EXPANDING A TOOLKIT FOR CHARACTERIZATION OF TRANSMEMBRANE PROTEINS

A Thesis
Presented to the Faculty of the Graduate School
Of Cornell University
In Partial Fulfillment of the Requirements for the Degree of
Master of Science

by
Richard Paul Pampuro
August 2014

ABSTRACT

Though transmembrane proteins constitute 30% genes in most organisms and the target of nearly 70% of therapeutics, they remain poorly understood. Two hurdles in membrane protein characterization are a dearth of antibodies against transmembrane protein targets and difficulties in purification. Interaction of antibody fragments with transmembrane targets is inferred through a protein complementation assay (PCA), where reassembly of two reporter fragments upon binding lends antibiotic resistance. Genetic selection of binding members allows engineering against novel targets. Oligosaccharyltransferases (OSTs) are chosen as a model for purification and characterization. The central enzyme of N-linked glycosylation, OSTs covalently attach a glycan to an asparagine residue in a protein substrate. Techniques are demonstrated first with bacterial OSTs. Characterization of *Lieshmania major* OST STT3A is the first case of a eukaryotic OST working *in vitro.* The abilities to target antibodies against and purify membrane proteins together constitute tools to elucidate the workings of membrane proteins.

BIOGRAPHICAL SKETCH

Richard Pampuro grew up as a veritable nomad, traveling from navy base to navy base across the United States. He and his mother, father and sister, Amanda, lived in a total of 12 places before pausing long enough for him to complete high school in East Lyme, Connecticut. From there, he attended Worcester Polytechnic Institute (WPI) to study chemical engineering. Under the direction of Susan Zhou at WPI, Richard spent two years researching the fabrication of thin carbon films for application in neural engineering. He also spent his summers working in research and industry, conducting research in computational fluid dynamics with Anthony Dixon at WPI in 2007, as an intern at Alstom Power in 2008, and a research intern at United Technologies in 2009. While at WPI, Richard won the Class of 1879 Prize for excellence in humanities and a Department of Chemical Engineering outstanding Major Qualifying Project award. In August of 2009, Richard started his graduate studies at Cornell and joined the DeLisa lab in October of that year. He won a NSF GRAD Fellowship honorable mention in his first year at Cornell. In addition to his studies, Richard was active in outreach at Cornell, participating in the Chemical Engineering Graduate Student Association, volunteering for Expanding Your Horizons and the Chemical engineering WOMEN event, tutoring at the McCormic Center for Juvenile Violent Offenders and directing a research team with AguaClara designing open source water treatment plants for third world communities. With direction from Matt DeLisa, he researched techniques to improve understanding of transmembrane proteins for applications in science and medicine. In May of 2014, Richard left Cornell with an MS in Chemical and Biomolecular Engineering.

ACKNOLWEDGEMENTS

I would like to thank my advisor, Matthew DeLisa, whose advice and support has been essential to my graduate studies. My committee members, Jeff Varner and Susan Daniel, have also been helpful in guiding my work.

My family and friends also deserve thanks, especially my parents and sister, without whose support this would not be possible. There have been many who have helped me with issues scientific and beyond, Cresten, Jason, Linxiao, Madeline, and Tom, to name a few. Last, but not least, I'd like to thank M-P Robinson, whose courage and calm is the very reason I am alive to write this thesis.

EXECUTIVE SUMMARY

Characterization of transmembrane proteins is currently one of the largest open problems in protein science. Though they constitute about 30% of most proteomes by sequence and are the target of upwards of 70% of therapeutic drugs, detailed structural and mechanistic understanding is lacking for all but a few dozen membrane proteins [1]. The relative lack of membrane protein crystal structures (475 of nearly 82,000) highlights the dearth of understanding. The characterization of membrane proteins has remained an open problem due to difficulties producing transmembrane proteins in large scales and a lack of antibodies targeted against them [1, 2]. **This work attempts to improve membrane protein characterization in two distinct ways: a suite of techniques for purifying and characterizing activity of transmembrane proteins *in vitro* and a method for targeting antibodies against transmembrane proteins *in vivo.***

One of the largest barriers to transmembrane protein characterization is difficulty overexpressing and purifying target proteins on scales suitable for crystallography and activity studies. Optimization of purification strategies and development of activity assays for membrane proteins opens doors to understanding how membrane proteins work [3]. In this work, Oligosaccharyltransferases (OSTs) are the primary characterization target because they are large and difficult to overexpress, have a detectible activity and have valuable applications in biotechnology [4]. Relevant applications of OSTs are centered on glycosylation of recombinant proteins produced in *Escherichia coli* to improve activity, stability and targeting [5]. This thesis seeks to develop a purification strategy and *in vitro* activity assay to characterize various OSTs that are interesting and valuable in biotechnology.

Current techniques to engineer antibodies that bind to membrane proteins rely heavily on modifications to the target protein's sequence and protein purification [2, 6]. Commonly, a solubilized target protein is used to engineer antibodies *in vitro*, either

through infection of a model animal or an immobilization and capture scheme. Detergent solubilized membrane proteins do not always fold into their *in vitro* conformations, and even when they do, these mutants may not always fold in the same conformations as membrane-integrated wild-type structures [7, 8]. The method used in this project relies on an *in vivo* selection that uses full-length membrane proteins in their native conformations. As such, antibodies against transmembrane proteins engineered with this new method are much more likely to bind to physiologically relevant conformations of target proteins.

With a suite of techniques for obtaining large quantities of high purity, functional transmembrane proteins and the ability to engineer antibodies against them, exciting new doors are opened for characterization of these difficult targets. Combining this project's two foci promises exciting new options in chaperone assisted membrane protein crystallography and a leap forward in protein drug development [9-11]. Beyond OSTs in this project, techniques for overexpression and characterization of transmembrane proteins can be applied to poorly characterized but therapeutically valuable proteins. Chief among future targets are G-Protein Coupled Receptors, which represent the target for 36% of therapeutic drugs [12]. With this project's proof of concept in robust characterization and targeting of antibodies, a suite of questions that has eluded protein science will be within reach.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

Membrane proteins are currently one of the most difficult classes of proteins to study. Though they represent 30% of protein sequences and are the target of nearly 70% of therapeutic drugs, very few unique structures have been solved [6]. Membrane proteins function as transporters, receptors, enzymes, and structural anchors, often acting as a means through which cells are able to perceive the world outside the cytoplasm [1].

One of the major hurdles to characterization of membrane proteins is difficulty producing sufficiently large and pure samples for crystallization [6]. Membrane proteins represent about half of lipid membranes by mass, but a small fraction of the cell mass as a whole [1]. Even with an efficient purification protocol, membrane proteins are expressed at low levels and overexpression is often toxic to native cells [8]. In the case of eukaryotic membrane proteins, many often require post-translational modifications that are not possible in bacteria, further complicating expression [8].

Bacteria are often used as an expression platform for membrane proteins because overexpression is often impossible in native hosts [13]. Because of difficulty in overexpressing membrane proteins that are properly folded and inserted into the lipid bilayer, a few strategies have arisen to circumvent challenges. Often, substantial gains can be achieved from selection of appropriate host strain, plasmid, and fusion partner [13, 14]. Codon substitutions have emerged as a means to improve membrane protein expression—though not using the classical approach of merely enriching for the most commonly used codons [15]. With improved expression, membrane protein purification must follow, bringing its own challenges.

The standard way to purify membrane proteins involves use of expensive detergents to solubilize structures in detergent micelles. One major downside to detergents is the inability to predict their success, and the high probability that they will inhibit protein function [8]. Most protein crystal structures contain lipid or detergent molecules in the high-

resolution structures, in very much the same way that water molecules are often included in the structures of soluble proteins. Indeed, studies have shed some light on the importance of detergent selection based on apparent protein structure [7]. For the most part, detergent selection comes down to trying a number of options and happenstance [6].

Transmembrane proteins that transmit signals from extracellular stimuli to elicit an intracellular response are convenient drug targets because they allow therapeutic molecules to act outside the cell [16]. Though they are currently the targets of most therapeutics, identification of new ligands remains an open problem [12]. Improved tools for overexpression and purification hold promise to facilitate characterization of thus far elusive transmembrane protein targets [17]. Of particular interest in membrane protein drug targets are G-protein coupled receptors (GPCRs).

Within membrane proteins, the superfamily of GPCRs constitutes roughly half of membrane protein drug targets, or 36% of all drug targets [12]. However, GPCRs remain one of the most elusive proteins to characterize because they are very difficult to study from expression to purification to ligand studies [17]. Because discovering new ligands is challenging and characterizing cross talk is even more difficult, emerging protein therapeutics seek to create a range of highly specific affinity proteins against GPCRs [18].

Currently, protein therapeutics represents one of the largest and fastest growing areas of pharmaceutical research. One of the primary areas of protein therapeutics is monoclonal antibodies (mAbs), of which there are currently 34 approved in the US or Europe. Indeed, mAbs have seen growing FDA approval over the past few years, and single drug valuations in the multi-billion dollar range [19]. Beyond therapeutic targets that may drive funding for much of the research, antibodies have a number of other applications in research and medicine.

Traditional methods of engineering antibodies for specificity or affinity maturation rely heavily on having soluble targets [20]. Infecting mice [21], phage display [22], yeast display [23], MAD-TRAP [24], and FLI-TRAP [25] were all developed with soluble target

proteins in mind. Because these methods all require either surface immobilization or soluble expression of targets, membrane proteins would have to be detergent solubilized.

With traditional antibody screening and production models failing to satisfy increasing demand, a variety of technologies to engineer antibodies against specific targets are in development [23]. Realizing membrane proteins as valuable targets, efforts to engineer antibodies against membrane proteins have adapted recent antibody design technologies. Yeast surface display has been used to target antibodies against detergent solubilized membrane proteins [2] and phage display has produced scFvs against nanodisc solubilized membrane proteins [26]. Lipoparticles may also be used to solubilize membrane proteins for antibody screening [27].

Emerging technologies to engineer antibodies against membrane proteins rely on solubilizing membrane protein targets in detergents, nanodiscs, or lipoparticles [2, 26, 27]. These techniques, however, are unable to control the orientation of protein targets, meaning that it is impossible to control whether antibodies are targeted against the intracellular/cytosolic or extracellular/periplasmic faces of the protein. Moreover, because detergents may alter the structure of the protein, it is difficult to guarantee that antibodies will bind to the native, *in vivo* folding state of the protein [7]. Development of an *in vivo* technique to target antibodies against membrane proteins would provide a means to target antibodies against specific parts of the protein in its physiological folding state.

Beyond therapeutic applications, antibodies have other very important applications in biotechnology. Diagnostics, activity modulation, and even chaperone-assisted crystallography would all benefit from improved antibodies against membrane proteins [9-11]. Such important opportunities for characterization are further improved by the ability to obtain large quantities of high purity membrane proteins.

In addition to designing antibodies against membrane proteins, there is value in improving techniques to overexpress and purify membrane proteins. One particular group of membrane proteins, oligosaccharyltransferases (OSTs), have emerged as an interesting and

valuable target for purification. OSTs constitute the central enzyme of the N-linked glycosylation pathway, catalyzing the covalent bonding of a glycan molecule to an asparagine residue in a protein [4]. N-linked glycosylation is one of the most important post-translational modifications in eukaryotes, and is essential for the creation of therapeutic antibodies [28].

Following the functional transfer of *C. jejuni* N-glycosylation machinery to *E. coli* opens new doors for protein production in *E. coli* [4]. Differences between eukaryotic and prokaryotic glycosylation pathways and OSTs highlight shortcomings in current capabilities. In particular, limitations such as the bacterial D/E – X_1 – N – X_2 – S/T consensus sequence compared to eukaryotic N – X – S/T, different glycans, and folding states make bacterial OSTs less desirable for therapeutic applications [5]. Recent efforts have sought to identify alternative bacterial OSTs that recognize a eukaryotic-like consensus sequence [29]. In the DeLisa lab, efforts to synthesize eukaryotic sugars *in vivo* have shown that bacterial OSTs may work with eukaryotic sugars [30]. Folding state requirements, however, have not been explored in detail.

Shortcomings in characterization make OSTs a prime target for development of a scalable and robust purification strategy [3]. To date, studies have focused on a single enzyme and do not show a robust purification strategy to compare OSTs from different species [29, 31, 32]). Techniques to rapidly analyze activity of a range of OSTs *in vitro* allow selection of ideal candidates for application in production of glycosylated proteins. Here, a scalable and robust purification strategy strives to overcome existing shortcomings.

Together, the abilities to target antibodies against and purify large quantities of membrane proteins at high purity allow membrane proteins to be characterized in ways previously impossible. Each problem seeks to overcome some of today's most difficult questions in protein engineering [2, 6]. Indeed, even incremental advancements would be of great value where membrane proteins are concerned.

CHAPTER 2: SPLIT BETA-LACTAMASE PROTEIN COMPLEMENTATION ASSAY FOR SELECTION OF ANTIBODIES AGAINST TRANSMEMBRANE PROTEINS

Targeting antibodies against transmembrane proteins is one of the largest open problems in protein engineering. Difficulties in expression, unique considerations of orientation, and challenges of purification limit the number of techniques that may be applied to engineer antibodies against membrane proteins. Here, a TEM-1 β-lactamase (Bla) protein complementation assay (PCA) is employed to build a genetic selection system to detect the binding of antibodies to transmembrane proteins *in vivo*. Binding of antibody to target leads to reassembly of to Bla fragments, lending resistance to beta-lactam antibiotics. Membrane-anchored PCA is demonstrated to be able to select for high binding and produce a large dynamic range of binding and selection with the use of leucine zippers. Early results for single chain antibody fragments (scFvs) limited resolution discerning between high and non-binding members.

BACKGROUND AND INTRODUCTION

Protein-protein interactions play an important role in nearly every cellular process. Assembly of multiple gene products into a single functional complex allows cells to spatially localize enzyme complexes, detect specific macromolecules, and selectively transcribe DNA, among countless other reactions. Study of protein-protein interactions has given rise to a suite of techniques to select for assembly of macromolecules. With the ability to select for interactions, engineering of specific binding molecules like antibodies is possible.

Because of their high specificity and affinity, antibodies have been used as therapeutic, diagnostic, and analytical reagents for decades [9]. However, conventional methods to engineer antibodies against new targets rely heavily on use of animal models and complicated hybridoma production techniques [21]. More recent advances have seen protein-protein interaction assays like surface display [23, 24], yeast 2-hybrid [33], and

FLI-TRAP [25] based systems applied to engineer antibodies, but even these have substantial limitations when applied to difficult targets like membrane proteins [2]. Here, we use a protein complementation assay (PCA) to select for binding [34].

Surface display methods rely on tethering antibody fragments in bacteria or even whole antibodies in yeast and capturing cells on surface-immobilized antigens or detecting binding of fluorescence-linked antigens using flow cytometry [24, 35]. These methods rely on surface immobilization or soluble expression of target proteins, making targeting membrane proteins problematic. Recent advances have seen yeast surface display applied to engineer antibodies against detergent solubilized membrane proteins [2], nanodisc solubilized proteins [26] and lipoparticle solubilized proteins [27]. These solubilized proteins, however, are difficult to prepare, may not be in their native conformations, and do not allow targeting against specific domains (eg. Intracellular or intercellular) of a protein [7].

2-hybrid systems work by assembly of a transcription promoter with the interaction of two target proteins and work in bacteria and yeast. One target (bait) is fused to a protein that binds to an upstream activating sequence and the second (prey) is fused to an activating domain that recruits the RNA polymerase, where successful binding of bait to prey results in expression of a reporter protein [36]. 2-hybrid techniques have been applied to investigate membrane protein interactions, allowing structures to remain in their *in vivo* conformations [37, 38]. One downside to 2-hybrid systems is demanding precise localization of protein complexes near DNA, a problem in organisms with more than one compartment. Also, they do not necessarily provide a proportional binding strength-reporter signal correlation, making affinity maturation difficult [39]. While 2-hybrid systems work for some membrane proteins, genome scale studies have highlighted that these techniques are inadequate for most membrane protein targets [40].

Because of shortcomings in surface display and 2-hybrid methods, we have chosen a PCA readout for this study. PCA methods use the reassembly of two reporter fragments into

a single, active reporter to create a selectable phenotype. This format relies on a genetic fusion of two complimentary fragments to each the bait and prey molecules (two interacting partners, eg. antibody and antigen epitope) [41]. Strong interaction between bait and prey brings the reporter fragments into close proximity and reconstitutes an active enzyme, as shown in Figure 1A [42]. Previous efforts to engineer binding molecules against transmembrane targets with the use of a Bla-based PCA have relied on soluble truncations of the target protein [43] and a mouse dihydrofolate reducase (mDHFR) PCA has been used to select high binding scFvs from large libraries. With PCA methods, target proteins will be folded in their native, membrane-integrated conformations with minimal disruption to the protein's structure [44]. PCA methods can use a variety of reporter proteins, fluorescent proteins and antibiotic resistance enzymes most commonly [41]. The activity of fluorescent proteins [45] and mDHFR [46] require folding in the cytoplasm, a feature that precludes extracellular/periplasmic domains as targets. Keeping extracellular domains of GPCRs in mind as an ultimate target, split beta-lactamase is selected as the interaction reporter [42].

