH 7.4	Enrich 70 SEC/ 50 μ L	trimer
CT663 (81)	50 mM K-Na-phosphate, pH 7.3	50 mM K-Na-phosphate, pH 7.3	Enrich 70 SEC/ 50 μ L	dimer
CT663 (81)	50 mM K-Na-phosphate, pH 7.3	1 mM SDS, 20 mM Tris, pH 7.4	Enrich 70 SEC/ 50 μ L	apparent pentamer
CT663 (20)	50 mM K-Na-phosphate, pH 7.3	10 mM SDS, 20 mM Tris, pH 7.4	Enrich 70 SEC/ 50 μ L	apparent pentamer

CT663 (81)	50 mM K-Na-phosphate, pH 7.3	0.01% Triton X-100, 20 mM Tris, pH 7.4	Enrich 70 SEC/ 50 uL	apparent tetramer
CT663 (81)	50 mM K-Na-phosphate, pH 7.3	0.01% Triton X-100, 50 mM K-Na-phosphate, pH 7.3	Enrich 70 SEC/ 50 uL	apparent dimer
CT663 (81)	50 mM K-Na-phosphate, pH 7.3	1% PEG, 50 mM K-Na-phosphate, pH 7.3	Enrich 70 SEC/ 50 uL	apparent dimer

4.3.6 SPR analysis of Ct663/Sccl binding to CopN

SPR was carried out at various concentrations of ligand and analyte to optimize the experiment. The Sccl/CT663 ligand concentration was 4 nM. The serial analyte concentrations used were 2.5, 5, 10, 20, and 40 nM of CopN. Raw data was collected in the form of a sensogram which plots response units (RU) vs time. The spectra corresponding to the blanks and the sample are automatically aligned for subtraction. The data is first manually processed by selecting which blanks to subtract from the sample run. Sharp peaks, artifacts that occur during injection, are removed from the resulting spectrum. The BIAevaluation software was used to fit the data to a curve based on a 1:1 binding motif. The rate of association and dissociation of the Sccl/CT663 complex and CopN is $1.01 \times 10^5 \pm 0.02 \times 10^5$ 1/Ms and $3.04 \times 10^{-4} \pm 0.16 \times 10^{-4}$ 1/s respectively. The equilibrium dissociation constant is 3.02 ± 0.17 nM, which suggests strong binding between CopN and the Sccl/CT663 complex.

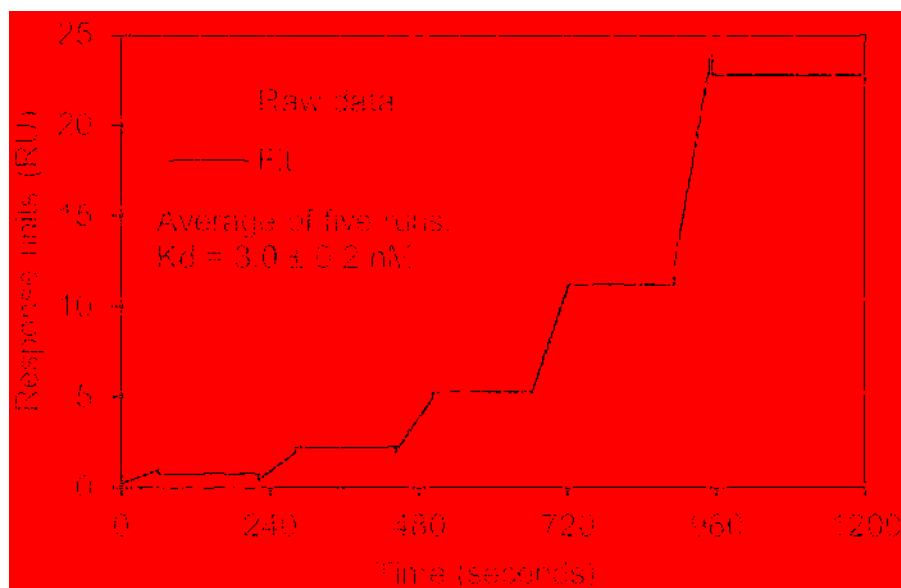


Figure 4-8 The SPR sensogram corresponding to the binding event between the Scc1/CT663 complex and CopN. The calculated $K_d=3.02 \pm 0.17$ nM.

4. 3. 7 The magical CT663 sample from never never land

CT663 prepared by our collaborators was sent to us for HSQC analysis. The sample was said to have contained 500 μ M CT663 without any purification tags and 20 mM sodium acetate at pH 4. 5. This sample is a good candidate for structural analysis because greater than 90% of the peaks are present. The peaks are well dispersed and of similar intensity, suggesting that the protein is fold and possibly in monomeric form. This spectrum has not been reproducible because the exact conditions were not known. There was a lack of communication from all parties involved in producing the sample and obtaining the spectrum. Recent meeting have revealed the appropriate conditions and attempts to reproduce this spectrum is underway.

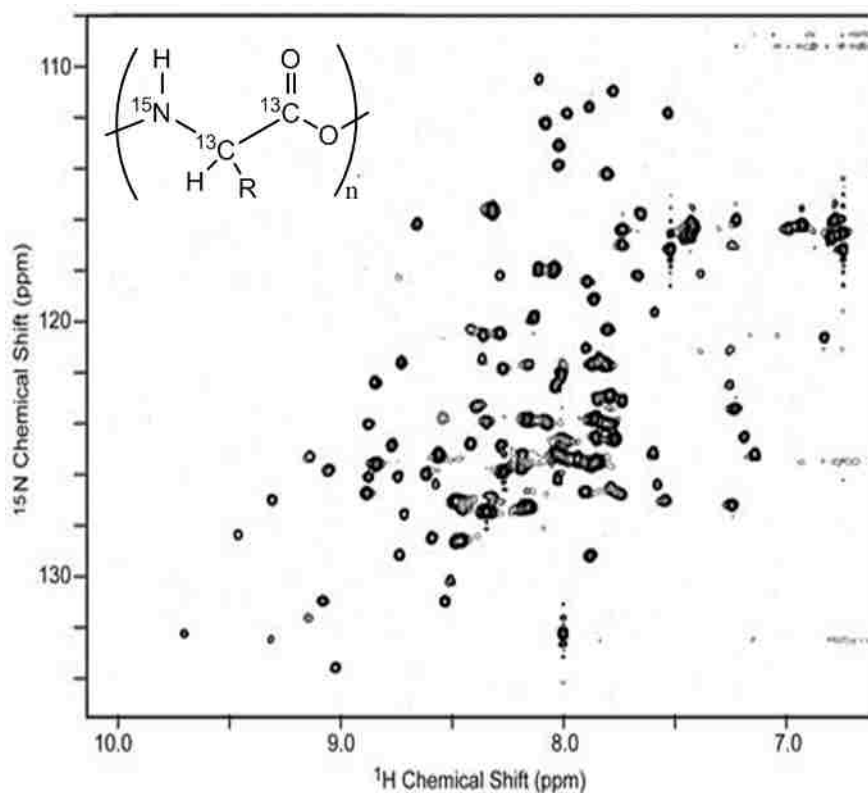


Figure 4-9 ^1H - ^{15}N HSQC of ^{15}N labeled CT663. Each peak corresponds to an ^1H - ^{15}N bond from the backbone or a side chain of the protein.

4. 3. 8 NMR buffer optimization for CT663

NMR conditions are optimized to obtain good data. The buffer conditions and protein concentration are varied until an appropriate spectrum is obtained. A good HSQC has well resolve peaks of similar intensity, with one peak for every amino acid except proline. Peak distribution is also important, as it gives information about the folding of the protein. Well disperse peaks are necessary for a structure determination. ^1H - ^{15}N HSQC of ^{15}N labeled CT663 in 20 mM sodium phosphate produced peaks that were dispersed. However, the peak intensities are not similar and not resolved between 8-8.5 ppm. CT663 in 20 mM tris produce less peaks that vary in intensity, that lack dispersion. CT663 in 20 mM MES produced peaks that were not dispersed or similar in

intensity. Tris is a terrible buffer for CT663 because it does not allow for the detection of an adequate number of amide peaks. There are 142 amino acids that compose CT663, accounting for the proline residues there should be 139 backbone amide peaks plus additional side chain peaks. There are more peaks for samples in MES and phosphate buffer; however, less than 90% of the amide peaks accounted. Peak dispersion is also a problem, the peaks between 8-8.5 ppm appear to be overlapping and of different intensity. These peaks indicated that the protein may be aggregated or unfolded.