A variation of the original interaction reporters with two soluble partners [41] involves replacing a soluble bait molecule with a membrane protein [44]. The interaction of the prey molecule with part of the membrane-integrated prey reassembles the split beta-lactamase, lending antibiotic resistance to the bacteria. For the reporter, TEM-1 beta lactamse is split into two segments, alpha 1-197 and omega 198-288, as developed by Wehrman. To demonstrate reassembly of Bla fragments, FosLZ and JunLZ, two leucine zipper peptides with a very high (1nM) affinity are chosen [25, 34]. These peptides have a number of well characterized intermediate binding mutants and have been previously applied to demonstrate relationships between affinity and phenotype in other protein-protein interaction assays [25, 34]. Antibody-antigen interactions are demonstrated with scFv anti-GCN4 and GCN4 respectively. This pair of interacting partners is well characterized within the DeLisa lab [25].

Building off of a design of a membrane protein folding reporter from Skretas, a membrane protein of unknown function, yegH is used as the membrane anchoring protein to which a reporter fragment is fused [47]. Successful reassembly of reporter Bla is first demonstrated with the use of the FosLZ and JunLZ interaction system, with FosLZ being fused to the N-terminus of yegH and coexpressed with the soluble JunLZ construct. Negative controls for binding will mirror soluble interactions, using GCN4 as a non-binding control and well characterized JunLZ mutants for low/intermediate binding. Replacing FosLZ with GCN4 and coexpressing the scFv anti-GCN4 construct demonstrates the ability of scFvs to act as the prey molecule and reassemble active Bla in the periplasm with a membrane protein bait (Figure 2A). Here, a negative binding control comes in the form of scFv D10, a non-specific anti-GPD antibody fragment that does not bind to GCN4 [46].

MATERIALS AND METHODS

Strains and plasmids

A set of plasmids coding for split beta-lactamse fused to Fos and Jun gifted from the Blau lab [34] served as templates for split beta-lactamase leucine zipper fusions. Intermediate binding Fos and Jun point mutants were made through quick change PCR, while GCN4 was inserted in place of Jun in the fusion αBla-Jun, between EcoRI and KpnI. PCRs of scFv anti-GCN4 and scFv D10 (anti-GPD) were similarly ligated between EcoRI and KpnI to form αBla-antiGCN4 and αBla-D10. Fos-ωBla-yegH was made by replacing full length Beta-lactamase in pBad18CM ssDsbA-Bla-yegH, "pBadBla" (gifted by the Georgiou lab, (Skretas & Georgiou, 2010)). Skretas selected yegH as a membrane-integrated anchor for full length Bla because it expresses well, and yields a selectable phenotype. Full length Bla was removed from pBadBla between SacI and XbaI and a PCR of ssDsbA-Fos-ωBla was ligated into the same restriction sites. A signal sequence, ssDsbA, was included on the N-terminus of membrane protein containing fusions to ensure that the protein is targeted to the sec machinery and inserted into the periplasmic membrane. ssDsba-GCN4-ωBla-yegH

and ssDsbA-GPD-ωBla-yegH were cloned between SacI and KpnI from primer overlap extension. GPD-yegH and ωBla-yegH were cloned via primer overlap extension inserts, between SacI and XbaI for ssDsbA-GPD-yegH, or SacI and KpnI for ssDsbA-ωBla-yegH.

Cloning was conducted in strain DH5α. MC4100 was used for soluble split Bla experiments. Membrane-anchored split Bla experiments used Walker's C43 strain to improve membrane protein expression [14].

The NNK library of a 3 residue span of CDR3 of the parent scFv were created as in [25]. Briefly, the NNK randomized residues were created with the use of primers that contain randomized nucleotide sequences in the range of the 3 CDR3 residues of interest (GFA in wild type). In NNK randomization, N represents an equimolar mixture of all four nucleotides and K is a mixture of G and T, together NNK exclude stop codons (TAA, TAG, or TGA), but permits all amino acids to be encoded [25]. PCR products were ligated into the puc-based plasmid from Blau between EcoRI and KpnI sites. Samples of 10 µL and 100 µL were plated on Kanamycin selective media to estimate library size, and pooled 1ml samples of transformants were frozen for later selection. From the plated samples estimating library size, colonies were selected and sequenced to estimate library diversity.

Determination of protein-protein interaction *in vivo*

To determine interaction of split-Bla fused to leucine zippers, freshly transformed *E. coli* MC4100 were grown in 2 ml LB with 100 µg/ml kanamycin and 20 µg/ml chloramphenicol overnight. Cells were collected in the morning and normalized to the lowest OD. A series of 10 fold dilutions were made from 10^{0} to 10^{-5} and 5µl of each dilution spotted onto plates containing 100 µg/ml kanamycin and 20 µg/ml chloramphenicol and a range of concentrations of ampicillin between 0 and 100 µg/ml were supplemented. Plates were grown at 30°C overnight, then at 22°C until the control (0 µg/ml ampicillin) plate grows to the 10^{-5} dilution.

Plasmids containing membrane-anchored split-Bla constructs were transformed into C43 cells plated on LB agar containing 100 µg/ml kanamycin and 20 µg/ml chloramphenicol. Fresh transformats were grown overnight in 2ml LB with 100 µg/ml kanamycin and 20 µg/ml to saturation. 5ml cultures of LB with 100 µg/ml kanamycin and 20 µg/ml were sub cultured to an OD600 of 0.1 and incubated at 37°C for roughly 4 hours. Once cultures reached OD_{600} 0.5-0.7, protein production was induced with 0.1% arabinose and cultures were incubated at 30°C. The OD_{600} was monitored periodically until cell growth rate recovered from the addition of arabinose and returned to log growth phase (usually 4-6 hours), cells were harvested with a 4,500x spin for 10 minutes and normalized by OD_{600}, usually to the range of 0.7-1.0. As in the case of soluble split Bla, 10 fold dilutions were prepared and 5ul spots were plated. Selective plates were grown at 22°C for 48 hours on LB agar containing 100 µg/ml kanamycin and 20 µg/ml and varying concentrations of ampicillin.

RESULTS

Interaction of solubly expressed FosLZ and JunLZ

To confirm that the split Bla genetic based interaction reporter demonstrates a proportional binding strength to antibiotic resistance phenotype, we analyzed fusions of c-Fos and c-Jun (FosLZ and JunLZ) leucine zippers to Bla fragments alpha and omega (αBla-JunLZ and FosLZ-ωBla. Peptide affinity was modulated with a series of point mutations to specific leucine residues, allowing a range of interaction strengths to be analyzed. All fusion proteins are expressed in the bacterial periplasm, where reassembly of Bla yields antibiotic resistance. Higher affinity wild type variants FosLZ and JunLZ were expected to yield the highest antibiotic resistance. FosLZ mutated to FosL2V and JunLZ to variants JunL3V and JunL4V reduce binding strength compared to wild type, and were expected to produce an intermediate resistance phenotype. A negative, non-binding control was made by replacing JunLZ with the peptide GCN4, which does not interact with FosLZ.

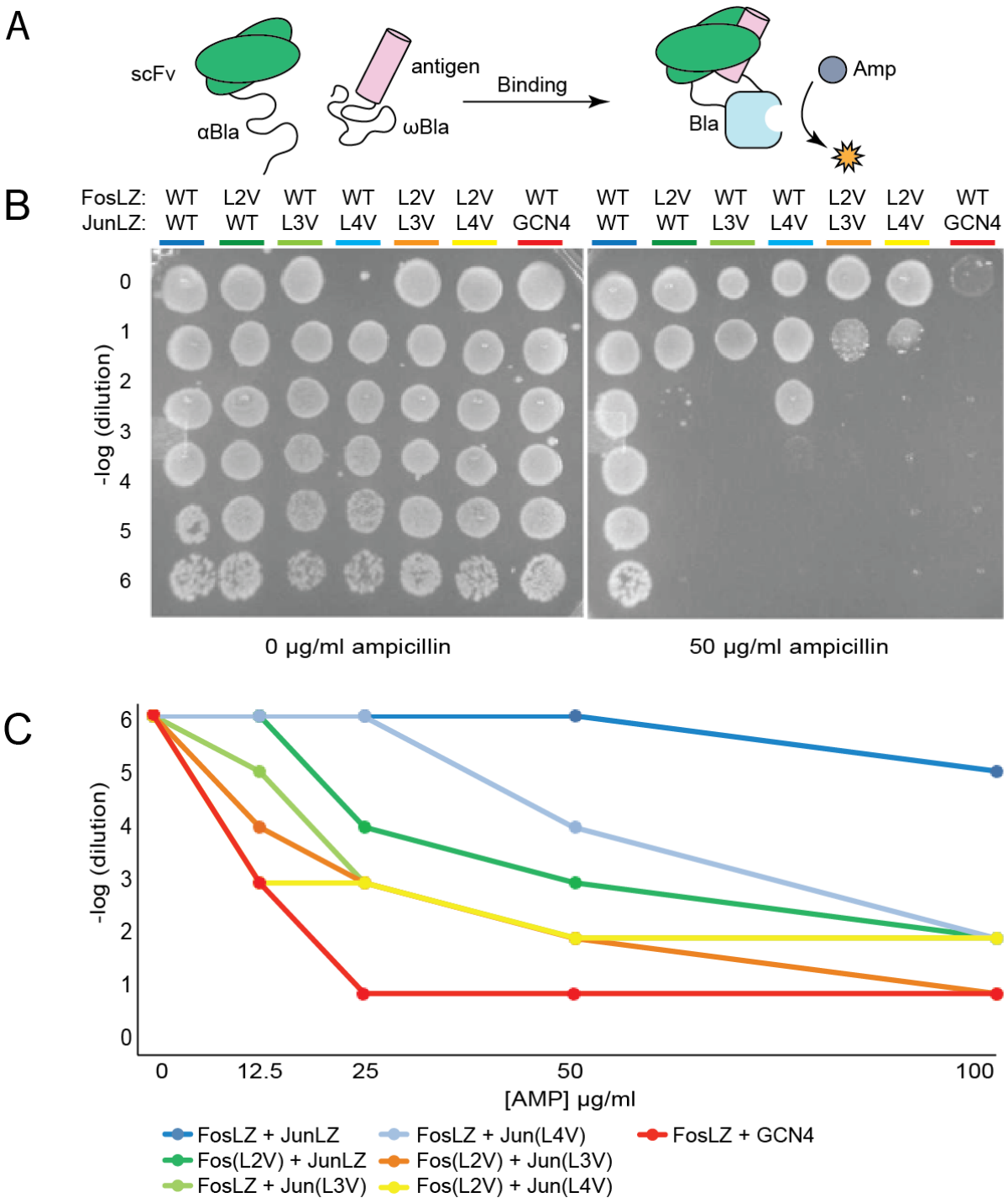


Figure 1:(A) Schematic of soluble beta-lactamse PCA. Association of the bait and prey molecules (green scFv and pink antigen here) brings the N-terminal bla fragment (αBla) and the C-terminal bla fragment (ωBla) into close proximity to reconstitute the active bla enzyme (in blue). In this system, reassembly of reporter enzyme beta-lactamse yields

resistance to ampicillin when localized properly in the bacterial periplasm.(B) Phenotypic selection of FosLZ-ωBla and αBla-JunLZ variants expressed in MC4100 and plated on control 0ug/ml ampicillin (left) and selective 50μg/ml ampicillin (right). Wild type variants are the strongest interacting, and show growth to lower dilutions. Modified leucine knockouts Fos(L2V),Jun(L3V), and Jun(L4V) and non-specific GCN4 peptide show lower resistance to antibiotic. A spectrum of interaction strength yields a proportional response of antibiotic resistance, demonstrating a high dynamic range for binding strength and resistance phenotype. (C) "Kill curve" of FosLZ and JunLZ variants over a range of ampicillin concentrations. Strong interactions (FosLZ/JunLZ and FosLZ/JunL3V) show the highest resistance over a range of ampicillin concentrations, with strong growth at high selective pressure (100ug/ml ampicillin). Weak interactions (combinations of Fos(L2V) and Jun(L3V)/Jun(L4V)), show intermediate levels of resistance over a range of selective pressures, especially at 50 μg/ml ampicillin, where there is a clear spectrum of binding and resistance phenotype. Non-interacting control (FosLZ/GCN4) show very poor growth, with no dilutions growing above 25 μ g/ml ampicillin. This interaction assay is able to successfully discern between binding and non-binding partners, while also providing a high dynamic range to discern between intermediate binding strength partners.

Wild type FosLZ/JunLZ interact with an K_D of 0.99 ± 0.30 nmol [48]. As expected, the positive interaction FosLZ/JunLZ has an antibiotic resistance far greater than the negative GCN4/FosLZ, growing to a 10^5 greater dilution at 50 μg/ml ampicillin (Figure 1B, 1C). In addition to being able to discriminate between binding and non-binding, intermediate binding point mutants show a dynamic range of interaction. Relative resistance of mid affinity FosLZ and JunLZ mutants (FosL2V/JunLZ, FosLZ/JunL3V, FosLZ/JunL4V) are in agreement with previous predictions of binding strength (Figure 1B) [25]. Though not characterized in this work, one well characterized Jun mutant Jun(V36E) has an affinity of 0.90 ± 0.13 μM [49]. FosL2V/JunL3V is observed to have a significantly weaker resistance

compared to FosL2V/JunlZ, while still between high and non-binding controls, satisfying expectations set in Waraho. Also as predicted, combinations of mutants (FosL2V/JunL3V, FosL2V/JunL4V) have lower resistance than mutants paired with wild type zippers.

The kill curve in Figure 1C illustrates the spectrum of resistance across several levels of selective pressure for soluble FosLZ/JunLZ fused split Bla. Strong correlation of predicted binding strength to resistance phenotype is evident in the 25 μg/ml and 50 μg/ml ampicillin pressures, with differences between the groups spaced out over several orders of magnitude of resistance. With an estimated difference of nearly three orders of magnitude in binding strength between wild type and mid-binding mutants [48], a quantitative correlation to resistance phenotype is observed. With dilution differences of 10^{-2}, 10^{-3}, and 10^{-4} for pairings FosLZ/JunL4V, FosL2V/JunLZ, FosLZ/JunL3V respectively, a strong correlation of binding strength and resistance phenotype is demonstrated.

Membrane-anchored Leucine Zippers

Fusing the model membrane protein yegH [47] to the C-terminus of FosLZ-ωBla anchors the FosLZ-ωBla fusion to the perisplasmic face of the inner membrane. This fusion serves to model a binding region of the membrane protein of interest and serves to demonstrate that 1) reassembly of Bla anchored to the membrane is functional and 2) preservation of the correlation between binding affinity and antibiotic resistance phenotype.

The same JunLZ mutant used to demonstrate affinity-resistance relationship in soluble split Bla is applied to the membrane-anchored split Bla. As shown in Figure 2B, the high binding FosLZ-yegH/JunLZ grows to high dilutions (10^{-6}) well, even at relatively high, 100 μg/ml, antibiotic concentrations. Mid-binding (FosLZ-yegH/Jun(L3V)) and non-binding (FosLZ-yegH/GCN4) illustrate range of the binding-resistance phenotype relationship.

The high dynamic range of 10^{3} dilutions resistance between groups holds promise for engineering improved binding of target molecules. High background noise in the form of relative resistance of non-binding cases shown in growth at 10^{-3} at 50 μg/ml membrane

linked compared to 10^{-1} μg/ml for soluble split Bla means that higher selective pressures, 100 μg/ml and up, are needed to discern between high- and non-binding cases. Selective pressures are increased to very high levels, 500 μg/ml (Figure 2B) to illustrate that the phenotype is a result of binding and not a false positive due to a recombination event, which would yield high levels of resistance up to 1,000 μg/ml ampicillin. Poor growth for positive binding cases (10^{-3} dilution) and non/mid-binding cases (10^{-1} dilution) suggest that resistance is due to binding assembly.

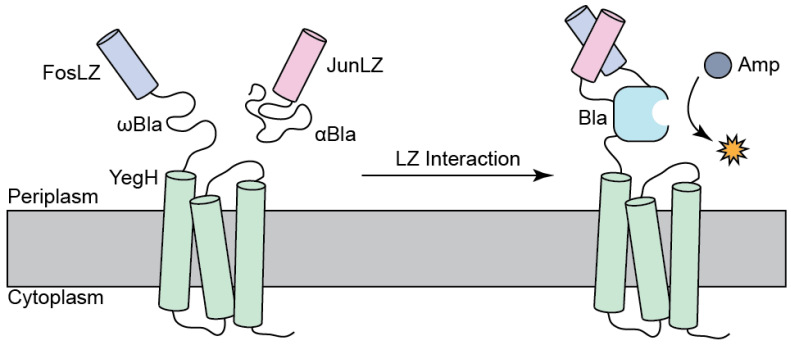
Membrane-anchored scFvs-Antigen Pairs

JunLZ is replaced with scFv anti-GCN4 (αGCN4) or scFv D10 to form αBla-αGCN4 and αBla-D10, while FosLZ in Fos-ωBla-yegH is replaced with the corresponding antigens GCN4 or GPD to form GCN4-ωBla-yegH and GPD-ωBla-yegH. Antibody fragment-αBla fusions are coexpressed with each of the membrane-anchored antigens and plated on selective media. GCN4-ωBla-yegH with αBla-αGCN4 and GPD-ωBla-yegH with αBla-D10 represent positive interacting partners, while non-specific scFv pairings GCN4-ωBla-yegH with αBla-D10 and GPD-ωBla-yegH with αBla-αGCN4. Pairing of membrane-anchored antigens with a non-interacting, non-scFv peptide αBla-JunLZ controls for spontaneous reassembly of Bla fragments. Further controls include GPD-yegH, lacking the ωBla fragment and offering no antibiotic resistance, controlling for intrinisic antibiotic resistance, while GPD-yegH lacks an antigen and controls for non-specific scFv binding.

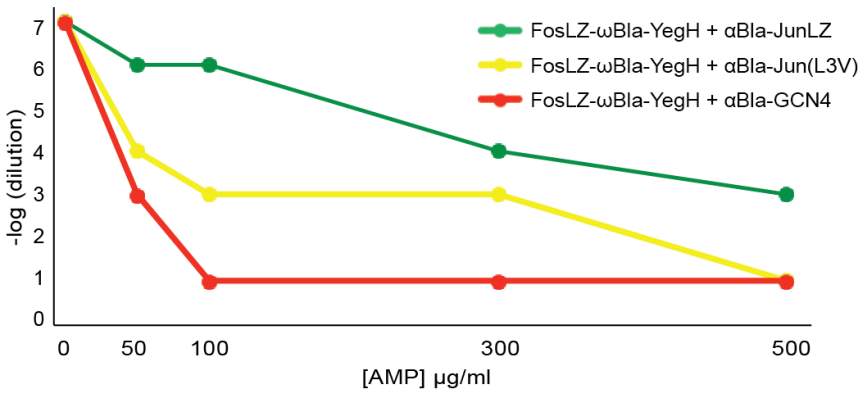
Shown in Figure 2C, pairing of specific scFvs with their corresponding antigens (GCN4-ωBla-yegH with αBla-αGCN4 and GPD-ωBla-yegH with αBla-D10) yields antibiotic resistance. However, non-specific pairings (GCN4-ωBla-yegH with αBla-D10 and GPD-ωBla-yegH with αBla-αGCN4) also grow to high antibiotic concentrations, with a resolution of only 10^{2}. GPD-ωBla-yegH tends to grow at higher antibiotic concentrations than GCN4-ωBla-yegH (10^{-6} for GPD vs 10^{-3} for GCN4 at 50 μg/ml ampicillin) regardless of scFv pairing. To determine whether this growth is due to inherent resistance or spontaneous reassembly of

Bla, a negative control for each is employed. The extent of growth on negative controls lacking both Bla fragments (GCN4-yegH with αBla-αGCN4 and GPD-yegH with αBla-D10) will show the inherent resistance of cells not harboring a complete copy of Bla. Non-binding control ωBla-yegH/αD10 growth shows to what extent reassembly of beta-lactamse is occurring in a binding-independent mechanism.

A



B



C

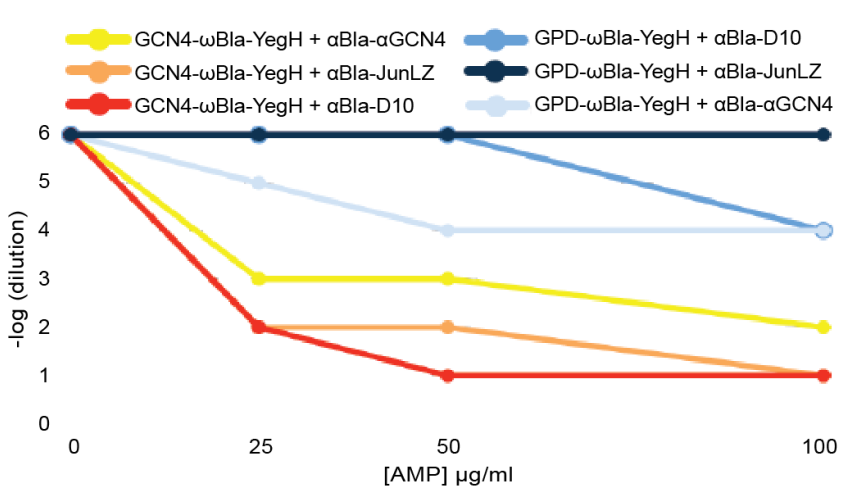


Figure 2: (A) Schematic of membrane-anchored split Bla setup. FosLZ-ωBla is fused to the N-terminus of a membrane protein to form FosLZ-ωBla-yegH, while the complimentary αBla-JunLZ is expressed solubly in the periplasm. Binding of FosLZ to JunLZ reassembles Bla anchored to the periplasmic face of the inner membrane and results in antibiotic resistance. (B) Kill curve of membrane-anchored split Bla with leucine zippers. As in the case of soluble split Bla, it is expected that high affinity wild-type partners yield higher antibiotic resistance. FosLZ-ωBla-yegH paired with αBla-JunLZ produces resistances on par with that of soluble split bla, while the non-binding αBla-GCN4/FosLZ-ωBla-yegH exhibits minimal resistance. Intermediate affinity Jun(L3V) lends resistance above non-binding and lower than high binding variants, demonstrating a high dynamic range of selection. (C) Kill curve of membrane-anchored split bla and scFv-antigen pairs. Two sets of scFv-antigen pairs are used: GCN4-ωBla-yegH with αBla-αGCN4 and GPD-ωBla-yegH with αBla-D10. Negative controls are formed by pairing with a non-specific scFv, substituting scFv αGCN4 or D10 for one another, as well as JunLZ as a non-specific peptide. In both cases, correct pairing (GCN4-ωBla-yegH with αBla-αGCN4 and GPD-ωBla-yegH with αBla-D10) offer the highest level of resistance compared to non-specific scFv (GCN4-ωBla-yegH with αBla-D10 and GPD-ωBla-yegH with αBla-αGCN4) and non-specific peptide (GCN4-ωBla-yegH with αBla-JunLZ and GPD-ωBla-yegH with αBla-JunLZ). Resolution between specific and non-specific scFv pairings is only a dilution of 10^2, and non-specific peptide αBla-JunLZ behaves similarly to αBla-D10 in both cases.

Discussion

Split beta-lactamase offers a selection platform for protein-protein interactions that is suitable for a variety of interactions and targets. Its ability to correlate resistance phenotype with binding strength proves valuable for improving binding of partner molecules, and obviates the issue of signal amplification in 2-hybrid systems [39]. Fragments with nanomolar affinities result in very strong growth, and a range of previously

characterized mutants [25] demonstrate a clear correlation between binding strength and antibiotic resistance phenotype. Capabilities of split Bla may be in the same range as FLI-TRAP, which has already been applied to select and mature antibodies [25].

Anchoring one binding fragment to the inner membrane does not inhibit reassembly of active beta-lactamase. Antibiotic resistances observed are in the same range as those of the soluble split Bla and the binding strength-antibiotic resistance phenotype correlation is preserved. Enhanced binding can be selected for because of the binding-strength resistance phenotype correlation present in both soluble and membrane-anchored formats, offering advantages over n-hybrid systems. Moreover, in allowing selection of binding to natively folded membrane proteins, the Bla-based PCA offers advantages over yeast surface display methods that rely on detergent solubilized targets [2]. In allowing binding strength-resistance phenotype correlation and natively folded structures, split Bla offers the ability to engineer antibodies against difficult membrane protein targets.

Non-binding reassembly of Bla or inherent resistance of C43 cells lends a high level of antibiotic resistance to non-binding scFv controls in the case of membrane-anchored antigens. In the event of non-binding reassembly of Bla, linking sequences between binding domains and Bla fragments can be altered to make spontaneous assembly more difficult. Inherent resistance of C43 cells under selection conditions demonstrated through non-assembling controls must be remediated by increasing resistance of the positive binding antibody-antigen pairs. Resistance of positive binding groups can be increased by tuning expression levels of antibodies and membrane-linked antigens, altering linker sequences, or selecting a different reporter.

At present, antibody-antigen pairs fail to discern between high and low binding pairs at a sufficiently high resolution. It is unclear why membrane-anchored GPD (GPD-ωBla-yegH) outperforms membrane-anchored GCN4 (GCN4-ωBla-yegH) under all conditions, it may be due to differences in expression level or reduced cell stress. The low dynamic range for each membrane-anchored antigen under these conditions does not allow enough

difference between binding and non-binding to allow for affinity maturation. For the NNK library, a resolution $O(10^4)$ is necessary, while the full random library would be best under conditions that offer even higher resolutions $O(10^6)$, the current $O(10^2)$ resolution is insufficient for use with these libraries.

CHAPTER 3: EXPRESSION, PURIFICATION, AND *IN VITRO* CHARACTERIZATION OF OLIGOSACCHARYLTRANSFERASES IN *ESCHERICHIA COLI*

Oligosaccharyltransferases (OSTs) are the core enzyme in asparagine-linked (N-linked) glycosylation, an essential process responsible for protein targeting and function in eukaryotes. OSTs function by covalently linking a glycan to an asparagine residue in a protein substrate. Originally thought to be exclusive to archaea and eukarya, discovery and functional transfer to *E. coli* of the *Campylobacter* N-linked glycosylation locus has opened doors for application of glycosylation in bacterial expression platforms. Characterization of these proteins is essential to application in biotechnology. Here, a range of OSTs are purified and characterized for substrate sequence and folding state specificity and activity *in vitro*. OSTs from *C. jejuni,* and *Desulfovibrio vulgaris* represent two bacterial OSTs. The *Leishmania major* STT3A is the first report of a eukaryotic OST being purified and functioning *in vitro*.

BACKGROUND AND INTRODUCTION

An important feature in protein folding, targeting, and cell-cell signaling, N-linked glycosylation is an essential process in eukaryotes. This post-translational modification is of key interest in biotechnology and medicine, particularly with regard to production of human-like proteins in bacteria [5]. Recent discovery of N-linked glycosylation machinery in bacteria and its functional transfer to the model organism *E. coli* has enabled the study and application of N-linked glycosylation in ways unattainable in eukaryotic systems [4]. In all domains of life, N-linked glycosylation follows the same reaction scheme where a glycan is assembled on a lipid carrier and then transferred to a target protein [5]. The central reaction where a glycan is moved from a lipid carrier onto the asparagine residue is catalyzed by the OST. Characterization of this enzyme is central to the understanding and application of N-linked glycosylation [4, 31].

Modification of an asparagine residue occurs in context of a N-X-(S/T) motif in Eukaryotes and Archeae and in an expanded (D/E)-X_1-N-X_2-(S/T) in some bacteria, where X is any amino acid other than proline (Eichler, 2013). The consensus sequence is necessary, but insufficient alone, to ensure glycosylation [50]. While there have been efforts to characterize the precise sequon requirements, separation of steric context and sequence identity can be difficult. Most eukaryotic glycosylation is performed cotranslocationally as the peptide chain passes through the sec translocon in the endoplasmic reticulum [51], and bacterial glyscosylation has been shown to occur post-translationally [52, 53]. Evidence from mammalian systems hints at two separate OST variants, where one glycosylates the unstructured N-terminus of a translating nascent protein chain and the second appears to modify the chain passing through the sec-translocon [54]. OSTs from the *Campylobacter* genus appear glycosylate both post-translocationally on substrates like AcrA *in vitro [55]* and co-translationally *in vivo [52]*. Peptide substrates offer a steric context-free method to evaluate OST activity on target sequons [56]. A mostly unstructured glycosylation domain attached to a soluble anchor protein [30] offers a means to measure glycosylation activity on weakly structured domains intended emulate a translocating nascent chain *in vitro*, and structured AcrA allows evaluation of glycosylation of fully folded substrates [32, 52].

With the publication of the crystal structure of PglB from *Campylobacter lari*, a detailed structural knowledge of *C. jejuni* PglB and its interaction with peptide substrates are well characterized [31]. Further studies have shown the role of the -2 D/E is necessary for binding of the peptide substrate [56]. Other bacterial OSTs, however, remain poorly characterized. Recent *in vivo* evidence from the DeLisa laboratory suggests that certain OSTs from the *Desulfovibrio* genus may not require the acidic D/E in the -2 position of the sequon [29]. To examine the OST outside of context of the glycan assembly pathway and in the absence of steric hindrance of the peptide substrate, the OST is purified and characterized *in vitro*. Additionally, *in vitro* reaction allows precise control of peptide and LLO substrates in ways not possible *in vivo* [55].

Beyond bacterial OSTs, eukaryotic OSTs offer a number of valuable features, primarily focused on compatibility with eukaryotic substrates and glycans [5]. However, these OSTs are likely incompatible with bacterial expression systems and translocation machinery necessary for cotranslocational glycosylation [57, 58]. Moreover, the majority of eukaryotic OSTs function within the context of an OST complex, which is responsible for keeping the translocating nascent protein chain unfolded and accessible while it is glycosylated [59]. In addition to requirements of the protein substrate, the limitations of LLO substrates are not fully understood [60]. Evidence has indicated that a group of STT3 catalytic core subunits from *Leishmania major* are able to functionally replace the OST complex in *Saccharomyces cerevisae [61].* Very little is known about the function of these STT3 catalytic core subunits, and bacteria represent a context-free platform to study these valuable enzymes.

Evidence from *in vivo* and *in vitro* experiments has contributed to a detailed understanding of *C. jejuni* and *C. lari* sequence requirements, as well as a rough understanding of steric context requirements [52, 56, 58, 62]. Detailed investigation of folding context requires substrates to model each of three folding states: highly structured folded protein, loosely structured nascent chain-like substrates, and unstructured, unhindered domains. Highly structured protein comes in the form of AcrA, a native substrate of *C. jejuni* PglB, which folds into a coherent structure that can be purified and used as a substrate *in vitro* [32, 55]. Loosely structured regions can be attached to a soluble anchor protein like scFv13-R4, approximating a translating nascent chain [52]. Finally, fluorescently labeled peptides offer minimal structure and allow investigation of pure sequence requirements in the absence of steric hindrance [55, 56].

While the *C. jejuni* PglB-peptide substrate interactions are well characterized, lipid liked oligosaccharide (LLO)-PglB interactions are only emerging as topic of investigation [31, 55, 56]. A rapid *in vitro* method to characterize OST activity opens new doors for high-throughput evaluation of reaction conditions and allows for artificially high concentrations of

LLOs not possible *in vivo* [55]. Previous studies have looked at the specificity of *C. jejuni* pgl pathway for polyisoprenoid specificity [50], but used membrane extracts rather than purified OSTs. More recent studies have used purified PglB, but investigations have focused on *C. jejuni*-like LLOs [32, 55]. Ideally, eukaryotic-like LLOs would be used to add therapeutically desirable sugars to proteins [30] without complicated purification and *in vitro* glycan elaboration [63].

MATERIALS AND METHODS

Membrane protein expression

OSTs with a 10x His-tag are cloned into pSN18 [62] and transformed into C43 cells. An overnight cultures were subcultured 1:500 into 2L of LB in a 4L Erlenmeyer flask and grown at 37°C until OD_{600} reached 0.5-0.7, then incubated at 16°C and, after a brief temperature equilibration, induced with 0.1% arabinose. Cells were grown overnight (16-18 hours) and harvested at 10,000xg for 10 minutes.

Harvested cells were resuspended in 60ml resuspension buffer (50mM tris, 1mM EDTA), and lysed with 3 passes in an Avastin microfluidizer at 10,000psi [64]. Unlysed cells were collected with a 10,000xg spin for 10 minutes. Cleared lysate supernatant was spun at 140,000g in a Beckman SW28 rotor for 90 minutes to pellet the membrane fraction. Membrane fraction pellets were resuspended in 10ml solubilization buffer (20mM, 1mM EDTA, 100mM NaCl, 1% cymal-7 W/V) and agitated at 4°C for 2 hours before a 10,000xg spin for 10 minutes. The resulting pellet was the insolubile membrane fraction and the supernatant was the solubilized membrane fraction.

Membrane protein purification

Purification methods are adapted from [32, 64]. Solubilized membrane fractions were mixed 1:1 with 10ml binding buffer (20mM tris, 300mM NaCl, 50mM imidazole, 0.1% cymal-7 W/V) to dilute EDTA in the solubilization buffer before being run over a nickel

immobilized metal affinity chromatography (IMAC) column. Purification was performed on a GE AKTA10 (GE Healthcare) using a 1ml Hi-Trap IMAC column (GE Healthcare) at 4°C, unless otherwise noted. 20ml of diluted membrane fraction mixture was flowed over the IMAC column at a rate of 0.5 ml/min, followed by a rinse with 6ml of rinsing buffer (20mM tris, 300mM NaCl, 100mM imidazole (Sigma), 0.1% cymal-7 W/V (Anatrace)) at 0.5 ml/min. Purified proteins were eluted in 5ml elution buffer (20mM tris, 300mM NaCl, 275mM imidazole, 0.1% cymal-7 W/V) at 0.5 ml/min. Elution fractions were desalted into storage buffer (50mM tris, 150mM NaCl, 0.05% cymal-7 W/V) with 2-5 ml GE HiTrap columns at a flowrate of 5 ml/min. Desalted samples were concentrated above a 3kDa mass cutoff spin column to a target concentration of 2.0 mg/mL (Amicron Ultra). Protein concentrations were measured using a bicinchronic acid (BCA) assay (sigma). The BCA assay was chosen in preference to the Bradford because BCA is compatible with cymal-7. For storing samples in excess of 2 days, 10% glycerol is added to the final concentrated sample. Final samples are stored at 4°C.