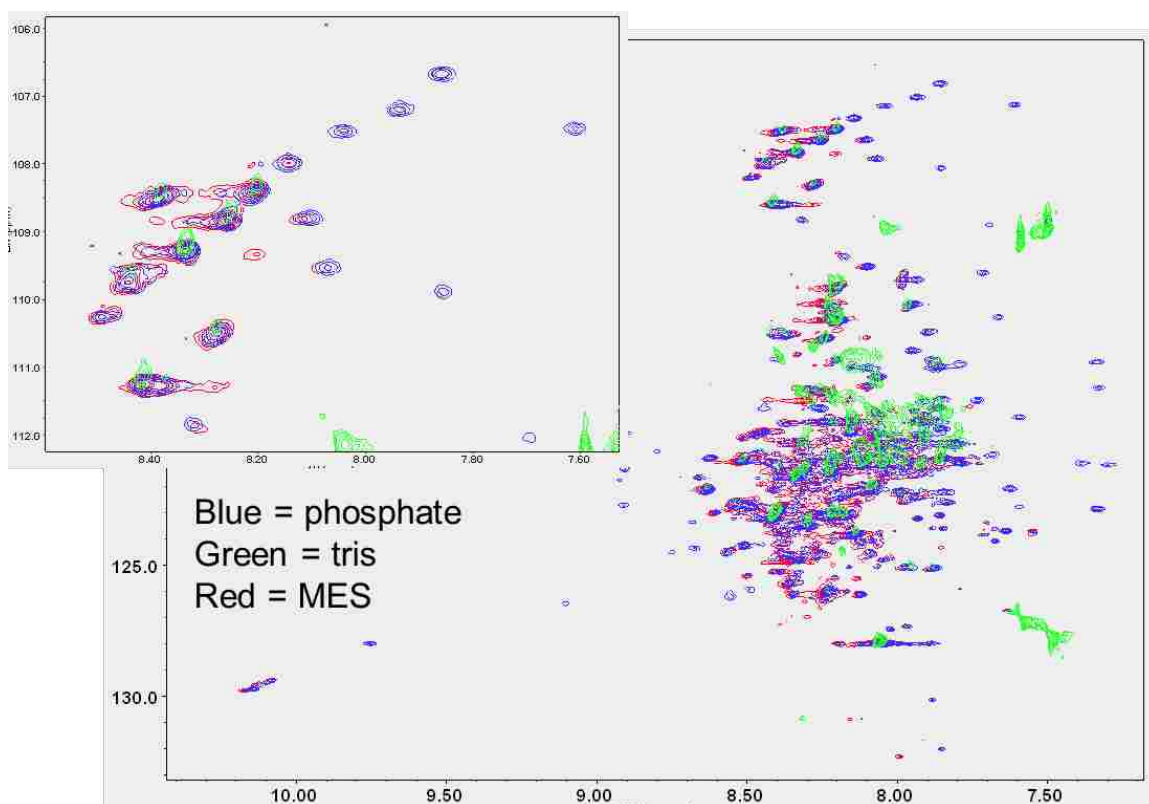


Figure 4-10 ^1H - ^{15}N HSQC of ^{15}N labeled CT663 in 20 mM sodium phosphate, 20 mM tris, and 20 mM MES.

4. 3. 8 ^{15}N -CT663:Scc1 NMR analysis

The nature of the CT663/Scc1 complex formation is not known. In an effort to analyze the effect of Scc1 on CT663, unlabeled Scc1 was added to ^{15}N CT663 and the ^1H - ^{15}N HSQC was analyzed. Figure 4-11 shows the overlapped ^1H - ^{15}N HSQC for ^{15}N

CT663 and ^{15}N CT663 complexed with Scc1. The spectra show shifts in the amide peaks that correspond to changes in the backbone structure. The inset of Figure 4-11 shows 13 peaks, 9 of which show varying degrees of shift, signifying a local structural change.

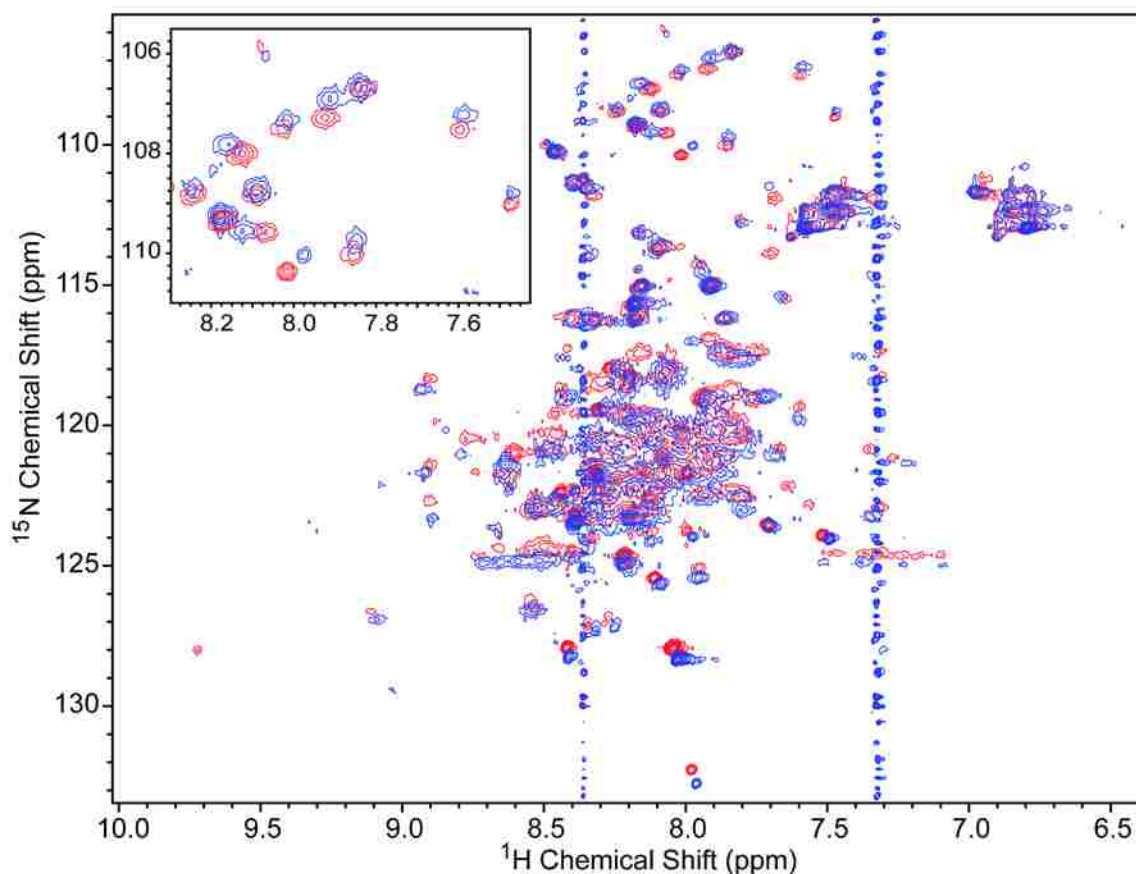


Figure 4-11 Scc1 perturbs the amide peaks of ^{15}N -CT663. An overlay of ^1H - ^{15}N HSQC spectra of ^{15}N -CT663 (red) and ^{15}N -CT663:Scc1 (blue). The inset illustrates the selective shifts of the putative glycine amide peaks upon binding Scc1. Note: No peaks are covered by the inset.

4. 3. 9 NMR analysis of the CT663/Scc1 complex

Preliminary NMR data suggested that the Scc1/CT663 complex is a good sample for structure determination, even though it is ~37 kDa. ^{15}N -Scc1/CT663 was expressed and purified. The sample was given to Abigail Songok, a graduate student in our lab, for further analysis. The sample was desalted into 20 mM ammonium acetate at pH 5.

5, frozen at $-80\text{ }^{\circ}\text{C}$, and lyophilized. The protein was reconstituted in 20 mM ammonium acetate and 10 mM DTT at pH 5.5 with 10% D_2O . The sample was concentrated to 250 μL and a ^1H - ^{15}N HSQC spectrum was collected. The spectrum looks great with 95% of the amide peaks accounted. The inset of Figure 4-12 shows the peaks are dispersed in the zoomed view.