LLO extraction

LLO extraction was adapted from [32]. Briefly, *E. coli* SCM6 cells transformed with pACYC-*pglΔB* were grown to saturation overnight. Cells were sub-cultured at 1:500 into 2L of LB in a 4L flask and grown at 30°C for 18 hours. Cells were centrifuged at 10,000g for 10min at 4°C and the pellets lyophilized to dryness at high vacuum overnight. All subsequent steps are conducted at room temperature unless otherwise noted, and with PTFE or glass tubes and pipettes. Lyophilized pellets were resuspended in 30ml of 10:20:3 $CHCl_3$:MeOH:H_2O and spun for 30min at 3,000g. The supernatant was collected and any remaining pellet residue removed after settling in a separatory funnel. The chloroform and methanol from the upper phase in the separatory funnel was collected and then evaporated in a rotary evaporator (Buchi) at 37°C until the resulting sample was a vicious yellow liquid. Trace solvent was evaporated in a vacufuge for several (8+) hours (Thermo) and the

resulting residue was lyophilized to dryness at high vacuum overnight. The lyophilized powder was weighted, then resuspended in reaction buffer (10mM HEPES (4-(2-hydroxylethyl)-1-piperazineethanesufonic acid), pH 7.5, 1mM MnCl2, 0.1% cymal-7 W/V) to a concentration of 0.12g/ml. Further dilutions before conducting reactions are performed in reaction buffer.

In the case of *S. cerevisiae* LLOs, cells from a 40L fermentation were collected and washed three times in PBS between spins at 300g for 5 min. About 10g cell paste was resuspended in 20 ml methanol and cells are lysed on a microfluidizer (Avestin) with three passes at 16,000 psi. Methanol was evaporated and the preparation proceeded as in the prokaryotic case after the first lyophylization step.

in vitro glycosylation reactions

Glycosylation reactions of purified acceptor proteins (PglB or R4-GT) were conducted in 50 µl volumes with 2 µg purified PglB, 1 µl of extracted LLOs (48 µg/ml) and 5µg of purified acceptor protein. Reactions were incubated for 12 hours at 30°C then heated to 95°C for 20 min to denature all proteins. 50 µl of 2x SDS PAGE loading buffer was added and samples were boiled for 20 min before loading 20 µl on a 10% SDS PAGE gel (Biorad) and run at 140V for 65 minutes. Gels were transferred to an immobilon-P PVDF membrane at 0.08 A per gel for 60 minutes. Transferred membranes were blocked in 5% non-fat milk for at least 1 hour before being probed with appropriate antibodies.

Peptide substrates were purchased from Thermo-Fisher peptide synthesis and were stored in PBS and frozen at -80°C until use. Two peptides were used: tamra-N-GDQNATAF-C and tamra-N-GAQNATAF-C. The first mimics the optimum *C. jejuni* PglB sequon [56], while the second models the eukaryotic glycosylation sequence. Reactions were conducted in 10 µl volumes with 1 µg purified OST, 0.2 µl of extracted LLOs (48 µg/ml) and 280 pmol peptide in a buffer of 10 mM HEPES, pH 7.5, 10 mM MnCl2 and 0.1% cymal-7 W/V. Reactions were run for at least 90 min at 30°C, followed by 20 min at 95°C to denature

OSTs. 2 µl of 6x SDS PAGE loading buffer is added and samples are boiled for 20 min. 3 µl of the boiled mixture is loaded on a Tricine SDS PAGE gel optimized for peptide resolution.

Tricine SDS PAGE gels are hand cast as discussed in [65]. Gels were cast in a BioRad Mini-PROTEAN 3 multicast chamber (BioRad) between glass plates with a gel thickness of 0.5 mm. Gels had a 16% AB-6 (49.5% acrylamide and 6% bisacrylamide) and 6M Urea resolving region and a 10% AB-3 (49.5% acrylamide and 3% bisacrylamide) stacking/loading region. Both resolving and stacking/loading regions have 1.0 M Tris and 0.1% SDS W/V. Poured gels were stored in 1x gel buffer (1.0 M Tris, 0.1% SDS W/V, pH 8.45) at 4°C for up to two weeks. Lanes were loaded with 3 µl of reaction mixture with loading dye, or 5 µl of BioRad Precison Plus Protein Dual Xtra standard (BioRad). Up to four gels were loaded in a BioRad Mini-Protean Tetra Cell (BioRad), with different anode (1.0 M Tris pH 8.9) and cathode buffers (1.0 M Tris, 1.0 M Tricine and 1% SDS). Using a BioRad PowerPac Universal Power Supply (BioRad), gels were run at at 15mA per gel for 60 minutes to ensure even stacking, followed by 135 minutes at 120V, or longer as needed for peptides to migrate near the bottom of the gel. Imaging was performed on a BioRad Chemi Doc EX using UV transillumination of naked gels.

RESULTS

Expression of bacterial and eukaryotic OSTs

Functional expression of *C. jejuni* and *C. lari* PglB has been long established [4]. Figure 3A Lane 1 of the immunoblot shows localization of *C. jejuni* PglB in the membrane fraction of *E. coli,* though significant amounts appeared in the membrane insoluble (lane 2) and lysate (lane 3). It is expected that closely related OSTs express similarly in *E. coli*. This was true of other members of the *Campylobacter* species, from which *C. lari* PglB appears in the membrane fraction as high mass bands, as though aggregated in inclusion bodies or detergent-resistant complexes (not shown). In figure 4A, *Desulfovibrio vulgaris* PglB expression is shown compared to *C. jejuni* PglB. *D. vulgaris* was expected to express, since

it appears to function *in vivo* in *E. coli* (Guarino, unpublished). The *D. gigas* OST showed low levels of expression, but *D. desulfuricans* did not seem to express (not shown).

Of the three *L. major* STT3 core enzyme variants able to supplant the *S. cerevisae* OST complex (STT3A, STT3B, and STT3D) [61], only the STT3A variant expressed in *E. coli* (Figure 4B). As shown in lanes 1-3 of Figure 3B, *L. major* STT3A localized as expected to the membrane fraction, though significant amounts are present in detergent-resistant high mass aggregates and membrane insoluable fractions (Figures 3B, 4B). The other three variants, however, did not show strong bands in any fraction other fractions (not shown). Successful expression of any eukaryotic OST was unexpected, and shows promise for study of this interesting and valuable protein. These *L. major* STT3 core enzymes some of the closest related eukaryotic OSTs to *C. jejuni* PglB (Supplemental Figure 1A).

Purification of bacterial and eukaryotic OSTs

Suitable expression level often indicators whether or not a protein can be purified. OSTs that do not express well do not purify in quantities large enough to be useful *in vitro*. *C. jejuni* PglB had previously been purified and serves as a model for development of a suitable in-house purification strategy [32, 64]. BCA measurements of purified proteins estimated yields of up to 0.8 mg/L despite only cursory optimization. Optimization of expression or scaling up culture strategy alone could drastically increase yield with few downstream changes [3].

Coomassie stained gels (Figure 3A left) show that the purification fractions of *C. jejuni* PglB were reasonably clean, with no major impurities within detection limits; while western blots (Figure 3A right) confirm that the coomassie stained bands are indeed His tagged PglB. PglB, like most membrane proteins, tends to accumulate in high-mass aggregates, and IMAC purification excluded many of the high-mass aggregates that appear in the lysate and membrane fractions (Figure 3A right). Desalting into storage buffer did not yield significant loss of protein, an example of which is shown in the last two lanes of Figure

3B. Purification was also effective with *D. gigsas* and *D. vulgaris* PglB, though not shown. These two particular *Desulfovibrio* OSTs had yields of 0.4-0.7 mg/L, while a third, *D. desulfuricans* did not express, nor purify (not shown). As demonstrated in the purification of OSTs from diverse species, the strategy developed here is robust for OSTs that express well.

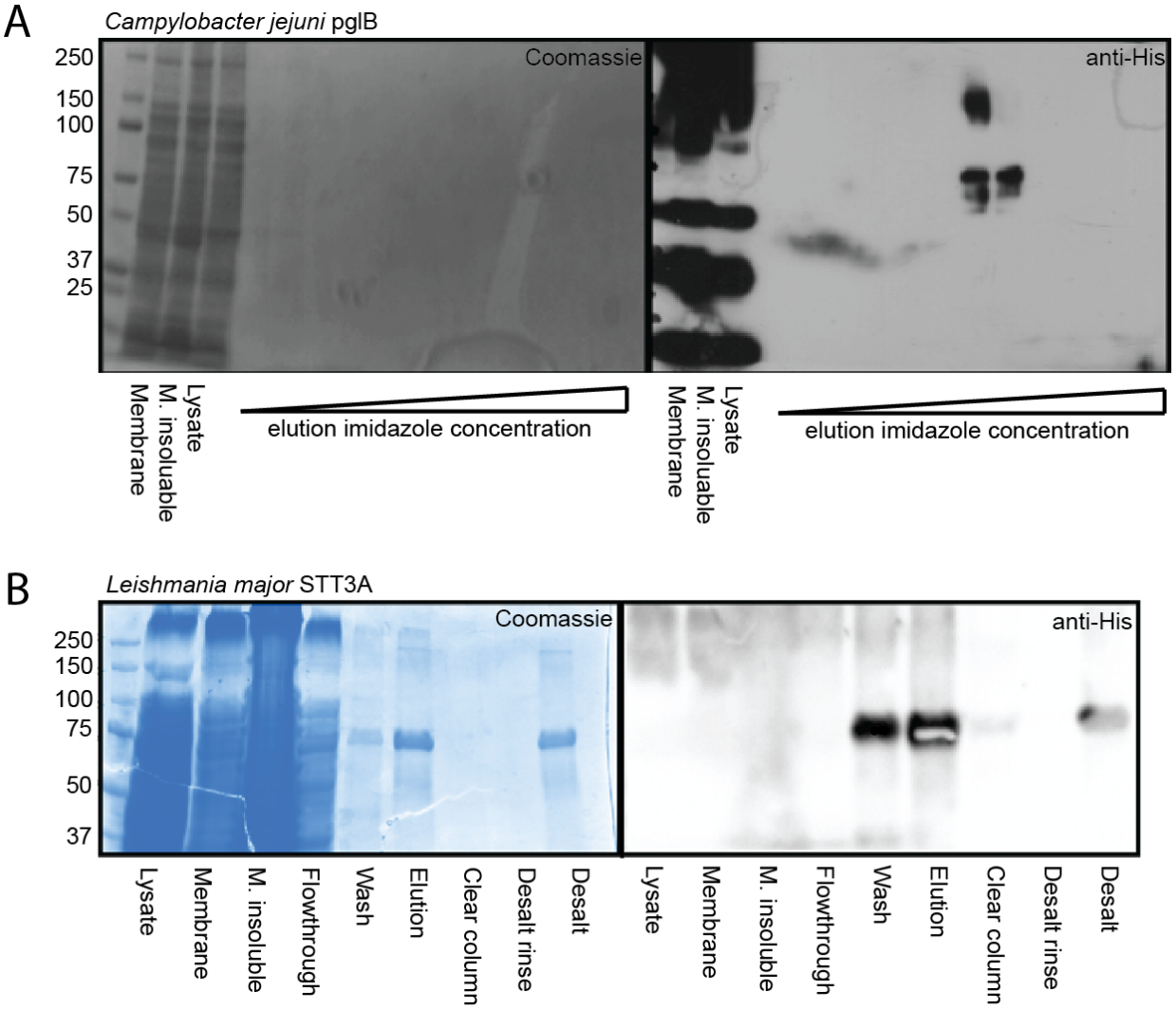


Figure 3: (A) Coomassie stain (left) and western blot (right) of *Campylobacter jejuni* PglB purification. Target PglB is abundant in high-mass aggregates and low mass degradation products in the membrane fraction, membrane insoluble fraction, and cleared lysate. After purification, PglB appears in the western blot at the expected mass (~80 kDa), and the coomassie shows clean fractions, though yields of PglB are below the coomassie detection

threshold when unconcentrated. Running at roughly 150 kDa, a potential dimer appears in the first successful elution fraction, it is unclear if this is an artifact of the purification process, or a property of PglB. A gradient of imidazole is used here to find the minimum elution concentration for subsequent purifications. Final yields are 0.8 mg/L (B) Coomassie stain (left) and western blot (right) of *Leishmania major* STT3A purification. Low expression levels in *E. coli* means that STT3A does not appear in lysate, membrane, or membrane insoluble fractions in this purification. A more concentrated membrane fraction can show expression of STT3A (Figure 3B). Target STT3A appears near the expected ~80 kDa mass in the desired elution lane (Lane 6). Relatively high imidazole concentrations and low affinity for the IMAC column results in an undesirable loss of target protein in the wash step (Lane 5). The coomasie stained gel shows that fractions containing STT3A (elution and desalt in particular) are relatively free of impurities. The final yield of *L. major* STT3A is 0.4 mg/L.

A portion of expressed *L. major* STT3A accumulated in high mass aggregates in membrane fractions of *E. coli*. In Figure 3B, IMAC purification yields proteins at the expected mass. Faint bands at roughly 160kDa in the coomassie gels of *L. major* STT3A purifications (Figure 3B) suggest the formation of dimers, confirming previous findings in yeast [61, 66]. It is unclear if these apparent dimers are active *in vitro,* or are artifacts of the preparation process, as migration of membrane proteins is highly variable with detergent, lipid, and Coomassie concentrations [67, 68]. Final yields were calculated to be 0.5 mg/L. This is the first reported expression and purification of a eukaryotic OST.

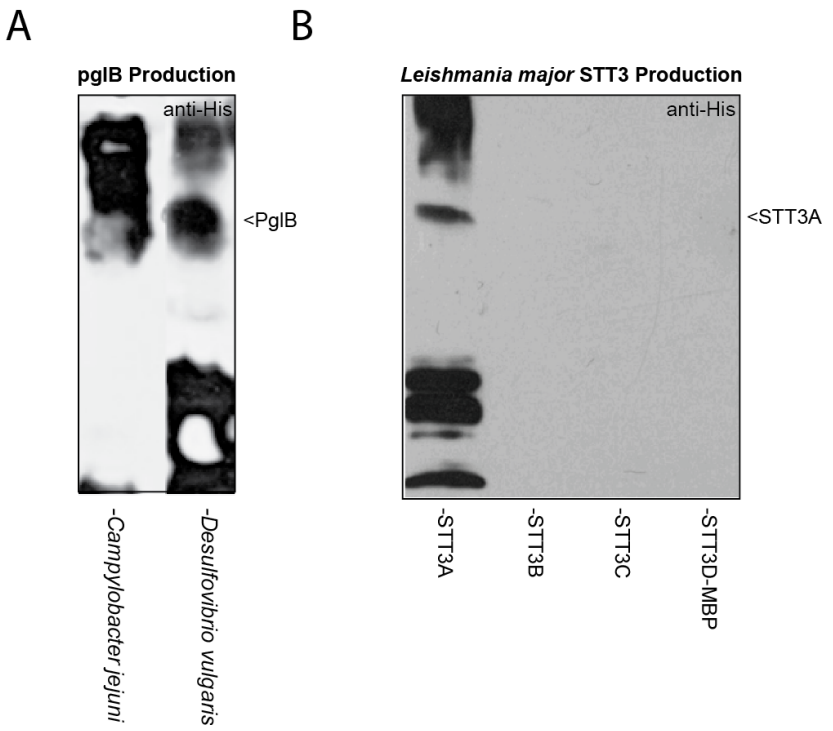


Figure 4: (A) membrane fraction from *E. coli* C43 cells expressing OSTs from *Campylobacter jejuni* and *Desulfovibrio vulgaris*. OSTs are expected to run around 80 kDa, but significant high mass aggregates characteristic of membrane protein overexpression obscures clean bands. OSTs that do not express often do not even have high mass aggregates (not shown). *C. jejuni* expression appears the strongest and has the least degradation product, and consequently has the highest purification yields of any OST expressed. *D. vulgaris* PglB shows the same characteristic high mass aggregates, but also high levels of low mass degredation products. (B) Expression of *L. major* STT3 core in *E. coli* C43. Three variants of the STT3 core are expressed: A, B and D, as well as D fused to MBP. High mass aggregates characteristic of membrane protein overexpression appear only in the STT3A case (Lane 1), with no trace of other variants detected. STT3A also appears as low mass aggregation products, though those do not appear in the purification process (Figure 3B). Because only STT3A expresses in *E. coli*, it is selected for purification and *in vitro* characterization.

in vitro activity of bacterial and eukaryotic OSTs

To evaluate OST activity, two readouts of *in vitro* assays are used: immunoblot of purified acceptor proteins and in-gel fluorescence of peptide substrates. Here, the purified acceptor proteins used were either AcrA, a native substrate of *C. jejuni* PglB or sfFv R4 with four glycosylation sequon repeats fused to the C-terminus (R4-4xGT). AcrA, as a fully folded protein offered the highest level of steric hindrance [32, 55], while R4-4xGT has reduced steric hindrance to model a translocating nascent chain [45]. The scFv R4-4xGT also may be easily recloned with an arbitrary sequence with relatively little effort. Here, R4-4xGT comes in three varities: 4xAQNAT, to mimic a eukaryotic consensus sequence, 4xDQNAT, the optimal *Campylobacter* sequence, and 4xDQAAT, a negative control which lacks the target asparagine residue [56, 62].

Peptide substrates offer high sequence flexibility for a relatively low cost, allow evaluation of substrate sequence requirements, and permit rapid, quantitative measure of glycosylation. These peptides were chemically synthesized by Thermo Fisher and have a longer shelf life than other purified proteins used here. Fluorescence measurements are highly sensitive and thus reactions were conducted in much smaller volumes than the immunoblot assays, meaning that highly valuable reagents (purified OSTs and LLOs) go a long way. Most importantly, the assays were easily made consistent and quantitative between independent sets of reactions. A simple standard curve illustrates that the assay was quantitative in ranges from O(1) pmol to O(100) pmol (supplemental Figure 1C).

To confirm that *in vitro* assays used here work properly, *C. jejuni* plgB was evaluated as a model enzyme because it is already well characterized [32]. Figure 5A illustrates the ability of purified *C. jejuni* PglB to function *in vitro* and specifically attach glycans to AcrA at both native sites. The left blot illustrates the presence of acceptor AcrA, while the right is specific for the *C. jejuni* glycan. Addition of PglB to a reaction mixture resulted in glycosylation, shown by a commensurate mass shift in AcrA on the His blot and appearance

of a band on the anti-glycan blot. The mass shift and specific detection of a glycan together demonstrate addition of a specific glycan to the target protein.

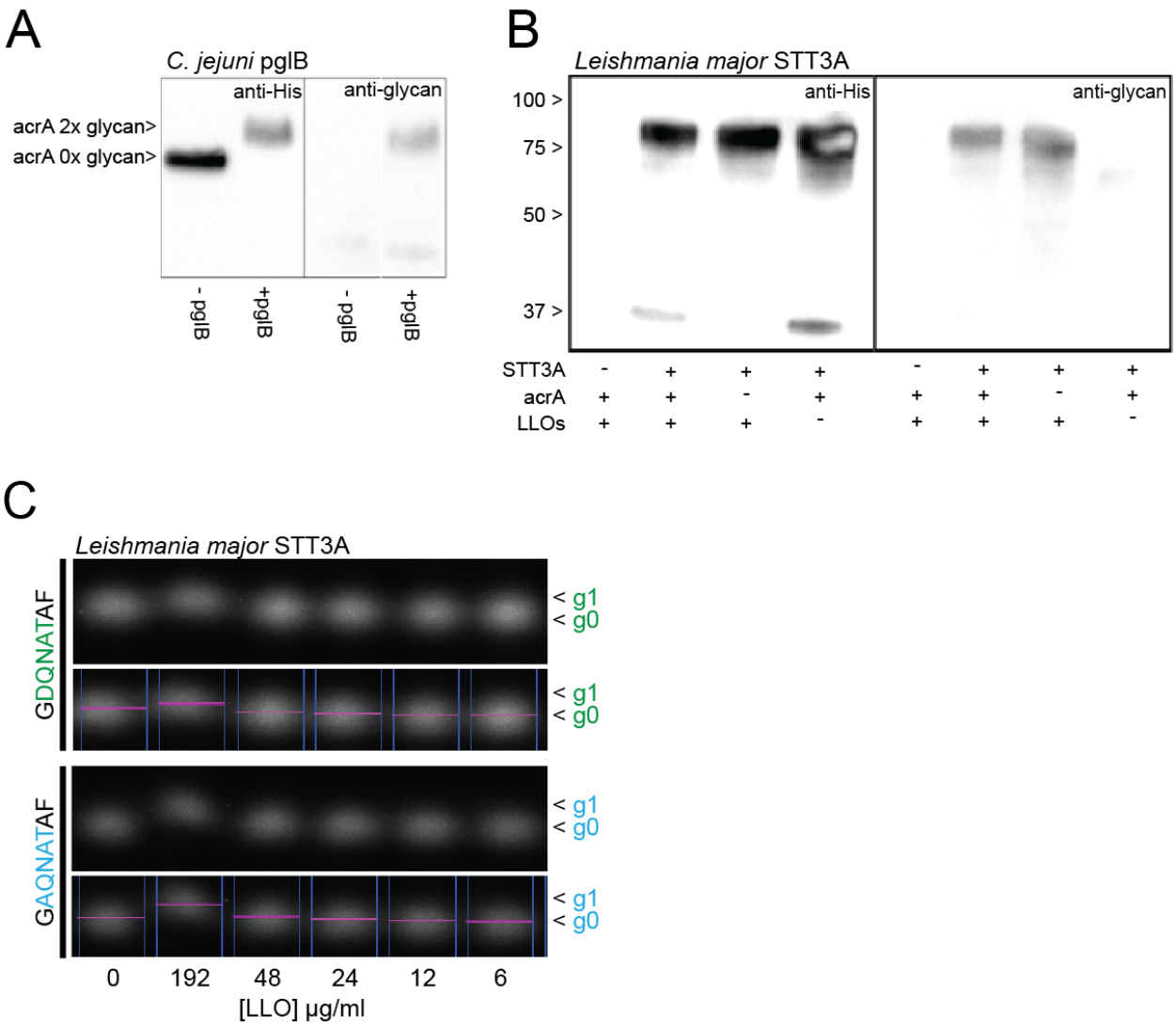


Figure 5: (A) Immunoblot glycosylation activity assay with *C. jejuni* PglB using purified AcrA as the acceptor protein. Purified AcrA and LLOs were incubated with (Lanes 2 and 4) and without (Lanes 1 and 3) *C. jejuni* PglB and the completed reactions were boiled and run on a 10% SDS page gel. The left membrane is probed with anti-His to show the mass shift of the AcrA acceptor protein and the right is probed against *C. jejuni* glycan to show the specific addition of the glycan. In the left blot, the AcrA band is shifted up approximately 2kDa, indicating the addition of two glycans, while the right blot shows a single band in the

reaction containing PglB, indicating specific addition of the *C. jejuni* glycan to PglB. (B) Glycosylation activity assay conducted with *L. major* STT3A using purified AcrA as the acceptor protein. Omission of each reagent, STT3A, LLO, and AcrA, illustrates that the reaction progresses to completion only with all present. Successful glycosylation of a target is shown with a mass shift in the His blot and the appearance of a band in the glycan blot. AcrA is the intended protein substrate of STT3A, but does not appear in the glycan blot, and does not show a mass shift in lane 2 compared to other lanes on the His blot. A band corresponding to glycosylated STT3A appears in lanes containing STT3A and LLOs (2 and 3), but not in those lacking one of the reagents. There is no observed mass shift of STT3A, but a 0.7 kDa difference on a ~80 kDa protein cannot be detected with this type of gel. Appearance of a ~80 kDa band in the anti-glycan blot suggests homoglycosylation of STT3A. (C) Peptide glycosylation assay with *L. major* STT3A. Tamra labeled GDQNATAF and GAQNATAF peptides are incubated with STT3A and varying concentrations of LLO. At sufficiently high concentrations of LLO (192 µg/mL, lane 2), STT3A appears to be able to glycosylate both the bacterial and eukaryotic glycosylation sequences. Pink lines are drawn through the calculated center of each band using Image Lab software (BioRad) to highlight the mass shift of peptides in lane 2.

Using this AcrA based assay to evaluate *L. major* STT3A core, it is clear that the enzyme does not glycosylate AcrA, shown by the absence a mass shift in the His blot and absence of any bands in the predicted range in the anti-glycan blot. However, the immunotblot shows that the *L. major* STT3A core was able to homoglycosylate, adding a *C. jejuni* glycan to another (or perhaps the same) STT3A molecule. In lanes 2 and 3 of Figure 5B, appearance of a glycan specific band in the range of STT3A suggests specific glycosylation of STT3A. These bands are present only in cases that contain both purified STT3A and LLOs, and appear independent of AcrA. No mass shift is observed for STT3A in the anti-His blot because the percent difference in mass (2%) is not sufficient to discern in

this region of the gel. A search of the amino acid sequence and rough estimates of backbone flexibility revealed that the protein is likely glycosylated on its flexible N-terminus at 30NTS. The *L. major* STT3A core was able to be functionally expressed and purified from *E. coli,* and maintained its activity *in vitro*. This is the first example of a eukaryotic OST being expressed functionally in *E. coli*, and the first example of a purified eukaryotic OST being functional *in vitro*.

Glycosylation of peptides reveals steric constraint-free sequence requirements of OSTs and allows simple quantification of extent of glycosylation. Figure 7a illustrates the ability of PglB to glycosylate peptides with DQNAT acceptor sequence but not those with AQNAT, confirming established findings [31, 56]. Reactions in lanes 1 and 2 were conducted with DQNAT, while 3 and 4 used the AQNAT acceptor sequence; together they represent a positive and negative control for the -2 acidic residue. Since differences in hydropathy and charge to mass ratio of aspartic acid and alanine significantly altered the way that the small peptide migrated through the gel matrix, the peptides appear as different masses in the gel, illustrated in comparing lanes 1 and 3 in Figure 6A. Addition of PglB to a reaction mixture containing LLOs and fluorescent linked GDQNATAF resulted in a portion of the peptide appearing at a higher mass than the PglB lacking negative control (illustrated in a higher mass band in lane 2, which is absent in lane 1). Figure 6D below shows the reaction's need for LLOs to proceed to completion. In the case of GAQNATAF, however, no mass shift is observed because *C. jejuni* PglB is unable to glycosylate substrates lacking an acidic residue in the -2 position due to an inability of R331 to coordinate peptide orientation via hydrogen bonding (Figure 7D). This confirms previous observations regarding signal sequence requirements for *C. jejuni* PglB and illustrates the ability of this assay to discern between glycosylated and unglycosylated peptides.

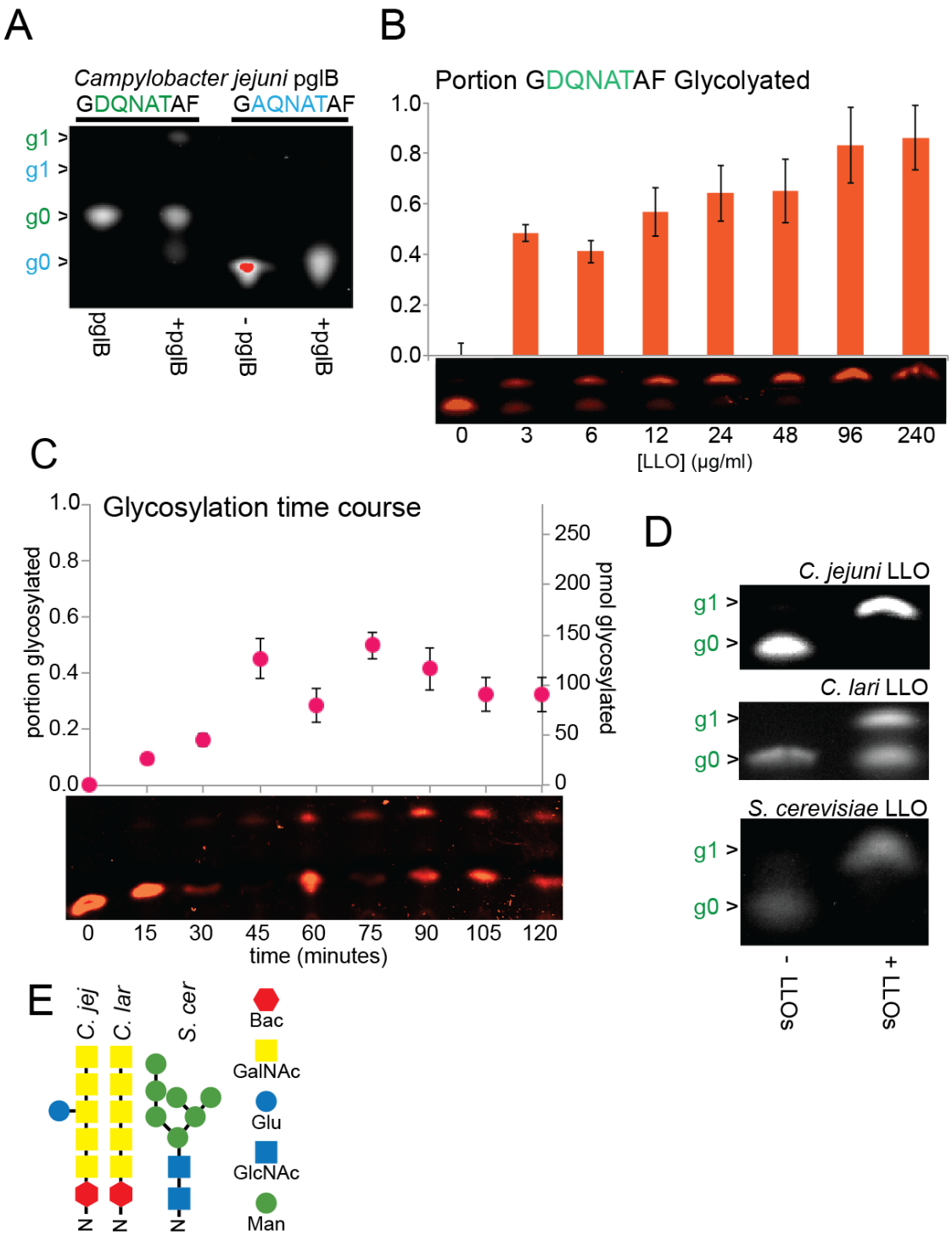


Figure 6: (A) Peptide glycosylation assay for *C. jejuni* PglB with tamra-labeled peptides. Peptide substrates with bacterial DQNAT or eukaryotic AQNAT are incubated with *C. jejuni* LLOs with and without the of *C. jejuni PglB.* Addition of *C. jejuni* PglB to the reaction mixture results in a mass shift in the peptide coding for DQNAT, seen as a portion of the

peptide fluorescing at a higher mass g1. This mass shift is not observed in the case of the eukaryotic consensus sequence AQNAT. (B) Peptide glycosylation assay investigating the role of LLO concentration on *C. jejuni* PglB activity. Tamra-labeled GDQNATAF is incubated with *C. jejuni* PglB and range of concentrations of *C. jejuni* LLOs. Portion glycosylated is calculated as the ratio of g1 band fluorescence to total lane fluorescence. Higher concentrations of LLO result in more glycosylation. Because 48 μg/ml is not statistically significant from 96 and 240 μg/ml, it was selected as the optimal LLO concentration. (C) Time course of *C. jejuni* PglB peptide glycosylation reaction. Samples of a reaction containing *C. jejuni* PglB, *C. jejuni* LLOs, and tamra-labeled GDQNATAF are incubated at 30°C, and reactions are quenched every fifteen minutes and run on a Tris-Tricine gel. Extent glycosylated (graph above) is calculated as ratio of g1 fluorescence to total lane fluorescence calculated with Image Lab (BioRad) from the fluorescence image (below). Reaction reaches completion in roughly 45 minutes, and subsequent time points are not statistically different (the 60 minute time is calculated as lower yield because the band is saturated). Subsequent experiments were run for 90 minutes to ensure they reach completion. (D) Activity of *C. jejuni* PglB with *C. jejuni, C. lari,* and *S. cerevisiae* LLOs. Tamra-labeled GDQNATAF is incubated with *C. jejuni* PglB and 48 μg/ml of either *C. jejuni* or *C. lari* LLOs. Mass shift of g0 band to g1 is caused by increased difficulty of migration through the matrix with the addition of a glycan. The difference in extent glycosylated between LLOS (eg. *C. lari* appears to only be about half glycosylated) cannot be attributed to the glycan identity, since purity and homogeneity of samples is unknown. This also offers the first demonstration of the glycosylation with a eukaryotic LLO extract, with LLOs prepared from *S. cerevisiae*. (E) Comparison of *C. jejuni* and *C. lari* glycans. *C. jejuni* features a branched hexose residue on the fourth sugar residue. Lipid carriers for both LLOs are expected to be identical. *S. cerevisiae* LLOs are extensively branched and contain different sugar residues. Like all eukaryotic LLOs, *S. cerevisiae* LLOs use a dolichol lipid carrier, unlike the bacterial undecaprenol.

The LLO concentration assay (Figure 7B) was used to investigate the role of LLO concentration on *C. jejuni* PglB and to inform the optimal reaction conditions of 48μg /ml LLO. Increased LLO concentration correlated to an increase in extent of glycosylation. Calculating the ratio of fluorescence between unglycosylated and glycosylated bands allows direct quantification of extent of glycosylation. Error was estimated based noise per unit area in regions outside of bands. Here, samples 24-240 μg/ml (lanes 5-8 in Figure 6B) were not statistically significant from one another. 48μg is selected because the unglycosylated signal was not different from that in higher concentration cases, but shows a greater signal than lower concentrations.

A time course experiment was conducted by quenching reactions with addition of β-mercatoethanol, SDS, and glycerol in the form of 6x SDS loading dye and boiling samples to denature PglB. Calculation of v_{max} from data in Figure 7C suggests a substrate turnover rate of once every four minutes. Compared to a previously reported enzyme rate [56] and assuming perfect purity, the sample was 2% active, an understandable finding given that the enzyme preparation was roughly 6 months old at the time of the experiment. It is important to note that this measurement is only valuable as a relative comparison between enzymes and is likely different than the *in vivo* rate, which may rely on association with additional protein machinery in the cell.

In addition to *C. jejuni* PglB, *L. major* STT3A was also able to use tamra-labeled peptides as a substrate. Figure 4C shows the peptide glycosylation assay applied to *L. major* STT3A. A mass shift is observed only in the case of high (192 μg/ml LLO, in lane 2) LLO concentrations. Peptide mass shift (difference between g0 and g1) appears smaller than other assays for unknown reasons, but it is suspected that it is an artifact of gel quality. The mass shift is not believed to be an artifact of lipid concentration because the presence of even higher lipid concentrations in Figure 6B shows the clear presence of bands at g0 and g1 glycosylation states.