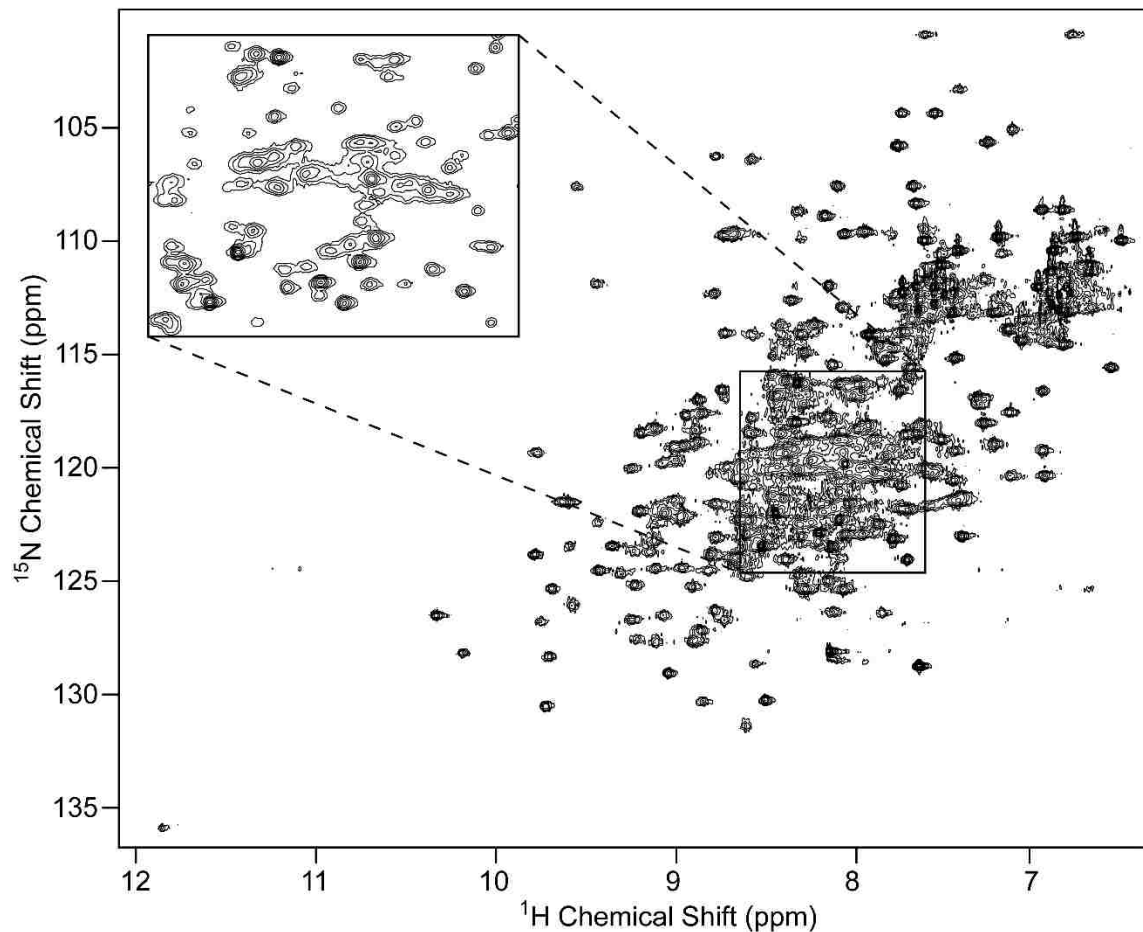


Figure 4-12 ^1H - ^{15}N HSQC of ^{15}N labeled Scc1/CT663. The inset is a zoomed view to show the peak dispersion.

4. 4 Conclusions

The goal of this research is to analyze the structure and ligand interactions of CT663. CT663 forms a tightly associated complex with Scc1 that cannot be disrupted with urea. This suggests that hydrophobic interactions play a minor role in keeping the

complex intact. A CopN pull down assay performed by Li Shen showed that CopN does not bind CT663 or Scc1 independently, but binds the Scc1/CT663 complex. What remains unclear is where, spatially, CopN binds the complex. The Phyre2 server predicted a linear structure, where CopN is bound mainly to CT663, and CT663 is bound to Scc1. This quaternary structure is possible if the binding of CT663 to Scc1 causes a conformational change of CT663 that is favorable for CopN binding. It is also possible that there is an area near the interface of the Scc1/CT663 complex where CopN binds to both proteins. The binding kinetics of Scc1/CT663 was analyzed by SPR. The interaction between the chaperone complex and CopN is strong bond with a K_d of 3.02 ± 0.17 nM.

In order to obtain good data for structural analysis with NMR, proteins need to be in their monomeric state or in small (< 25 kDa) multimeric state. The results of the gel filtration experiments indicate that CT663 forms a dimer or a higher order oligomer depending on the buffer conditions. Stabilizing the CT663 monomer is not straightforward. The overlapped peaks in the center of the ^1H - ^{15}N HSQC of ^{15}N -CT663 in sodium phosphate and tris support the finding that CT663 is aggregating under these conditions. The ^1H - ^{15}N HSQC of the ^{15}N -CT663 sample sent to us by our collaborators showed a very different scenario. The buffer conditions were never replicated because the protein we produced precipitated when the pH was lowered to 5. After several meetings, it was discovered that the CT663 in the magical sample lacked the His-tag used for purification. The His-tag increases the isoelectric point (pI) of CT663 causing it to precipitate at pH 5 and may also be the source of the aggregation issues. The CT663 that is in complex with Scc1 does not have a His-tag, thus sarkosyl was used to

separate CT663 from the complex for NMR analysis. Abigael Songok is continuing this work. If the NMR spectra of CT663 cannot be replicated using sarkosyl to purify the protein, then CT663 without the His-tag can be produced and purified using a combination of chromatography methods.

Initial NMR experiments showed that when Scc1 is complexes CT663, there is a localized change in the backbone structure of CT663. In order to get a better understanding of how the structure of CT663 changes, the structure of the complex will be analyzed through complete and mixed isotopic labeling. The ^1H - ^{15}N HSQC of ^{15}N Scc1/CT663 is very promising—it should lead to a full 3D structure of the complex. Adding CopN to the isotopically labeled complex may give information, such as the residues of CT663 and Scc1 that bind CopN and changes to the CT663:Scc1 structure when CopN is bound. Knowledge of the binding, even between the complex and CopN, could lead to inhibitors that could be used in conjunction with current therapeutics to treat *Chlamydia trachomatis*.

References

1. Cooper, G. M. *The cell: A molecular approach*. The Molecular Composition of Cells 2000; 2nd edition:[Available from: <http://www.ncbi.nlm.nih.gov/books/NBK9879/>].
2. Nester, E. W. , *Microbiology: A human perspective*. 3rd Edition ed. 2001, Michigan: McGraw Hill Higher Education.
3. Lehninger, A. L. , D. L. Nelson, and M. M. Cox, *Lehninger principles of biochemistry*. 5th Edition ed. 2005, New York: W. H. Freeman.
4. Varki, A. and M. J. Chrispeels, *Essentials of glycobiology*. 1999, New York: Cold Spring Harbor Laboratory Press.
5. Banerjee, P. S. , G. W. Hart, and J. W. Cho, *Chemical approaches to study o-glcnacylation*. Chemical Society Reviews, 2013. **42**(10): p. 4345-4357.
6. Santoyo-Ramos, P. , M. C. Castañeda-Patlán, and M. Robles-Flores, *The role of o-linked beta-n-acetylglucosamine (glcnac) modification in cell signaling*. 2012, Romania Intech.
7. Schleicher, E. D. and C. Weigert, *Role of the hexosamine biosynthetic pathway in diabetic nephropathy*. Kidney International Supplement, 2000. **58**(S77): p. S13-S18.
8. Darley-USmar, V. M. , L. E. Ball, and J. C. Chatham, *Protein o-linked β -n-acetylglucosamine: A novel effector of cardiomyocyte metabolism and function*. Journal of Molecular and Cellular Cardiology, 2012. **52**(3): p. 538-549.
9. Fang, W. , Du, T. , Raimi, O. G. , Hurtado Guerrero, R. , Mariño, K. , Ibrahim, A. F. M. , Albarbarawi, O. , Ferguson, M. A. J. , Jin, C. and Van Aalten, D. M. F. , *Genetic and structural validation of aspergillus fumigatus n-acetylphosphoglucosamine mutase as an antifungal target*. Bioscience. Report, 2013. **33**(5): p. 11.
10. Hart, G. W. , M. P. Housley, and C. Slawson, *Cycling of o-linked beta-n-acetylglucosamine on nucleocytoplasmic proteins*. Nature, 2007. **446**(7139): p. 1017-22.
11. Whelan, S. and G. Hart, *Identification of o-glcnac sites on proteins*. Methods Enzymol, 2006. **415**: p. 113 - 133.
12. Butkinaree, C. , K. Park, and G. Hart, *O-linked n-acetylglucosamine (o-glcnac): Extensive crosstalk with phosphorylation to regulate signaling and transcription in*