Using LLO preparations from *E. coli* SCM6 cells harboring a plasmid coding for the pgl locus from *C. lari*, it is demonstrated that *C. jejuni* PglB is able to glycosylate a tamra-labeled DQNAT peptide with both *C. lari* and *C. jejuni* glycans (Figure 6D). Incubating *C. jejuni* PglB with either *C. jejuni* or *C. lari* LLOs resulted in glycosylation of peptide substrates, shown in the mass shift from g0 to g1 of GDQNATAF peptide in Figure 6D. Shown in Figure 6E, the *C. jejuni* and *C. lari* LLO N-glycans differ by the presence of a branching hexnac residue [69]. Addition of LLOs to a reaction mixture containing peptide substrate and PglB results in the characteristic mass shift of the peptides for both species of LLO (Figure 6E). Quantitative measurements indicate that *C. jejuni* PglB was about twice as efficient with is native LLOs compared to *C. lari* LLOs, though the concentration of mature LLOs within a lipid extract sample is unknown and difficult to measure.

LLO preparations from eukaryotic *S. cerevisiae* are also shown to work with *C. jejuni* PglB (Figure 6D). Recent work has shown that the *C. jejuni* PglB can work with dolichol substrates at low efficiency [55], and can use eukaryotic glycans [30]. This result combines the two findings into a single substrate to show relaxed specificity for lipid and oligosaccharide at the same time. This is the first report of *C. jejuni* PglB using a eukaryotic lipid extract as an LLO substrate.

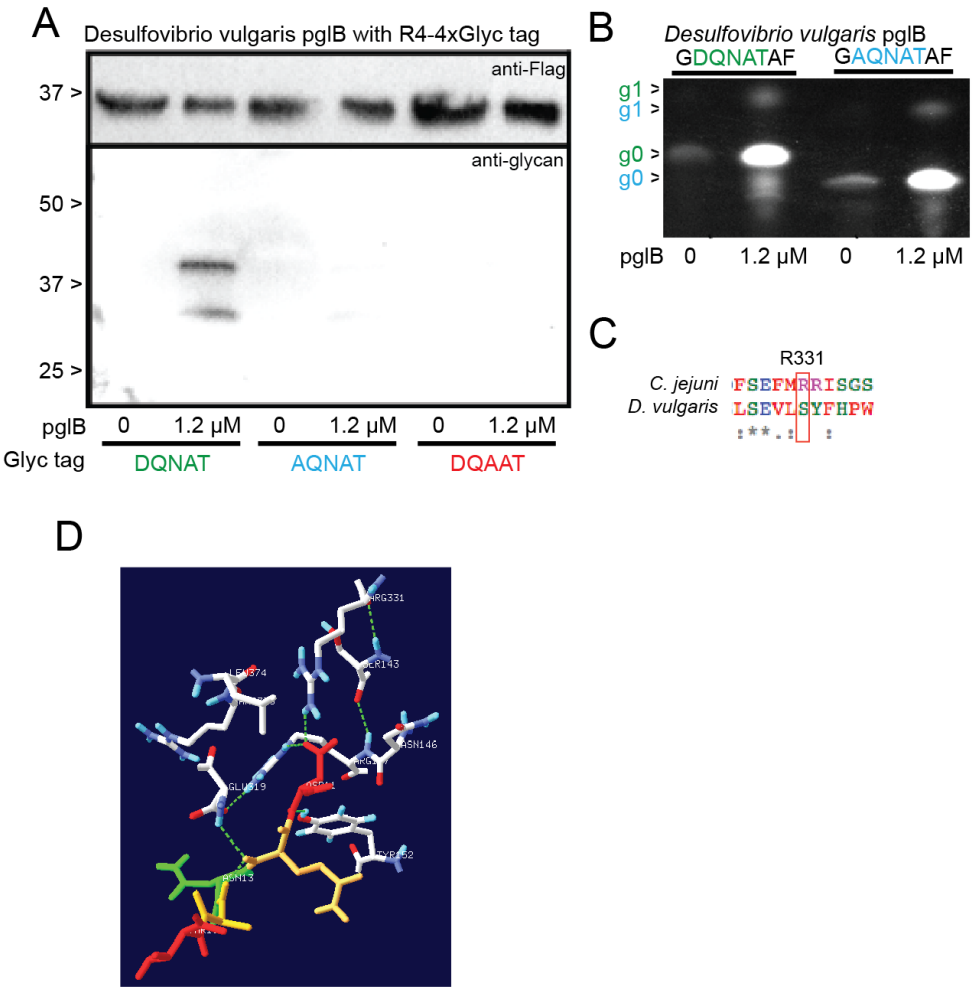


Figure 7: (A) Glycoylsation of R4-4xGT with *D. vulgaris* PglB. Purified R4-4xGT is incubated with *C. jejuni* LLOs and with or without *D. vulgaris* PglB. Western blot probed against flag tag demonstrates even expression in each sample group, while probing against glycan shows specific addition of *C. jejuni* glycan. Addition of *D. vulgaris* PglB results in glycosylation only in the case of bacterial glycosylation sequence DQNAT. eukaryotic sequence AQNAT and negative control DQAAT are not glycosylated, since no band appears at the predicted mass of R4-4xGT. (B) Glycosylation of tamra labled bacterial GDQNATAF and eukaryotic GAQNATAF peptides with *D. vulgaris* PglB. Labeled peptides and LLOs are incubated with and without *D. vulgaris* PglB. Addition of enzyme is necessary for glycosylation and is expected to result in a mass shift from g0 to g1. In both cases a band

appears at g1. Band intensity between g0 for cases with and without PglB is not constant because this image is a crop of a peptide dilution series experiment. *D. vulgaris* PglB required higher peptide concentrations than *C. jejuni* PglB when paired with *C. jejuni* LLOs and tamra labeled peptides. (C) Sequence alignment of *Campylobacter lari* PglB and *Desulfovibrio vulgaris* PglB near R331 and R332. R331 and R332 are necessary for coordination of protein substrate and mutations in this region may have consequences for substrate specificity. Here, *D. vulgaris* PglB lacks an arginine residue in that area, leading to predictions of altered substrate specificity. (D) Crystal structure of *Campylobacter lari* PglB in the catalytic cleft showing the peptide substrate and residues responsible for coordinating its orientation. Of particular are of interest is the region around R331 and R332, which participates in a hydrogen bonding network that stabilizes the acidic -2 residue of the protein substrate.

The fluorescent peptide glycosylation has also been applied to other bacterial OSTs. Figure 7B shows the glycosylation of DQNAT and AQNAT peptide substrates with *Desulfovibrio vulgaris* PglB. With the addition of *D. vulgaris* PglB to the reaction mixtures containing LLO and peptide substrates, a mass shift indicative of an added glycan from g0 to g1 appeared. Recent *in vivo* data suggests that this enzyme has relaxed specificity for substrate sequences when expressed in *E. coli* (Guarino, unpublished). While the enzyme appears to operate at a low efficiency, there are a number of factors that may contribute to this, including LLO specificity, purity and reaction conditions. Also apparent in Figure 7B is the greater fluorescence intensity in lanes 2 and 4 compared to lanes 1 and 3. This is due to higher concentration of peptide in lanes 2 and 4, as part of an assay intended to optimize reaction conditions.

The R4-4xGT fusion is intended to model the loosely structured translocating nascent chain. Here, only *D. vulgaris* has successfully been used on this substrate. In Figure 7A, an anti-FLAG Blot (above) shows even loading of R4-4xGT between reactions, while specific

probing for the *C. jejuni* glycan shows glycosylation of R4-4xGT. It was expected that DQNAT and AQNAT should be glycosylated, as in the case of peptide assays and *in vivo* evidence (Figure 7B, Guarino, unpublished). However, only R4-4xDQNAT was glycosylated, and shown via anti-glycan blots. It remains unclear why R4-4xAQNAT is unglycosylated against expectation. R4-4xAQNAT was observed to be, as expected, unglycosylated. No mass shift was observed in the anti-FLAG blot, perhaps due to glycans blocking access to the FLAG tag, meaning that glycosylated proteins do not appear in the above blot.

DISCUSSION

The primary advantages of characterizing OST function *in vitro* is that the concentrations of substrates and cofactors can be precisely tuned and that the enzyme functions independent of the glycan synthesis pathway, sec translocon and chaperones. These two combined features allow study of the enzyme in ways that cannot be controlled *in vivo* [56]. Moreover, the simplicity of the system allows reactions to be conducted in the absence of inhibiting substances and proteases, ensuring high quality and homogeneity of glycosylated product. Indeed, efficiency and extent of glycosylation *in vitro* is often higher than that observed *in vivo* [28, 56, 58]. Enzyme substrates including both LLOs and acceptor can be evaluated in a much higher throughput manner than *in vivo*, with hundreds of conditions easily tested in a single day. This constitutes not only a tool for rapid characterization, but optimization of reaction conditions. Moreover, the ability to precisely control both LLO and peptide substrate concentrations and measure product concentrations opens new doors for understanding reaction kinetics in ways otherwise inaccessible.

Expression and purification of OSTs from an *E. coli* expression platform has proven successful for targets from a range of species. Well understood *C. jejuni* (Figure 3A) and *C. lari* (not shown) PglB were used to optimize expression and purification methods, before scaling to other well expressed bacterial OSTs, like *D. vulgaris*. Techniques that were successful for bacterial OSTs were applied to *L. major* STT3A, the first eukaryotic OST to be

expressed and purified from *E. coli.* Activity assays *in vitro* confirm function of *C. jejuni* and *D. vulgaris* OSTs with a selection of target sequences and steric constraints. Experiments with purified AcrA and peptide substrates and *L. major* STT3A represent the first reported *in vitro* activity of a eukaryotic OST. Of the OSTs expressed and purified, the *L. major* STT3A is the most distant evolutionarily from any other target, suggesting that the technique should be robust for this class of protein.

The previously established requirement of *C. jejuni* PglB substrates to possess the canonical bacterial glycosylation sequence (D/E)-X_1-N-X_2-(S/T) (where X_1 and X_2 are any amino acid except proline) is confirmed by both the immunoblot and peptide based assays using DQNAT and AQNAT sequences. *D. vulgaris* PglB seems to have a preference for the bacterial DQNAT using the R4-GT substrate and no preference for peptide substrates. This may be due to steric constraints in the catalytic cleft of the enzyme. An alignment of sequences around the R331 (Figure 7C), the residue important for substrate coordination in *C. jejuni* (Figure 7D) shows differences in sequence that may explain relaxed substrate specificity and steric constraints.

Immunoblot assays show that *L. major* STT3A is able to homoglycosylate, but is unable to glycosylate fully folded AcrA. These observations, given previous findings of eukaryotic and mammalian OSTs are not expected to be able to glycosylate structured domains [51, 54]. Sequence analysis reveals a probable glycosylation site on the highly flexible N-terminus residues 30NTS. Observations of *L. major* STT3A activity with peptide substrates confirms expectations that STT3A is able to glycosylate the minimal eukaryotic sequon of N-X-T in unstructured substrates. *L. major* seems to require high concentrations of LLO for function, and does not produce as large a mass shift as other experiments. The lower mass shift may be due to a preference for rare, unelaborated glycans that more resemble eukaryotic glycans. Use of eukaryotic OSTs in bacteria holds promise for extensive applications in biotechnology.

With the ability to purify and characterize a diverse set of OSTs in a robust and scalable way new insight about these valuable proteins can be gained rapidly. Characterization of each enzyme's preference for sequon identity and steric requirements permits identification of ideal candidates for applications protein production. The ultimate goal will be to identify enzymes that are able to glycosylate eukaryotic acceptor sites on sterically constrained peptide substrates, allowing for the most diverse range of acceptor proteins. Enzymes that perform as desired could then be engineered for improved expression and find application in large scale protein production.

Differing specificities for sequence and lipid carrier can be leveraged to select a set of enzymes that is able to place different glycans on unique sites with high specificity. In eukaryotic systems, elaboration and completion of the glycan occurs after translation and translocation, requiring a range of enzymes that may not be an option in an *E. coli* host [5, 30]. Such specificity will allow for glycan diversity far beyond current capabilities, and open doors for production new and interesting proteins that are currently difficult targets, including proteins from arcaea.

The high throughput nature and quantitative output of the in-gel peptide fluorescence assay allow for measurements that are otherwise difficult with immunoblot assay, including the role of LLO concentration and kinetic rates. To date, Michaelis Mentin kinetics have been used to model OST activity [56], an assumption that requires fixed concentration of either LLO or protein substrates. A more accurate ternary-complex model can be applied to better understand the kinetics of substrate binding and reaction mechanism. Details like substrate binding order and release order can be explored in ways previously not possible. Understanding of these mechanisms can be used to optimize both *in vivo* and *in vitro* systems and provide guidance for further optimizations.

Previous studies have shown flexibility of *C. jejuni* PglB with truncated glycans and varying lipid substrates [30, 50, 55], but this is the first demonstration with fully exogenous glycans. Other results have confirmed that this enzyme is able to function with diverse

glycans, including eukaryotic-like substrates [30]. Emerging techniques to synthesize other interesting glycans [55, 70] may allow *in vitro* glycosylation with sugars that are otherwise difficult to synethsize *in vivo*. Intersting sugars, particularly azido-sugars, can be used to functionally label proteins in specific locations with high efficiency and minimal impact to a protein's structure. The demonstration of function with *S. cerevisiae* LLOs opens doors to use lipid extracts from valuable sources. Previously established relaxed specificity for glycans [50,55] and glycans [30] invited the idea of using mixed glycans that may be valuable in biotechnology. *S. cerevisiae* serves as a proof of concept and more complex substrates may be possible.

Evaluation of substrate requirements for protein substrates and homology modeling of the catalytic core to the *C. lari* structure may lend insights into how different OSTs interact with their substrates. Understanding protein-OST interactions can allow engineering of bacterial OSTs that will glycosylate eukaryotic substrates. It may also be the case that folding state and acceptor sequences are closely related features, where the acidic -2 residue coordinates fully folded proteins, performing the job of the chaperones that unfold and coordinate protein substrates [56]. In presenting substrates with a variety of different folding states, levels of steric hindrance, and amino acid sequence, this work allows answering of questions that have previously evaded researchers.

Application of these purification and characterization techniques to a broad range of OSTs will lead to an enhanced understanding of one of the most important post-translational modifications. Improved understanding may even lend insight into diseases caused by malfunction of glycosylation machinery [71]. With enhanced fundamental understanding, and a means to produce large, high purity quantities of OSTs, additional crystal structures are within reach. With a number of parallel advances in glycan design and LLO production, eukaryotic therapeutic proteins will be accessible targets for production using *E. coli* as a platform.

CHAPTER 4: NOTABLE ACHIEVEMENTS

First demonstration of beta-lactamase based reporter for interaction between soluble and membrane proteins capable of genetic selection

Demonstration of high dynamic range for strong associating peptides between soluble and membrane proteins

Development of a scheme to engineer antibodies exclusively against periplasmic/extracellular domains of membrane proteins

First example of eukaryotic oligosaccharyltransferase expression in *E. coli*

First example of eukaryotic oligosaccharyltransferase purification from *E. coli*

First example of eukaryotic oligosaccharyltransferase activity *in vitro*

First use of yeast membrane extracts as an LLO substrate for *C. jejuni* PglB

Evaluation of a diverse selection of oligosaccharyltranferases with peptide substrates to examine substrate sequence specificity

CHAPTER 5: FUTURE DIRECTIONS

Split Beta-Lactamase Protein Complementation Assay for Selection of Antibodies Against Transmembrane Proteins

The ultimate goal of engineering antibodies against an arbitrary membrane protein target requires three intermediate goals. First, selection conditions that offer an appropriate dynamic range between binding and non-binding cases must be found. Second, scFv libraries must be selected against a contrived target using the new selection conditions. Finally, binding of antibodies is to be confirmed *in vitro* to ensure that observed resistance is not an artifact of a non-specific binding interaction. These three aims together yield a methodology that enables the targeting of scFvs against arbitrary transmembrane protein targets.

Specific Aim 1: Optimizing Positive-Negative Signal Resolution for scFv-antigen interactions

Having demonstrated the ability of split membrane-anchored beta-lactamase to confer antibiotic resistance correlated to fragment binding strength, a platform for engineering binding molecules is ready for application. Presently, scFv based binding systems are not sufficiently orthogonal to give rise to conditions that allow selection of appropriately large libraries (Figure 2C). Current shortcomings may be due to spontaneous reassembly or inherent resistance, as demonstrated with the non-binding and inherent resistance controls. Inherent resistance is overcome by increasing the positive binding cases and non-binding assembly is remediated by increasing the barrier of interaction for reporter fragments.

Optimization is necessary to find conditions that allow appropriate selection resolution, by either increasing positive case resistance or decreasing non-binding interaction of reporter molecules. Here, optimization in either case relies on tuning two

facets of the system: bait/prey configuration and expression level. Should these methods fail, alternative scFv frameworks and reporters can be explored. Reporter configuration is altered by changing the order of fusion proteins, and identity of reporter-fusion pairs. Expression level is tuned by exploring alternative expression platforms and assay conditions. Lastly, alternative scFv frameworks and reporters can be considered.

Sub aim A: Altering Bait/Prey Configuration

It has been shown that the order of bait and prey in reference to reporter fusion (N-terminal reporter vs C-terminal reporter) and the identity of the reporter fragment can influence the effectiveness of reporter assembly [44]. To explore these configurations, constructs will be recloned replacing the reporter fragment and/or changing the order of reporter and bait/prey. This reconfiguration is accomplished with simple directional cloning into the already modular vector. Testing of new combinations is accomplished by applying the proven plating selection methods, and will use the already proven controls to demonstrate effectiveness.

For example, current cases rely on N-antigen-ωBla-membrane protein-C and N-scFv-αBla-C as the bait and prey respectively. Altering the prey configuration to N-αBla-scFv-N, or bait and prey to N-αBla-antigen-membrane protein-C and N-scFv-ωBla-C may offer greater resolution between binding and non-binding cases by reducing spontaneous assembly of reporter fragments. By making assembly slower or more difficult, the background noise is reduced and signal to noise ratio of positive cases is increased.

Sub Aim B: Tuning expression level

Expression levels of associating fragments is a key aspect of optimizing PCA methods [42]. Given proven sensitivity to expression level, the membrane-anchored component is the more difficult of the two constructs to alter. Thus, relative expression is best tuned by recloning the scFv portion into a vector under the control of a highly tunable promoter. This will allow the relative concentrations of bait and prey to be tuned over a much wider range

than the current puc-based expression platform [34]. With this, it will be possible to explore the stoichiometry of the antibody-antigen interaction and offer new directions to optimize this particular selection platform.

Current options for a scFv-reporter fusion plasmid are only limited by bacterial resistance. The chloramphenicol resistant pBad-18 that forms the backbone of the yegH expressing plasmid and the ampicillin resistance of the binding reporter exclude those two antibiotics from the new plasmid's resistance. Viable options for such a plasmid include a Kan resistant pET vector (high expression) or Kan resistant tetracycline inducible (highly tunable expression) vector.

Sub Aim C: Framework matching

Changes to scFv framework can have dramatic consequences for expression level and folding stability [24]. The current configuration using D10 and scFv anti-GCN4 frameworks may create irreconcilable differences in expression level. Selecting controls made from the same scFv framework should balance expression levels and more accurately model library selection conditions. The existing platform for this comparison utilizing so called GLF and GFA mutants of scFv anti-GCN4 has kD values that are too close to be discerned between [25]. Alternative mutations to the CDR3 may further reduce binding to GCN4 without significantly altering expression level.

Selecting a scFv anti-GCN4 mutant with very low affinity to GCN4 may allow discerning between non-binding and high binding scFvs. Another option is to use scFv D10 with GPD anchored to the membrane. A negative, non-binding control with the same framework, J21, does not bind to GPD and can serve as an alternative to the current scFv anti-GCN4/GCN4 system [46]. Parts of this alternative scheme is already cloned (ssDsbA-GPD-ωBla-yegH and αBla-D10), leaving only αBla-J21 to be cloned. Implimentation of αBla-J21 will hopefully increase selectable range for scFv binding above the existing $O(10^2)$

(Figure 2C). This simple switch to an alternative contrived scFv configuration may offer sufficient resolution to demonstrate selection between binding and non-binding scFvs.

Sub Aim D: Alternative reporters.

While split beta-lactamase has proven effective for soluble associating pairs (Figure 1C) [34], limitations in selection resolution may not be suited for membrane-anchored targets. As such, it may be worthwhile to explore alternative reporter proteins. One key advantage of the Bla-based reporter is the ability to function in the periplasm, permitting scFv targeting against extracellular domains of human proteins. When considering alternative reporters, periplasmic activity precludes a number of proven options including split mDHFR [46], split CRE [72], and split YFP [44]. A perisplasmically active mutant of split superfolder GFP offers a periplasmic alternative [73, 74].

Just as with altering bait/prey configuration, exploring alternative reporters requiring recloning existing vectors. A split GFP format developed by Zhou uses GFP1-10, the first 10 beta-sheets of superfolder GFP fused to the prey molecule and the engineered M12 peptide, an analogue to the 11^{th} beta-sheet of GFP, fused to the bait molecule. In place of αBla and ωBla, GFP1-10 and GFPM12 [74] will be inserted. Earlier work has demonstrated that the GFP1-10/M12 reporter configuration yields a detectable signal with Fos/Jun bait/prey fusions (not shown). GFP1-10 will be genetically fused to the scFv, while the M12 peptide will be inserted in place of ωBla to form ssDsbA-GCN4-M12-yegH. Similar controls to previous cases will be employed in an effort to discern between high and non-binding cases.

Specific Aim 2: Screening scFv Libraries Against Transmembrane Proteins

The goal of developing conditions that discern between binding and non-binding scFvs leads to being able to screen scFv libraries against membrane proteins to select binding members. This aim proceeds in a series of three cases, building to the ultimate goal of engineering antibodies against arbitrary transmembrane proteins. First, a contrived library that has been demonstrated to function in other selection systems [25] will confirm

that the selection is able to create desirable results. From there, a naïve library may be screened against native *E. coli* proteins that are understood to have high expression levels and tolerate fusion proteins. Finally, scFvs will be engineered against interesting and biotechnologically relevant membrane proteins.

Sub Aim A: Contrived membrane linked antigen substrates

Before true naïve libraries can be screened, a previously characterized library will demonstrate the ability to select known binding members. Previous studies in the DeLisa laboratory, have selected libraries of scFv anti-GCN4 against a GCN4 epitope using the FLI-TRAP selection scheme [25]. In Waraho's study, an NNK library of three residues in CDR3 of scFv anti-GCN4 was constructed and screened to select high binding mutants. Following a similar methodology, this project seeks to screen the same NNK library of CDR3 of scFv anti-GCN4 against the existing GCN4-ωBla-yegH to find scFvs that bind to GCN4. Known high binders have a GLF motif in the CDR3 domain, while non-binding or low-binding mutants have other sequences, GFA is one such well characterized low-binding mutant [25]. Testing positive hits against a negative control of ωBla-yegH will ensure that members bind only to GCN4.

The library of scFv anti-gGCN4 NNK CDR3 has already been created and sequencing was underway. Pending confirmation of library size and diversity, a series of trial selections of library members will be used to determine conditions that yield desired numbers (O(10)) of colonies, as in [25]. Selected colonies will be retransformed and screened to determine precise resistance, and high resistances will be assumed to be false positives. Positive hits suspected to be true results will be cloned into fresh vectors via PCR and rescreened against target and negative controls to show that members bind only to the GCN4 epitope. Recloning makes sure that no recombination events reconstitute full length Bla between the two vectors. Further characterization *in vitro* will confirm direct binding to the target region, and provide affinity measurements.

Sub Aim B: Antibodies against native E. coli transmembrane proteins

It is best to screen naïve libraries against membrane proteins with well-characterized proteins that are known to express at high levels in E. coli. Selection of high expression targets ensures that weak signals can be attributed to non-binding rather than an absence of target molecule. Moreover, it is important to select targets that are less toxic to cells, and native proteins often satisfy this requirement. In this line of thinking, Skretas and Georgiou selected yegH as their native *E. coli* membrane protein anchor for their membrane protein folding reporter [47].

To find antibodies against membrane proteins, ssDsbA-ωBla-yegH will be expressed as the target protein. Binding of an scFv-αBla prey molecule to yegH will reconstitute full length Bla, lending antibiotic resistance. As with the contrived anti-GCN4 scFv library, successful binding scFvs will be recloned into fresh vectors to confirm binding. In addition to confirming antibiotic resistance with the selection target of ssDsbA-ωBla-yegH, a negative control of ssDsbA-ωBla will serve to remove any library members that might bind to ωBla or a linker sequence. Successful scFv mutants that yield resistance when expressed with ssDsbA-ωBla-yegH, but not ssDsbA-ωBla, will be further characterized *in vitro* to demonstrate and measure binding.

Sub Aim C: Antibodies against biotechnologically relevant transmembrane proteins

With the ability to target antibodies against native *E. coli* proteins, targeting antibodies against an arbitrary transmembrane protein target is limited only by the ability to express target proteins. Chief targets include OSTs and GPCRs for their importance in biotechnology, and ability to be expressed and purified in *E. coli*. A range of bacterial and the only reported eukaryotic OST have been expressed in this project. Earlier work in the DeLisa laboratory successfully expressed a range of human GPCRs from [75]. Balancing expression level with cell stress of overexpression will present another challenge. Examples from Skretas [47] clearly illustrate that there is no universal method for improving expression of membrane proteins.

Lacking a guaranteed method to improve expression without increasing toxicity, tuning expression levels will require a heuristic approach for each individual family of membrane protein. Such techniques will begin with simple optimization of growth conditions, strain, and expression plasmid. It is expected that these techniques will be sufficient for the inhouse GPCR (CB1, CB2, BR2, and NKR1 [75]) and OST (*C. jejuni, C. lari, D. vulgaris, D. gigas,* and *L. major*) since their expression can already be detected in membrane fractions (not shown). Further optimization of membrane protein expression level will demand implementation of a selection platform, as espoused by Skretas [47].

Tuning scFv library expression will also prove beneficial, because carefully tuning stoichiometry of bait-prey reporter assembly interactions can improve resistance. This is achieved through implementation of alternative expression vectors for the αBla-scFv construct. Highly tunable tet-promoter based vectors are desirable in a case like this, since they permit a broad range of expression levels. Such vectors are likely to be implemented to overcome current failures to create sufficient selectable range, so minimal cloning will be required to optimize expression level for each new membrane protein target.

Specific Aim 3: Develop *in vitro* Assay to Confirm Binding of scFvs to Membrane Targets

To ensure that selected resistance phenotypes are the result of specific binding interactions, binding must be confirmed *in vitro*. Purified scFvs and target membrane proteins will be used to show that the antibiotic resistance phenotype is not an artifact of an alternative or non-specific binding interaction. The primary options available are: Enzyme-linked immunosorbent assay (ELISA) and surface plasmon resonance (SPR). These methods offer a means for measuring scFv binding outside the context of the PCA machinery and cell milieu.

ELISA or SPR with membrane proteins begins with purification of detergent solubilized targets, a process which will borrow proven techniques from the OST project.

Native *E. coli* proteins will be purified by applying techniques developed for more difficult, exogenous membrane proteins (See OST characterization chapter). Initial targets for naïve libraries will be constrained to those that have been purified effectively—namely oligosaccharyl transferases and GPCRs. With the exception of methods used to purify membrane protein targets, *in vitro* scFv binding measurements rely heavily on long established and well-understood methods [76, 77].

Sub Aim A: Enzyme-Linked Immunosorbent Assay (ELISA)

A common way to confirm that selected scFvs are true positives is the Enzyme-Linked Immunosorbent Assay (ELISA) [24]. In this assay, the target membrane protein is purified and captured on the solid phase of a 96-well plate, and is incubated with the purified scFv. Binding of the target protein to the immobilized scFv can be detected by probing against a tag specific to the target protein [76]. In this format, membrane proteins are incubated on coated ELISA plates to immobilize them on the surface. Next, a scFv with a unique protein tag is incubated with the immobilized protein and, after rinsing away excess, is detected with an antibody specific for the unique tag. This technique may detect binding, and can even be used to compare binding of different variants of an scFv [24].

Sub Aim B: Surface Plasmon Resonance (SPR)

Surface plasmon resonance makes use of a highly sensitive optical detection scheme to measure binding of molecules near a high precision gold surface [78]. Measuring the index of refraction of an incident laser on a gold surface, it is possible to detect minute changes in interface properties. This technique has been applied to antibody binding to protein targets [77] and for detecting ligand binding to immobilized membrane proteins [79].

To demonstrate binding of selected scFvs, two schemes of detection may be employed: immobilized membrane proteins and scFv captured, or immobilized scFvs with membrane protein captured. Membrane proteins may be immobilized in a number of ways

including antibody capture of detergent solubilized membrane proteins [79, 80]. Additionally, NTA may be used to immobilize His-tagged membrane proteins [81]. Conversely, scFvs may be immobilized using either free-amine coupling or His-NTA [82]. High sensitivity has been achieved detecting small molecule binding to immobilized GPCR targets [80, 81], making detection of macromolecules like scFvs a probable outcome.

Expression, Purification, and Characterization of Oligosaccharyltransferases in *Escherichia coli*

A scalable platform for OST expression and purification from *E. coli* has been developed and preliminary characterization of *C. jejuni* PglB demonstrated. Further applications of these techniques will allow new levels of understanding for not only the well understood *C. jejuni* and *C. lari* OSTs [31, 50, 53, 56], but also interesting and valuable OSTs that appear to have relaxed or altered substrate specificity (Guarino unpublished). In addition to investigating protein substrate preferences, the developed immunoblot and peptide fluorescence assays allow investigation of OST lipid and glycan requirements. Characterization of substrate requirements allows a number of expanded applications including therapeutically relevant sugars and a variety of glycan-linked reporters.

Specific Aim 1: Purification of new OSTs

Sub Aim A: R331 PglB mutants

Mutations to R331 in *C. jejuni* PglB alter protein substrate specificity *in vivo* by changing the hydrogen bonding pattern between the OST and its substrate(Figure 7C) (Guarino, unpublished). It is unclear whether this mutation has consequences for expression level, substrate binding strength, or catalytic activity. It is hypothesized that this mutation modifies the enzyme's ability to bind the substrate, since R331 forms a hydrogen bond with the -2 residue of the protein substrate (Figure 7B) [31]. Expression and purification of mutants with altered specificity will offer insight into how the enzyme interacts differently with the substrate. Combining *in vitro* assays with binding studies, the efficiency of binding and catalysis can be measured. This also serves as a platform to engineer variants of *C. jejuni* PglB with different protein specificities, which could allow different glycans to be put on a single protein. For example, a protein can be glycosylated at DQNAT *in vivo*, then purified and glycosylated at AQNAT *in vitro*, yielding a protein with two distinct glycans.

To evaluate the activity of these enzymes, select R331X mutants will be transformed into C43 cells, and evaluated for expression just as PglB has been previously. With successful expression, enzymes will be purified and tested for activity on tamra-labeled DQNAT and AQNAT sequon peptides. Enzymes will be compared for rate and extent of glycosylation under a variety of concentrations of substrate. The goal of these experiments is to confirm *in vivo* findings about the ability of R331X variants to glycosylate AQNAT *in vitro*.

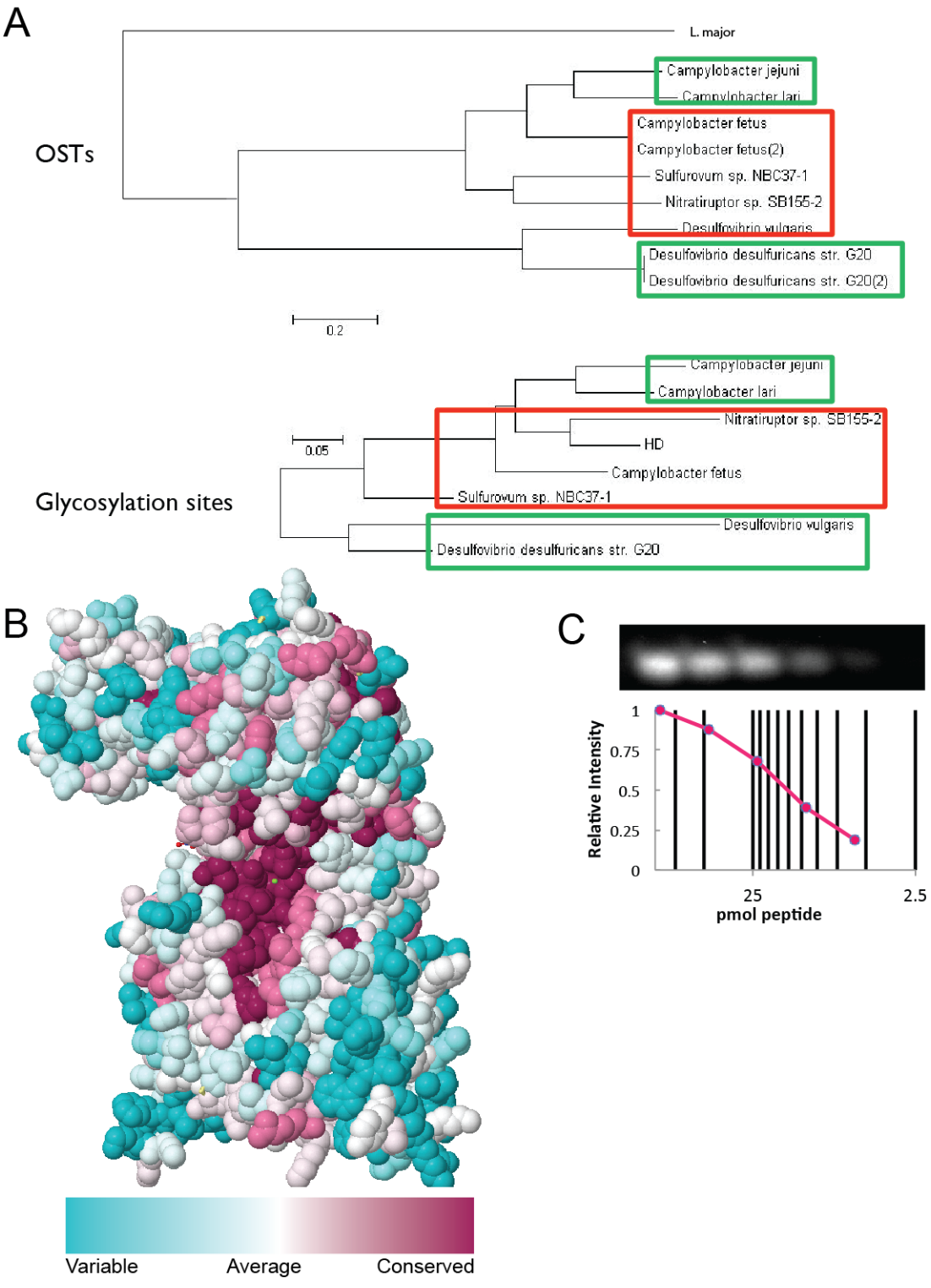
Sub Aim B: OSTs from other bacterial species

Recent work in the DeLisa laboratory has uncovered some evidence that a number of OSTs from the *Desulfovibrio* genus are able to glycosylate a Eurkayotic-like acceptor site AQNAT *in vivo* (Guarino, unpublished). To confirm this relaxed acceptor site specificity, *D. vulgaris* has been purified and reacted *in vitro* with peptide substrates. Two more enzymes from *D. gigas* and *D. desulfuricans* appear to have similar activity [29] and preliminary evidence indicates that *D. gigas* has similar performance to *D. vulgaris*. To date, *D. desulfuricans* has evaded purification.