- response to nutrients and stress*. Biochem Biophys Acta, 2010. **2010**(1800): p. 96 - 106.
13. Torres, C. and G. Hart, *Topography and polypeptide distribution of terminal n-acetylglucosamine residues on the surfaces of intact lymphocytes. Evidence for o-linked glcnac*. J Biol Chem, 1984. **259**: p. 3308 - 3317.
 14. Liu, Y. , X. Li, Y. Yu, J. Shi, Z. Liang, X. Run, Y. Li, C. -I. Dai, I. Grundke-Iqbal, K. Iqbal, F. Liu, and C. -X. Gong, *Developmental regulation of protein o-glcnacylation, o-glcnac transferase, and o-glcnacase in mammalian brain*. Plos One, 2012. **7**(8): p. e43724.
 15. Ma, J. and G. Hart, *O-glcnac profiling: From proteins to proteomes*. Clinical Proteomics, 2014. **11**(1): p. 8.
 16. Hanover, J. A. , M. W. Krause, and D. C. Love, *The hexosamine signaling pathway: O-glcnac cycling in feast or famine*. Biochimica et Biophysica Acta (BBA) - General Subjects, 2010. **1800**(2): p. 80-95.
 17. Khidekel, N. , S. Arndt, N. Lamarre-Vincent, A. Lippert, K. Poulin-Kerstien, B. Ramakrishnan, P. Qasba, and L. Hsieh-Wilson, *A chemoenzymatic approach toward the rapid and sensitive detection of o-glcnac posttranslational modifications*. Journal of the American Chemical Society, 2003. **125**: p. 16162 - 16163.
 18. Rexach, J. , P. Clark, D. Mason, R. Neve, E. Peters, and L. Hsieh-Wilson, *Dynamic o-glcnac modification regulates creb-mediated gene expression and memory formation*. Nature Chemical Biology, 2012. **8**: p. 253 - 261.
 19. Borman, S. *Memories' sweet origins*. Available from: <http://cen.acs.org/articles/90/web/2012/02/MemoriesSweet-Origins.html>.
 20. Rexach, J. E. , P. M. Clark, and L. C. Hsieh-Wilson, *Chemical approaches to understanding o-glcnac glycosylation in the brain*. Nature Chemical Biology, 2008. **4**(2): p. 97-106.
 21. Klein, S. B. and B. M. Thorne, *Biological psychology*. 2006, New York: Worth Publisher.
 22. Yu, Y. , L. Zhang, X. Li, X. Run, Z. Liang, Y. Li, Y. Liu, M. H. Lee, I. Grundke-Iqbal, K. Iqbal, D. J. Vocadlo, F. Liu, and C. -X. Gong, *Differential effects of an o-glcnacase inhibitor on tau phosphorylation*. PLoS ONE, 2012. **7**(4): p. e35277.
 23. Yuzwa, S. A. , X. Shan, M. S. Macauley, T. Clark, Y. Skorobogatko, K. Vosseller, and D. J. Vocadlo, *Increasing o-glcnac slows neurodegeneration and stabilizes tau against aggregation*. Nature Chemical Biology, 2012. **8**(4): p. 393-399.

24. Yuzwa, S. , A. Yadav, Y. Skorobogatko, T. Clark, K. Vosseller, and D. Vocadlo, *Mapping o-glcnaC modification sites on tau and generation of a site-specific o-glcnaC tau antibody*. Amino Acids, 2011. **40**: p. 857 - 868.
25. Kamigaito, T. , T. Okaneya, M. Kawakubo, H. Shimojo, O. Nishizawa, and J. Nakayama, *Overexpression of o-glcnaC by prostate cancer cells is significantly associated with poor prognosis of patients*. Prostate Cancer Prostatic Disease, 2014. **17**(1): p. 18-22.
26. Gu, Y. , W. Mi, Y. Ge, H. Liu, Q. Fan, C. Han, J. Yang, F. Han, X. Lu, and W. Yu, *GlcnaCylation plays an essential role in breast cancer metastasis*. Cancer Research, 2010. **70**(15): p. 6344-6351.
27. Issad, T. , E. Masson, and P. Pagesy, *O-glcnaC modification, insulin signaling and diabetic complications*. Diabetes & Metabolism, 2010. **36**: p. 423 - 435.
28. Kahn, B. B. and J. S. Flier, *Obesity and insulin resistance*. The Journal of Clinical Investigation, 2000. **106**(4): p. 473-481.
29. Konrad, R. J. , I. Mikolaenko, J. F. Tolar, K. Liu, and J. E. Kudlow, *The potential mechanism of the diabetogenic action of streptozotocin: Inhibition of pancreatic beta-cell o-glcnaC-selective n-acetyl-beta-d-glucosaminidase*. Biochemical Journal, 2001. **356**(1): p. 31-41.
30. Lameira, J. , C. u. N. Alves, V. Moliner, S. Martí, R. Castillo, and I. a. Tuñón, *Quantum mechanical/molecular mechanical molecular dynamics simulation of wild-type and seven mutants of cpnagj in complex with pugnac*. The Journal of Physical Chemistry B, 2010. **114**(20): p. 7029-7036.
31. Yuzwa, S. A. , M. S. Macauley, J. E. Heinonen, X. Shan, R. J. Dennis, Y. He, G. E. Whitworth, K. A. Stubbs, E. J. McEachern, G. J. Davies, and D. J. Vocadlo, *A potent mechanism-inspired o-glcnaCase inhibitor that blocks phosphorylation of tau in vivo*. Nature Chemical Biology, 2008. **4**(8): p. 483-490.
32. Dorfmüller, H. C. , V. S. Borodkin, M. Schimpl, and D. M. F. van Aalten, *GlcnaCstatins are nanomolar inhibitors of human o-glcnaCase inducing cellular hyper-o-glcnaCylation*. Biochemical Journal, 2009. **420**(2): p. 221-227.
33. Rexach, J. , C. Rogers, S. Yu, J. Tao, Y. Sun, and L. Hsieh-Wilson, *Quantification of o-glycosylation stoichiometry and dynamics using resolvable mass tags*. Nature Chemical Biology, 2010. **6**: p. 645 - 651.
34. J T Gallagher, A. M. , and T M Dexter, *Identification of two binding sites for wheat-germ agglutinin on polyactosamine-type oligosaccharides*. Biochemical Journal, 1985. **231**(1): p. 115-122.
35. Stalnaker, S. H. , S. Hashmi, J. -M. Lim, K. Aoki, M. Porterfield, G. Gutierrez-Sanchez, J. Wheeler, J. M. Ervasti, C. Bergmann, M. Tiemeyer, and L. Wells,

- Site mapping and characterization of o-glycan structures on α -dystroglycan isolated from rabbit skeletal muscle.* Journal of Biological Chemistry, 2010. **285**(32): p. 24882-24891.
36. Wells, L. , K. Vosseller, R. N. Cole, J. M. Cronshaw, M. J. Matunis, and G. W. Hart, *Mapping sites of o-glcnac modification using affinity tags for serine and threonine post-translational modifications.* Molecular & Cellular Proteomics, 2002. **1**(10): p. 791-804.
 37. Sekine, O. , D. C. Love, D. S. Rubenstein, and J. A. Hanover, *Blocking o-linked glcnac cycling in drosophila insulin-producing cells perturbs glucose-insulin homeostasis.* Journal of Biological Chemistry, 2010. **285**(49): p. 38684-38691.
 38. Chung, S. , D. -O. Kang, J. Yamzon, D. Warburton, and C. J. Koh, *O-glcnac mediated glycosylation down-regulation in mice with cyclophosphamide induced cystitis.* The Journal of Urology, 2010. **183**(1): p. 351-356.
 39. Yee, A. , X. Chang, A. Pineda-Lucena, B. Wu, A. Semesi, B. Le, T. Ramelot, G. M. Lee, S. Bhattacharyya, P. Gutierrez, A. Denisov, C. -H. Lee, J. R. Cort, G. Kozlov, J. Liao, G. Finak, L. Chen, D. Wishart, W. Lee, L. P. McIntosh, K. Gehring, M. A. Kennedy, A. M. Edwards, and C. H. Arrowsmith, *An nmr approach to structural proteomics.* Proceedings of the National Academy of Sciences, 2002. **99**(4): p. 1825-1830.
 40. Higman, V. A. *Protein nmr: A practical guide.* 2012; Available from: <http://www.protein-nmr.org.uk/>.
 41. Johannessen, M. , M. P. Delghandi, and U. Moens, *What turns creb on?* Cellular Signalling, 2004. **16**(11): p. 1211-1227.
 42. Zubenko, G. S. , H. B. Hughes, B. S. Maher, J. S. Stiffler, W. N. Zubenko, and M. L. Marazita, *Genetic linkage of region containing the creb1 gene to depressive disorders in women from families with recurrent, early-onset, major depression.* American Journal of Medical Genetics, 2002. **114**(8): p. 980-987.
 43. Gaiarsa, J. -L. , O. Caillard, and Y. Ben-Ari, *Long-term plasticity at gabaergic and glycinergic synapses: Mechanisms and functional significance.* Trends in Neurosciences. **25**(11): p. 564-570.
 44. Josselyn SA1, S. C. , Carlezon WA Jr, Neve RL, Nestler EJ, Davis M. , *Long-term memory is facilitated by camp response element-binding protein overexpression in the amygdala.* The Journal of Neuroscience, 2001. **21**(7): p. 2404-2412.
 45. Baba, T. , T. Ara, M. Hasegawa, Y. Takai, Y. Okumura, M. Baba, K. A. Datsenko, M. Tomita, B. L. Wanner, and H. Mori, *Construction of escherichia coli k-12 in-frame, single-gene knockout mutants: The keio collection.* Molecular Systems Biology, 2006. **2**(16738554): p. 11.

46. Cormier, C. , J. Park, M. Fiacco, J. Steel, P. Hunter, J. Kramer, R. Singla, and J. LaBaer, *Psi: Biology-materials repository: A biologist's resource for protein expression plasmids*. Journal of Structural and Functional Genomics, 2011. **12**(2): p. 55-62.
47. Lazarus, B. D. , D. C. Love, and J. A. Hanover, *Recombinant o-glcnae transferase isoforms: Identification of o-glcnae, yes tyrosine kinase, and tau as isoform-specific substrates*. Glycobiology, 2006. **16**(16434389): p. 415-421.
48. Abràmoff, M. D. , P. J. Magalhães, and S. J. Ram, *Image processing with imagej*. Biophotonics international, 2004. **11**(7): p. 36-43.
49. Frey, P. A. , F. C. Kokesh, and F. H. Westheimer, *A reporter group at the active site of acetoacetate decarboxylase. I. Ionization constant of the nitrophenol*. Journal of the American Chemical Society, 1971. **93**(5127415): p. 7266-7269.
50. Projan, S. J. and R. P. Novick, *Reciprocal intrapool variation in plasmid copy numbers: A characteristic of segregational incompatibility*. Plasmid, 1984. **12**(6494316): p. 52-60.
51. Lim KH, H. C. , Chang HI, *Production of o-glcnae modified recombinant proteins in escherichia coli*. Journal Of Microbiology And Biotechnology, 2002. **12**.
52. Riu, I. -H. , I. -S. Shin, and S. -I. Do, *Sp1 modulates ncogt activity to alter target recognition and enhanced thermotolerance in e. Coli*. Biochemical and Biophysical Research Communications, 2008. **372**(18486602): p. 203-209.
53. Lim, K. -H. and H. -I. Chang, *O-linked n-acetylglucosamine suppresses thermal aggregation of sp1*. FEBS Letters, 2006. **580**(16879824): p. 4645-4652.
54. Liu G, H. Y. , Xiao R, Wang D, Acton TB, Montelione GT, *Nmr structure of f-actin binding domain of arg/abl2 from homo sapiens*. Proteins: Structure, Function, and Bioinformatics, 2010. **78**(5): p. 1326–1330.
55. Bhonagiri, P. , G. R. Pattar, K. M. Habegger, A. M. McCarthy, L. Tackett, and J. S. Elmendorf, *Evidence coupling increased hexosamine biosynthesis pathway activity to membrane cholesterol toxicity and cortical filamentous actin derangement contributing to cellular insulin resistance*. Endocrinology, 2011. **152**(21712361): p. 3373-3384.
56. Hedou, J. , B. Bastide, A. Page, J. Michalski, and W. Morelle, *Mapping of o-linked beta-n-acetylglucosamine modification sites in key contractile proteins of rat skeletal muscle*. Proteomics, 2009. **9**: p. 2139 - 2148.
57. Gupta, R. and S. Brunak, *Prediction of glycosylation across the human proteome and the correlation to protein function*. Pacific Symposium On Biocomputing. , 2002. **7**: p. 310 - 322.

58. Lubas, W. A. and J. A. Hanover, *Functional expression of o-linked glcnac transferase. Domain structure and substrate specificity*. Journal Biological Chemistry, 2000. **275**(10753899): p. 10983-10988.
59. Yem, D. W. and H. C. Wu, *Isolation of escherichia coli k-12 mutants with altered level of beta-n-acetylglucosaminidase*. Journal of Bacteriology, 1976. **125**(1107312): p. 372-373.
60. Cheng, Q. , H. Li, K. Merdek, and J. T. Park, *Molecular characterization of the beta-n-acetylglucosaminidase of escherichia coli and its role in cell wall recycling*. Journal of Bacteriology, 2000. **182**(10940025): p. 4836-4840.
61. Votsch, W. and M. F. Templin, *Characterization of a beta -n-acetylglucosaminidase of escherichia coli and elucidation of its role in muropeptide recycling and beta -lactamase induction*. Journal Biological Chemistry, 2000. **275**(10978324): p. 39032-39038.
62. Beer D, M. J. , Rast DM, Vasella A. ,, *Synthesis of 2-acetamido-2-deoxy-d-gluconhydroximolactone- and chitobionhydroximolactone-derived n-phenylcarbamates, potential inhibitors of β -n-acetylglucosaminidase*. Helvetica Chimica Acta, 1990. **73**(7): p. 1918–1922.
63. Haltiwanger, R. S. , K. Grove, and G. A. Philipsberg, *Modulation of o-linked n-acetylglucosamine levels on nuclear and cytoplasmic proteins in vivo using the peptide o-glcna β -n-acetylglucosaminidase inhibitor o-(2-acetamido-2-deoxy-d-glucopyranosylidene)amino-n-phenylcarbamate*. Journal Biological Chemistry, 1998. **273**(9452489): p. 3611-3617.
64. Stubbs, K. A. , M. Balcewich, B. L. Mark, and D. J. Vocadlo, *Small molecule inhibitors of a glycoside hydrolase attenuate inducible ampc-mediated beta-lactam resistance*. Journal Biological Chemistry, 2007. **282**(17439950): p. 21382-21391.
65. Horsch, M. , L. Hoesch, A. Vasella, and D. M. Rast, *N-acetylglucosaminono-1,5-lactone oxime and the corresponding (phenylcarbamoyl)oxime. Novel and potent inhibitors of beta-n-acetylglucosaminidase*. European Journal of Biochemistry, 1991. **197**(2029909): p. 815-818.
66. Deng, M. -D. , D. K. Severson, A. D. Grund, S. L. Wassink, R. P. Burlingame, A. Berry, J. A. Running, C. A. Kunesh, L. Song, T. A. Jerrell, and R. A. Rosson, *Metabolic engineering of escherichia coli for industrial production of glucosamine and n-acetylglucosamine*. Metabolic Engineering, 2005. **7**(3): p. 201-214.
67. Rodríguez-Díaz, J. , A. Rubio-del-Campo, and M. J. Yebra, *Metabolic engineering of lactobacillus casei for production of udp-n-acetylglucosamine*. Biotechnology and Bioengineering, 2012. **109**(7): p. 1704-1712.

68. Uehara, T. and J. T. Park, *The n-acetyl-d-glucosamine kinase of escherichia coli and its role in murein recycling*. Journal of Bacteriology, 2004. **186**(21): p. 7273-7279.
69. Cheng, Q. , H. Li, K. Merdek, and J. T. Park, *Molecular characterization of the β -n-acetylglucosaminidase of escherichia coli and its role in cell wall recycling*. Journal of Bacteriology, 2000. **182**(17): p. 4836-4840.
70. Triassi, A. J. , M. S. Wheatley, M. Savka, H. M. Gan, R. C. J. Dobson, and A. O. B. Hudson, *L,I-diaminopimelate aminotransferase (dapI): A putative target for the development of narrow-spectrum antibacterial compounds*. Frontiers in Microbiology, 2014. **5**.
71. Kreppel, L. K. and G. W. Hart, *Regulation of a cytosolic and nuclear o-glcnae transferase: Role of the tetratricopeptide repeats*. Journal of Biological Chemistry, 1999. **274**(45): p. 32015-32022.
72. Baba, T. , T. Ara, M. Hasegawa, Y. Takai, Y. Okumura, M. Baba, K. A. Datsenko, M. Tomita, B. L. Wanner, and H. Mori, *Construction of escherichia coli k-12 in-frame, single-gene knockout mutants: The keio collection*. 2006.
73. Álvarez-Añorve, L. I. , M. L. Calcagno, and J. Plumbridge, *Why does escherichia coli grow more slowly on glucosamine than on n-acetylglucosamine? Effects of enzyme levels and allosteric activation of glcn6p deaminase (nagb) on growth rates*. Journal of Bacteriology, 2005. **187**(9): p. 2974-2982.
74. Barton, L. , *Structural and functional relationships in prokaryotes*. 2005, New York: Springer.
75. Gurcel, C. , A. -S. Vercoutter-Edouart, C. Fonbonne, M. Mortuaire, A. Salvador, J. -C. Michalski, and J. Lemoine, *Identification of new o-glcnae modified proteins using a click-chemistry-based tagging*. Analytical and Bioanalytical Chemistry, 2008. **390**(8): p. 2089-2097.
76. Brunham, R. C. and J. Rey-Ladino, *Immunology of chlamydia infection: Implications for a chlamydia trachomatis vaccine*. Nature Reviews Immunology, 2005. **5**(2): p. 149-161.
77. CDC. *Cdc fact sheet*. 2013 [cited 2014 October 1,]; Available from: <http://www.cdc.gov/std/stats/sti-estimates-fact-sheet-feb-2013.pdf>.
78. Paavonen, J. and W. Eggert-Kruse, *Chlamydia trachomatis: Impact on human reproduction*. Human Reproduction Update, 1999. **5**(5): p. 433-447.
79. Somani, J. , V. B. Bhullar, K. A. Workowski, C. E. Farshy, and C. M. Black, *Multiple drug-resistant chlamydia trachomatis associated with clinical treatment failure*. Journal of Infectious Diseases, 2000. **181**(4): p. 1421-1427.

80. Scott LaMontagne, D. , K. Baster, L. Emmett, T. Nichols, S. Randall, L. McLean, P. Meredith, V. Harindra, J. M. Tobin, G. S. Underhill, W. Graham Hewitt, J. Hopwood, T. Gleave, A. K. Ghosh, H. Mallinson, A. R. Davies, G. Hughes, and K. A. Fenton, *Incidence and reinfection rates of genital chlamydial infection among women aged 16–24 years attending general practice, family planning and genitourinary medicine clinics in england: A prospective cohort study by the chlamydia recall study advisory group*. Sexually Transmitted Infections, 2007. **83**(4): p. 292-303.
81. Wenman, W. M. and R. U. Meuser, *Chlamydia trachomatis elementary bodies possess proteins which bind to eucaryotic cell membranes*. Journal of Bacteriology, 1986. **165**(2): p. 602-607.
82. Bao, X. , B. E. Nickels, and H. Fan, *Chlamydia trachomatis protein grga activates transcription by contacting the nonconserved region of sigma(66)*. Proceedings of the National Academy of Sciences of the United States of America, 2012. **109**(42): p. 16870-16875.
83. Bloom, B. R. and P. H. Lambert, *The vaccine book*. 2003, California: Academic Press.
84. Archuleta, T. L. , Y. Du, C. A. English, S. Lory, C. Lesser, M. D. Ohi, R. Ohi, and B. W. Spiller, *The chlamydia effector chlamydial outer protein n (copn) sequesters tubulin and prevents microtubule assembly*. Journal of Biological Chemistry, 2011. **286**(39): p. 33992-33998.
85. Pais, S. V. , C. Milho, F. Almeida, and L. J. Mota, *Identification of novel type iii secretion chaperone-substrate complexes of chlamydia trachomatis*. Plos One, 2013. **8**(2).
86. Büttner D , a. H. S. Y. , *Type iii protein secretion in plant pathogenic bacteria*. Plant Physiology, 2009. **150**(4): p. 1656-1664
87. Betts-Hampikian, H. J. and K. Fields, *The chlamydial type iii secretion mechanism: Revealing cracks in a tough nut*. Frontiers in Microbiology, 2010. **1**.
88. Parsot, C. , C. Hamiaux, and A. L. Page, *The various and varying roles of specific chaperones in type iii secretion systems*. Current Opinion in Microbiology, 2003. **6**(1): p. 7-14.
89. Rao, X. C. , P. Deighan, Z. Y. Hua, X. M. Hu, J. Wang, M. Luo, J. Wang, Y. M. Liang, G. M. Zhong, A. Hochschild, and L. Shen, *A regulator from chlamydia trachomatis modulates the activity of rna polymerase through direct interaction with the beta subunit and the primary sigma subunit*. Genes & Development, 2009. **23**(15): p. 1818-1829.



90. Huang, J. , C. F. Lesser, and S. Lory, *The essential role of the copn protein in chlamydia pneumoniae intracellular growth*. Nature, 2008. **456**(7218): p. 112-115.
91. Santini, S. , S. Di Agostino, E. Coppari, A. R. Bizzarri, G. Blandino, and S. Cannistraro, *Interaction of mutant p53 with p73: A surface plasmon resonance and atomic force spectroscopy study*. Biochimica et Biophysica Acta (BBA) - General Subjects, 2014. **1840**(6): p. 1958-1964.
92. Izore, T. , V. Job, and A. Dessen, *Biogenesis, regulation, and targeting of the type iii secretion system*. Structure, 2011. **19**(5): p. 603-612.
93. Krieger, E. , S. B. Nabuurs, and G. Vriend, *Homology modeling*, in *Structural bioinformatics2005*, John Wiley & Sons, Inc. p. 509-523.
94. Cooper, C. , Zhang, K. , Andres, S. N. , Hannemann, M. , Junop, M. S. , Coombes, B. , *Structural analysis and quantitative proteomic interactome of a novel virulence chaperone in salmonella*. Protein Data Bank, 2009.
95. Galán, J. E. and A. Collmer, *Type iii secretion machines: Bacterial devices for protein delivery into host cells*. Science, 1999. **284**(5418): p. 1322-1328.
96. Phan, J. , J. E. Tropea, and D. S. Waugh, *Structure of the yersinia pestis type iii secretion chaperone sych in complex with a stable fragment of yscm2*. Acta Crystallographica Section D, 2004. **60**(9): p. 1591-1599.
97. Luo, Y. , M. G. Bertero, E. A. Frey, R. A. Pfuetzner, M. R. Wenk, L. Creagh, S. L. Marcus, D. Lim, F. Sicheri, C. Kay, C. Haynes, B. B. Finlay, and N. C. J. Strynadka, *Structural and biochemical characterization of the type iii secretion chaperones cest and sigE*. Nature Structural & Molecular Biology, 2001. **8**(12): p. 1031-1036.
98. Stebbins, C. E. and J. E. Galan, *Maintenance of an unfolded polypeptide by a cognate chaperone in bacterial type iii secretion*. Nature, 2001. **414**(6859): p. 77-81.
99. Schubot, F. D. , M. W. Jackson, K. J. Penrose, S. Cherry, J. E. Tropea, G. V. Plano, and D. S. Waugh, *Three-dimensional structure of a macromolecular assembly that regulates type iii secretion in yersinia pestis*. Journal of Molecular Biology, 2005. **346**(4): p. 1147-1161.
100. Deane, J. E. , P. Roversi, C. King, S. Johnson, and S. M. Lea, *Structures of the shigella flexneri type 3 secretion system protein mxic reveal conformational variability amongst homologues*. Journal of Molecular Biology, 2008. **377**(4): p. 985-992.

101. Bennion, B. J. and V. Daggett, *The molecular basis for the chemical denaturation of proteins by urea*. Proceedings of the National Academy of Sciences, 2003. **100**(9): p. 5142-5147.

Appendix

Letter of Permission

This is the permission to use my publication in chapter 2 of this thesis.



Thank You For Your Order!

Dear Dr. Octavia Goodwin,

Thank you for placing your order through Copyright Clearance Center's RightsLink service. Elsevier has partnered with RightsLink to license its content. This notice is a confirmation that your order was successful.

Your order details and publisher terms and conditions are available by clicking the link below:
<http://s100.copyright.com/CustomerAdmin/PLF.jsp?ref=75cfbe9f-8231-4bef-bf84-c80c893d44e4>

Order Details


Licensee: Louisiana State University
License Date: Nov 15, 2014
License Number: 3510481191108
Publication: Journal of Biotechnology
Title: E. coli sabotages the in vivo production of O-linked β -N-acetylglucosamine-modified proteins
Type Of Use: reuse in a thesis/dissertation
Total: 0.00 USD

To access your account, please visit <https://myaccount.copyright.com>.

Please note: Online payments are charged immediately after order confirmation; invoices are issued daily and are payable immediately upon receipt.

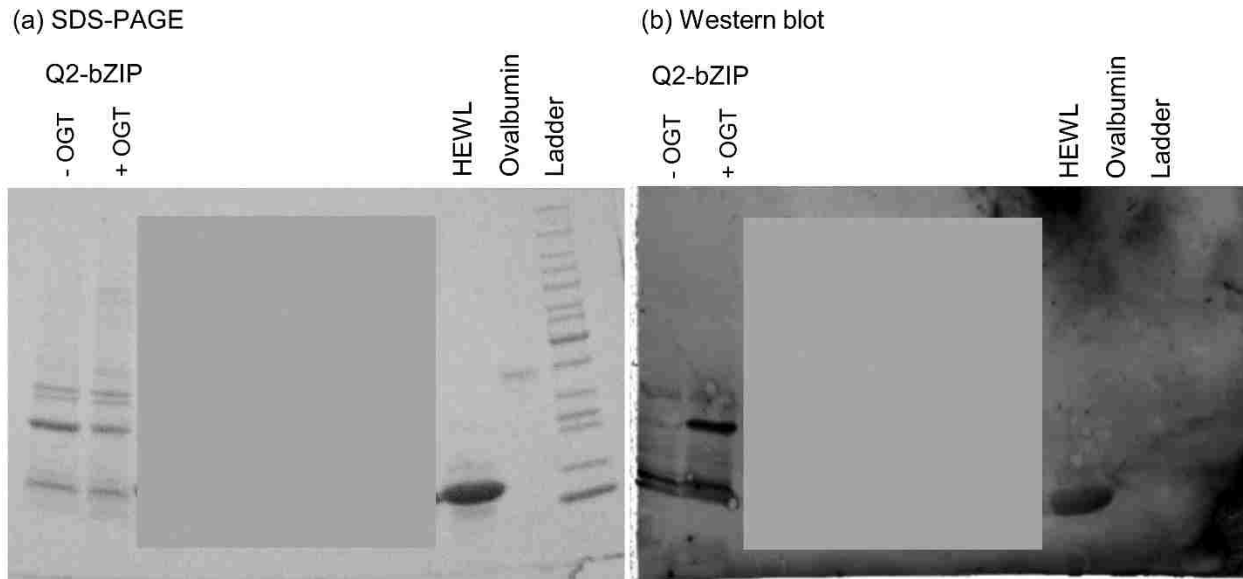
To ensure that we are continuously improving our services, please take a moment to complete our [customer satisfaction survey](#).

B.1:v4.2

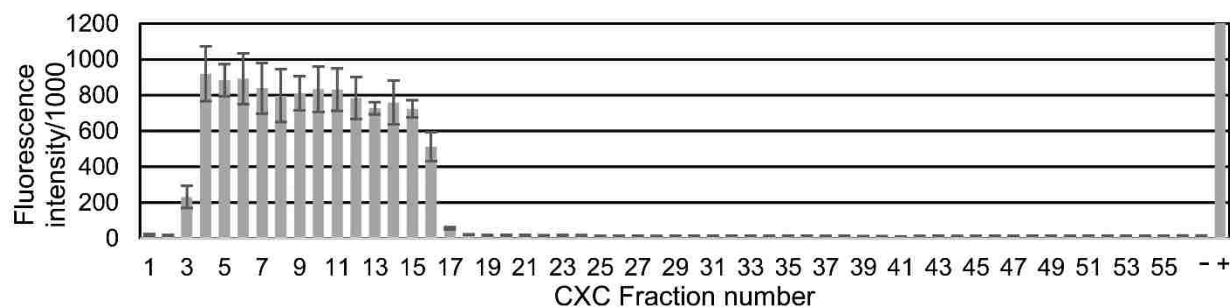


+1-855-239-3415 / Tel: +1-978-646-2777
customer care@copyright.com
<http://www.copyright.com>

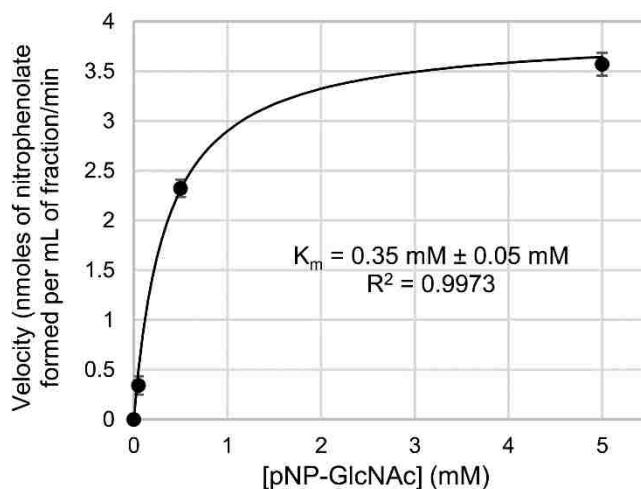
Supplementary Figures



Supplementary Figure 1. Ni-affinity purified, resolubilized pellet from the expression of CREB1-Q2-bZIP (pET-28b-*CREB1-Q2-bZIP*) without (lane1) and with (lane2) co-expression of OGT (pET-43. 1-*OGT*) in *E. coli* BL21-Gold(DE3) analyzed with (a) Coomassie-stained SDS-PAGE and (b) anti-O-GlcNAc MAb CTD 110. 6 western blot. HEWL, ovalbumin, and protein ladder are shown in the last three lanes, respectively. HEWL acts as a positive control for the western blot because the secondary antibody (DyLight 550-conjugate Goat anti-Mouse IgG) binds non-specifically to HEWL. Residual HEWL can be seen in lanes 1 and 2 because it was used to lyse the cells in these preparations. HEWL was not used to lyse cells in any of the other experiments presented in this manuscript because of its interference in western blot analysis. The western blot of Q2-bZIP (lanes 1 and 2) demonstrates that OGT is modifying the protein with O-GlcNAc, consistent with the analysis of Q2-bZIP in the lectin blot in Figure 2. The pellets were resolubilized in 6 M urea, centrifuged at 14,000 g at 4 °C for 30 min. Q2-bZIP was purified from the supernatant using a Bio-Rad BioLogic DuoFlow chromatography system with a 1 mL His-Select Ni-affinity gel column. After loading the supernatant, the column was washed with 10 mL of 6 M urea in buffer (50 mM sodium phosphate, 300 mM NaCl, pH 8. 0), 30 mL gradient (6 – 0 M urea in buffer) at a flow rate of 0. 1 mL/min, 10 mL of buffer, and 10 mL of 300 mM imidazole in buffer.

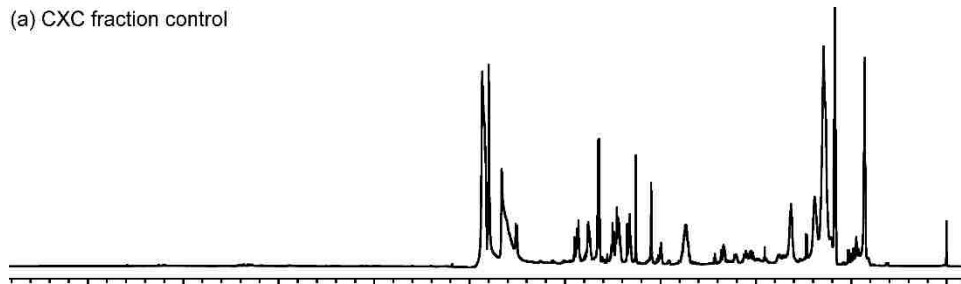


Supplementary Figure 2. The CXC fractions of cleared lysate from a culture of BL21-Gold(DE3) cells expressing empty pET-28b were analyzed using the MU-GlcNAc well-plate assay with the β -*N*-acetylglucosaminidase negative (-) and positive (+) controls. The active CXC fractions (#4-16) were combined and used to determine the enzyme kinetic parameters and activity towards the glycopeptide. Error bars represent the standard deviation of three replicate samples. The positive control intensity was truncated to better visualize the sample data.

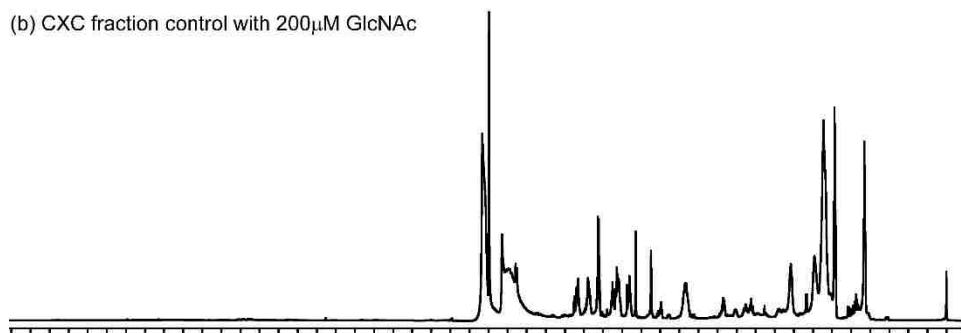


Supplementary Figure 3. Velocities of pNP-GlcNAc hydrolysis by *E. coli* β -*N*-acetylglucosaminidase, partially purified with CXC from BL21-Gold(DE3) cleared lysate, at varying pNP-GlcNAc concentrations. The velocities (nmoles of nitrophenolate formed per mL of fraction/min) were calculated from the absorbance at 405 nm of samples measured after 4, 9, and 14 min of reaction time. The curve was fitted to the Michaelis-Menten equation (1) using nonlinear regression analysis to give a K_m value of 0.35 mM.

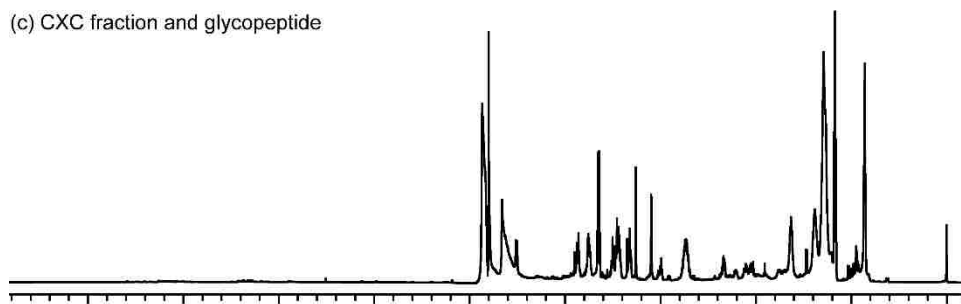
(a) CXC fraction control



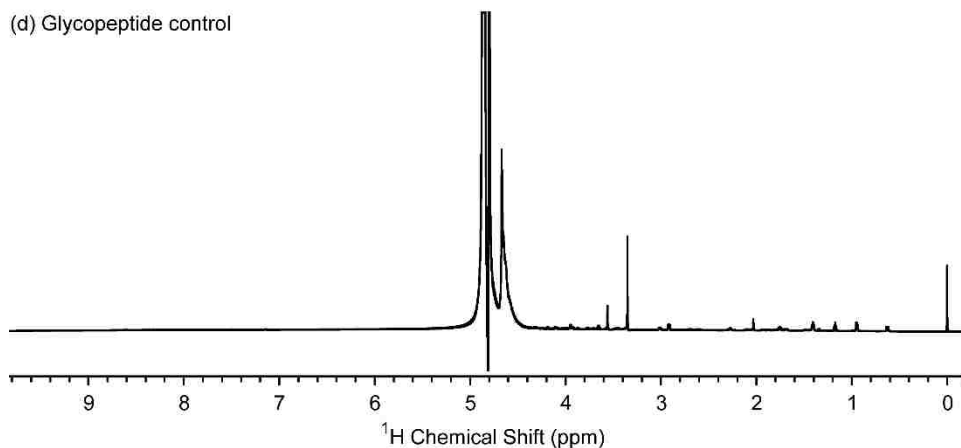
(b) CXC fraction control with 200 μM GlcNAc



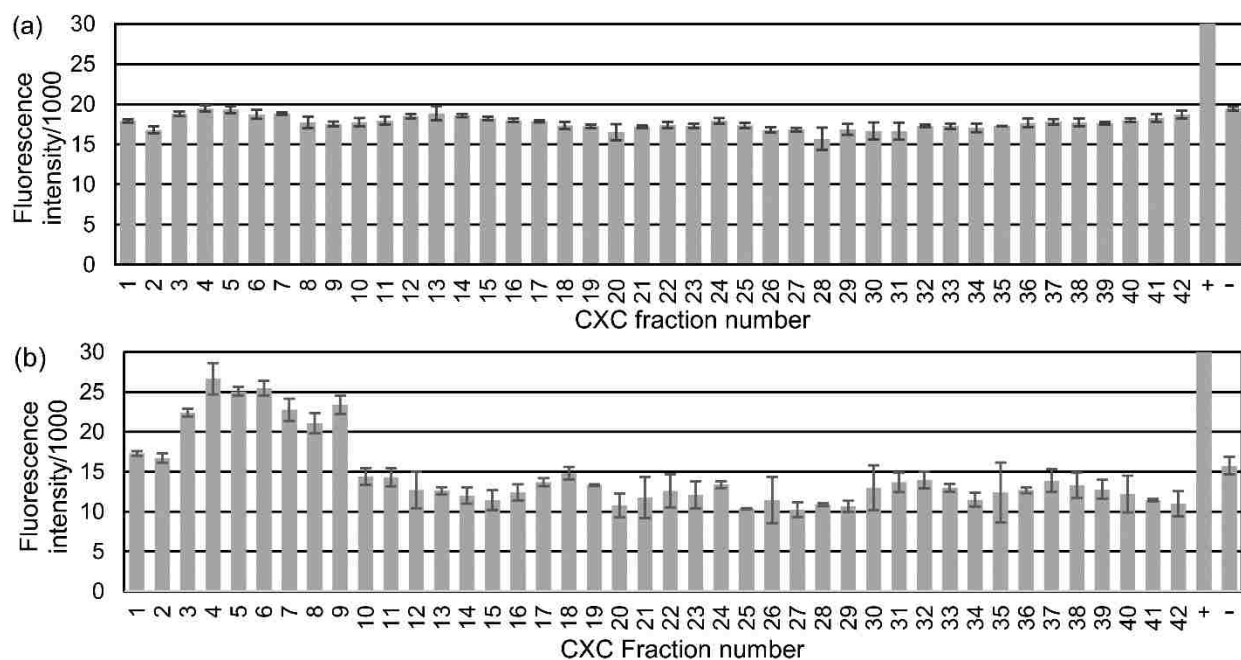
(c) CXC fraction and glycopeptide



(d) Glycopeptide control

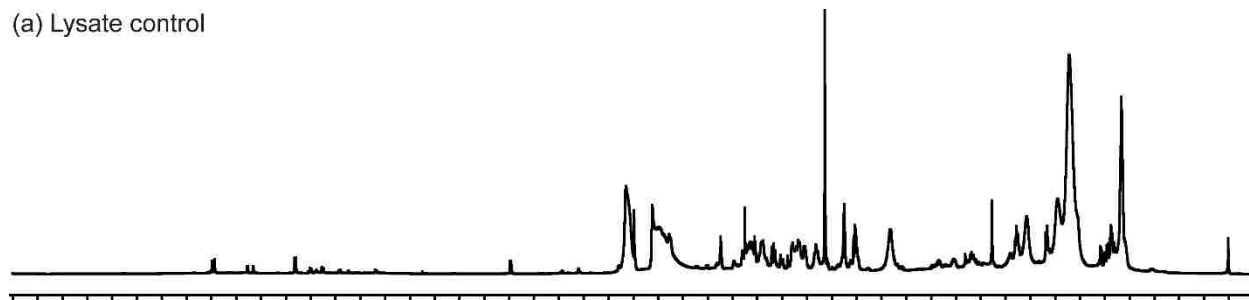


Supplementary Figure 4. Full ^1H NMR spectra of samples from testing the *E. coli* β -*N*-acetylglucosaminidase activity of CXC fractions towards the synthetic *O*-GlcNAc-modified peptide, SVES(β -*O*-GlcNAc)GSADAK-NH₂. Each spectrum was treated identically for direct comparison.

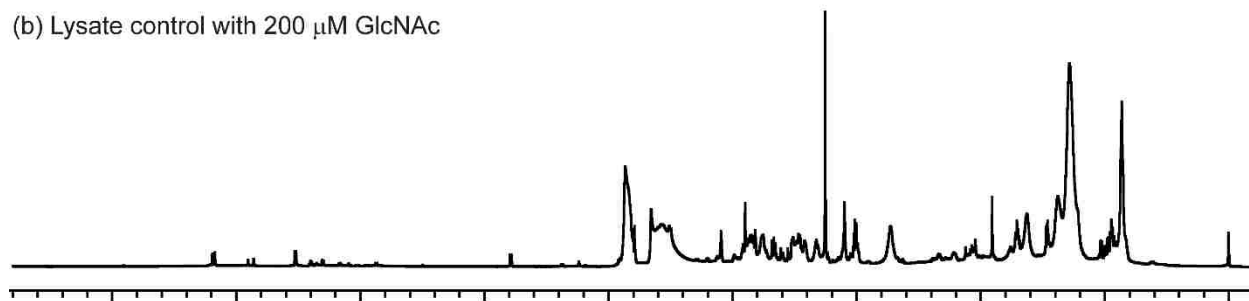


Supplementary Figure 5. Cleared lysate from cultures of (a) $\Delta nagZ$ K12 strain and (b) parent K12 strain were separated with CXC and analyzed with the MU-GlcNAc well-plate assay with the β -N-acetylglucosaminidase positive control (+) and negative control (-). Error bars represent the standard deviation of three replicate samples. The positive control intensity was truncated to better visualize the sample data.

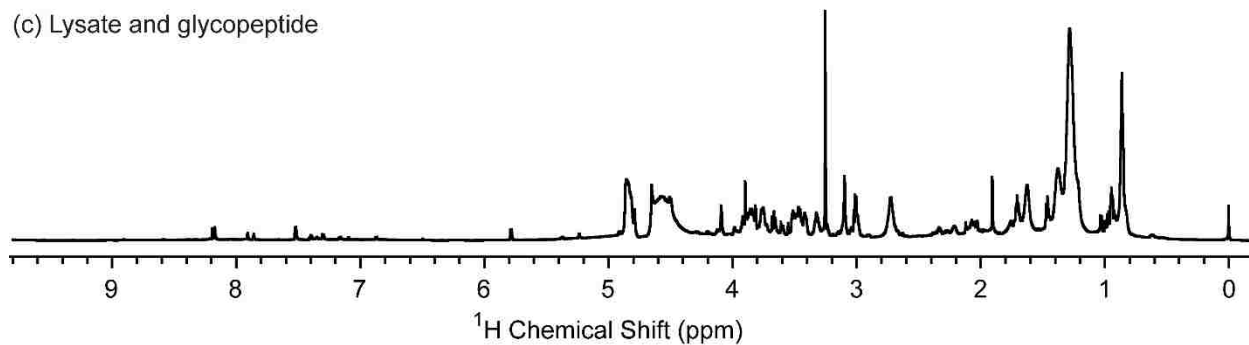
(a) Lysate control



(b) Lysate control with 200 μ M GlcNAc



(c) Lysate and glycopeptide



Supplementary Figure 6. Full ¹H NMR spectra of samples from testing for *E. coli* β -*N*-acetylglucosaminidase activity in Δ *nagZ* K12 lysate towards the synthetic *O*-GlcNAc-modified peptide, SVES(β -*O*-GlcNAc)GSADAK-NH₂. Each spectrum was treated identically for direct comparison.

Vita

Octavia Y Goodwin earned her B. S. in chemistry from Georgia Southern University in May of 2008. Octavia started her career as a research scientist in 2006. Under the supervision of Dr. Karen Welch, Octavia synthesized azo dyes. She also gained some computational molecular modeling experience while working on a glycomenttic ligand docking project. In January of 2009 Octavia joined Dr. Megan Macnaughtan's research group. While working for Dr. Macnaughtan, she bio-engineered a system to express near milligram quantities of O-GlcNAc-modified protein for structural analysis. Octavia will be receiving a Doctor of Philosophy in chemistry from Louisiana State University in December of 2014.