Here, the scheme follows as before, evaluation of expression is followed by purification of target OSTs. Successfully purified OSTs will be evaluated for ACRA, R4-4xGT and peptide reactions will be used to investigate folding state requirements of each enzyme. Acceptor site sequence requirements will be conducted using peptide assays to rule out the influence of steric constraints. Peptide assays will also be used to calculate rates and evaluate LLO preferences by measuring extent of glycosylation over a short time course.

Ideal OSTs have higher catalytic rates and eukaryotic-like sequence specificity compared to the current standard of *C. jejuni* PglB. Sequences of candidates can be aligned with well-characterized OSTs to evaluate the role of mutations in variable regions on catalytic activity and substrate binding. Preliminary evidence shows high conservation in substrate binding and catalytic domains (Supplemental Figure 1B), but it remains unclear

which domains are most important to improve performance. Such insights could be valuable for creating chimeras with high catalytic rates and specifically designed substrate binding.



Supplemental Figure 1: (A) Distance tree of OSTs (above) and acceptor sites (below). OST distance trees are calculated based on sequence identity and assembled using Molecular Evolutionary Genetics Analysis. (Tamura, Stencher, Peterson, Filipski, & Kumar, 2013). Glycosylation sequence distance is calculated as the probability for enrichment of a specific amino acid in the X_1 and X_2 residues (-1 and +1 relative to the asparagine residue) in context of a glycosylation site compared to an arbitrary asparagine residue for bacterial hosts. (B) Conversation of residues in bacterial OSTs mapped onto the three dimensional structure of *C. lari* PglB from (Lizak C. , Gerber, Numao, Aebi, & Locher, 2011). Highly conserved regions (maroon) tend to be focused in the catalytic cleft (maroon pocket in center), while outer regions, as expected, tend to be highly variable. A degree of variability near R331 and R332 permits invites expectation of differences in substrate specificity. (C) calibration curve of peptide fluorescence in Tris-Tricine SDS PAGE gels. A simple dilution series illustrates a direct correlation between concentration and signal strength over several orders of magnitude and sets a minimum detection threshold on the order of single pmol of peptide.

Sub Aim C: eukaryotic OSTs

Studies suggest that eukaryotic OSTs are highly sensitive to steric hindrance of the protein substrate and to the identity of the LLO lipid carrier [51, 54]. Further *in vitro* characterization of *L. major* STT3A will help to expand understanding of not only this enzyme's properties, but also eukaryotic OSTs in general. Relatively little is known about eukaryotic OST substrate specificity and LLO requirements [5]. The current three model protein substrates allow investigation of structured, loosely structured and unstructured domains, which will be useful in tuning reaction conditions. eukaryotic LLOs are desirable in biotechnology in their ability to place eukaryotic glycans on eukaryotic acceptor sequences [5, 30].

Further characterization of *L. major* STT3A will begin with purified R4-GT, to see if the enzyme is able to glycosylate loosely structured domains. Since eukaryotic OSTs are known to associate with the sec translocon and glycosylate cotranslocationally [51, 58] it is expected that eukaryotic OSTs can glycosylate loosly structured domains. More detailed peptide assays with an expanded set of LLOs will begin to investigate lipid and glycan requirements for the enzyme. At present, only *C. jejuni* LLOs have been evaluated, and *C. lari* and *W. succinognenes* are likely next targets. Advances in LLO preparation (Specific Aim 3) will enable use of eukaryotic LLOs as well.

A number of eukaryotic OSTs are closely related to the *L. major* OSTs, specifically those from the *Trypanosoma* genus. OSTs from *T. cruzi* and *T. bruci* will be cloned into bacterial expression vectors. These OSTs will be evaluated for expression and purification. Targets that are successfully purified will be characterized for their substrate specificity, paralleling work with *L. major* STT3A and bacterial OSTs.

Specific Aim 2: Characterization of lipid requirements

Lipid requirements for OST function have been characterized only in limited ways, often using crude membrane isolates rather than purified OSTs [50, 59]. Peptide based i*n vitro* characterization of purified OSTs allows precise control of LLO, enzyme and substrate concentrations in ways that are impossible *in vivo* or with crude membrane fractions [55]. It is possible to examine precise LLO requirements for OSTs in terms not only of concentration, but also phase behavior, and identity.

The Imperiali lab has succeeded in isolating isoprenoid species from plants and testing their function as a substrate with *C. lari* PglB [50]. Similarly, the Davis lab has been able to synthesize rationally designed polyisoprenoid carriers to investigate LLO requirements of PglB [55]. Findings have suggested that *C. lari* PglB prefers unsaturated head groups [50, 55], while investigations in yeast have demonstrated a preference for saturated head groups in eukaryotic OSTs [59].

Sub Aim A: Lipid phase behavior

It has long been established that lipid phase behavior is highly sensitive to even small changes in concentration or hydropathy [8]. Despite the recent publishing of a crystal structure of *C. lari* PglB, it is unclear how the lipid carrier binds to the hydrophobic domain of the OST [31]. Early results have indicated that LLO concentrations above critical levels decrease glycosylation efficiency (not shown). Moreover, the integration of LLOs into OST-harboring detergent micelles is an uncharacterized that is open to optimization.

Phase behavior can be evaluated with varying concentrations of detergent, lipid carriers, and LLOs in the reaction mixture [8]. Light scattering or small angle X-ray scattering measurements can lend insight into lipid phase behavior with varying compositions, for example dispersed, micelle, or bicelle behavior [83]. Current theory predicts that purified membrane proteins are present in detergent micelles, but it remains unclear how LLOs integrate into these micelles and bind to OSTs. Detailed investigation of lipid exchange kinetics with the use of radiolabeled lipids can lend insight into interactions between LLOs and OSTs *in vitro* [84]. Such experiments, however are difficult and not suited to work in the DeLisa lab.

Detergent identity is easily investigated with modifications to the buffers used to purify protein and the reaction buffer. Other target detergents are DDM [32, 64], zwitterionic detergents and other maltopyranosides [8, 85]. Concentrations in the reaction mixture are trivial to change, and approximate estimates for detergent phase behavior (critical micelle concentration, aggregation number, etc.) available from suppliers lend insight for choosing ranges of conditions. These experiments are to be conducted by creating reaction mixtures with detergent concentrations in three main phase domains: dispersed, micelle, and bicelle.

Lipid carriers (non-glycan-conjugated polyisoprenols) can be isolated from *E. coli* SCM6 harboring and empty pACYC plasmid [32]. Alternate options include *pglΔFEDC* plasmids, which don't express the genes to build undecaprenol-bacilosamine pyrophosphate

[86], yielding unlabeled undecaprenol. Mixing LLO carrying membrane fractions with polyisoprene species lacking glycans allows the concentration of LLO to remain constant while changing the total polyisoprene concentrations. Such variations may alter phase behavior or interactions between polyisoprene and detergent micelle phases.

Sub Aim B: Alternate polyisoprene LLO substrates

In addition to *C. jejuni* and *C. lari pgl* loci, *W. succinogenes* and a few other bacterial *pgl* loci have been cloned into pACYC vectors. Transforming these plasmids into SCM6 cells allows isolation of LLOs from an arbitrary bacterial species. With a large diversity of bacterial glycans, it becomes possible to investigate if OSTs have preference for glycan types. Evidence of glycosylation of O-antigens onto AcrA by *C. jejuni* PglB suggests that glycan size alone does not inhibit glycosylation [87]. However, O-antigens tend to be highly linear and little is known of PglB's ability to use highly branched glycans beyond studies with simple eukaryotic glycans [30]. Full eukaryotic glycans tend to be highly branched, which may create problems. Only limited branching of yeast glycans have been examined so far.

Because antibodies against glycans tend to be highly specific, it may be possible to perform competitive glycosylation and evaluate preference for one glycan over another. However, the highly variable nature of antibodies requires a thorough calibration and quantification to ensure that comparisons are valid. However, if glycans are different sizes, or significantly change peptide migration rate through a gel matrix, it may be easy to evaluate glycosylation efficiency using the simpler peptide fluorescence assay.

Bacterial OSTs could possibly use eukaryotic LLOs *in vitro*. This allows LLOs to be prepared from an organism that offer a higher yield or homogeneity of specific glycans than current bacterial systems, namely yeast. Using the bacterial OST's ability to glycosylate folded proteins and eukaryotic LLOs allows eukaryotic core man3 glycans to be added with high efficiency to purified proteins. Other LLO targets may have glycans that can not presently be synthesized in bacteria, sialylated human glycans, for example.

Specific Aim 3: Glycosylation with therapeutically relevant and labeled sugars

Current projects in the DeLisa lab are focusing on the use of synthetic sugars with interesting functional groups. Azido sugars like Ac4GlcNAz and Ac4GalNAz can be integrated into bacterial N- and O- glycans (Stevenson, unpublished). Such sugars can be functionalized with the use of copper-free click chemistry. Copper-free click chemistry allows labeling of any azide group with a diverse range of chemical labels under physiological conditions. Reaction between a DIFO motif with attached functional group and an azido sugar yields a covalent linkage between sugar and desired functional group.

Sub aim A: glycosylation with eukaryotic sugars

eukaryotic glycans are of particular interest in therapeutic applications. Recent *in vivo* studies have shown that *C. jejuni* PglB is able to glycosylate AcrA with undecaprenol-linked Man3 core eukaryotic glycan [30]. Other findings suggest that *C. jejuni* PglB is able to use dolichol as a substrate, albeit at 25% the rate of undecaprenol-linked glycans [50]. There are thus two ways to isolate LLOs with eukaryotic glycans: synthesis in *E. coli* and extraction from eukaryotic sources.

At present, yields of the Man3 core eukaryotic glycan in *E. coli* are low and homogeneity is poor. This is likely due to poor expression of the *Saccharomyces cerivisae* enzymes responsible for the synthesis [30]. With improved yield of these glycans on LLOs, *E. coli* based synthesis of eukaryotic LLOs will become a viable option. For the foreseeable future, *E. coli* remain a platform suitable for synthesis of bacterial LLOs. Eukaryotes will have to suffice as a source of eukaryotic LLOs.

Methods for isolation of large quantities of relatively high purity and high homogeneity LLOs have long been established [50, 88]. Preliminary results suggest that existing in-house methods are able to isolate LLOs, albeit with low yield and low purity. Refinement of this technique will allow use of OSTs to glycosylate peptides and proteins with

fully formed yeast core glycans. While many bacterial glycans have good antibodies against them, eukaryotic glycans rely on lectins for identification [30].

Sub aim B: labeling with Azido sugars

If LLOs with Azido sugars are viable substrates for OSTs, a number of extraordinary new protein labeling techniques become available. Potential labels include protein tags, fluorescent labels, detergents, and virtually any active group imaginable. Protein tags are useful for capture and purification, and can even be used to create branched proteins or scaffolding enzymes. Virtually any fluorescent tag can be attached to a DIFO motif, allowing specific addition of easily identifiable tags, rapid imaging, and simple measurements of glycosylation efficiency. Conceivably, detergents or acyl groups can be attached specifically to N-glycans to produce acylated proteins that can anchor in cell membranes, or assemble into small clusters.

Sugar labeling offers a number of advantages over current chemical labeling and unnatural amino acid techniques. Chemical labeling tends to be either non-specific for amine or carboxyl groups, or limited to the termini of a protein. Glycosylation ensures that labels are attached within the specific sequence. Unnatural amino acid additions typically rely on an amber-suppressor mutation and have low incorporation rates[89]. By contrast, *in vitro* glycosylation has high efficiency (upwards of 80% for peptides, Figure 6) and high specificity.

Simple demonstrations of *in vitro* labeling of proteins with N-linked azido sugars is performed by isolating LLOs from SCM6 cells harboring the pglΔB plasmid and fed either Ac4GlcNAz or Ac4GalNAz. The resulting LLOs will be used in a peptide reaction mixture in place of wild type *C. jejuni* LLOs. Reactions will be boiled to denature PglB and click chemistry will be employed in the reaction mixture to add a fluorescein dye to the azide group. Peptides will be run on a gel as in established protocols, and imaged using a 2-color capable imaging system to discern between red TAMRA-labeled peptides and green

fluorescein-labeled sugars. Successful labeling will be demonstrated by coinciding red and green bands corresponding to the mass of a glycan conjugated to a peptide.

Sub aim C: labeling with fluorescent labeled sugars

LLOs can be labeled with fluorescent dye molecules before being used as reagents in an *in vitro* glycosylation reaction. In this case, proximity of TAMRA molecules on peptides and fluorescein labeled sugars can be calculated with Forester Resonance Energy Transfer (FRET) measurements. FRET is the non-radiative energy transfer between two chormaphores. Energy transfer from the donor chromaphore is able to excite the receptor, which releases a photon upon relaxation. In practical terms, illumination with a wavelength that excites one chromophore yields fluorescence of the second chromaphore at a rate that is proportional to the distance and alignment of the two molecules, and overlap of the absorbtion spectra. TAMRA and fluorescein are well-characterized FRET pairs, with a R_0 of 55nm, or a distance of roughly 4 amino acids and half the length of the *C. jejuni* glycan. Glycosylation of a TAMRA-labeled peptide with a fluorescein-labeled glycan will yield an increase of FRET signal, allowing real-time measurements of glycosylation rates.

Specific aim 4: Kinetics

Current descriptions of PglB kinetics rely on a simple Micalis-Mentin model [56]. This model ignores the enzyme's dual substrate nature and keeps LLO concentration fixed. A more accurate consideration of the substrate binding kinetics and enzymatic action can give rise to a more detailed understanding of not only overall enzymatic action but also binding domains for specific substrates. At present, it is not known if the order of substrate binding is important. A detailed kinetic understanding that is able to control availabilities of both substrates will be an important step in understanding the precise kinetic mechanism of OSTs. Moreover, a better understanding of kinetic mechanisms allows the consequences of mutations to be better understood.

Advanced optics techniques like fluorescence anisotropy allow measurement of substrate binding in real time. Fluorescence anisotropy measures polarized fluorescence, effectively measuring a molecule's rotational diffusion. Such techniques were used to measure peptide substrate affinity of PglB [56], but have yet to be applied to LLO binding. Use of fluorescence anisotropy with fluorescent labeled peptides and glycans allows calculation of binding rates for both substrates to be calculated. FRET based real-time catalytic rate measurements and fluorescence anisotropy together allow measurement allows the measurement of the binding rates for both substrates, catalytic rates and product release rates. In short, these techniques enable a complete kinetic characterization of OSTs.

The peptide-based assay also allows evaluation of efficiency against rationally designed peptide substrates. Peptides can be designed for predicted binding characteristics like solvation energy and flexibility by selecting amino acid sequences or even peptidomimetic residues for desired binding cleft characteristics. Energy of solvation may be tuned to change the affinity of the peptide for the catalytic cleft of PglB, which allows evaluation of a potential new model for enzyme-substrate interaction.

WORKS CITED

1. Opella, S.J., *Structure determination of membrane proteins in their native phospholipid bilayer environment by rotationally aligned solid-state NMR spectroscopy.* Acc Chem Res, 2013. **46**(9): p. 2145-2153.
2. Tillotson, B.J., et al., *Antibody affinity maturation using yeast display with detergent-solubilized membrane proteins as antigen sources.* Protein Eng Des Sel, 2013. **26**(2): p. 101-112.
3. Jaffee, M.B. and B. Imperiali, *Optimized protocol for expression and purification of membrane-bound PglB, a bacterial oligosaccharyl transferase.* Protein Expr Purif, 2013. **89**(2): p. 241-250.
4. Wacker, M., et al., *N-linked glycosylation in Campylobacter jejuni and its functional transfer into E. coli.* Science, 2002. **298**(5599): p. 1790-1793.
5. Eichler, J., *Extreme sweetness: protein glycosylation in archaea.* Nat Rev Microbiol, 2013. **11**(3): p. 151-156.
6. Hagn, F., et al., *Optimized phospholipid bilayer nanodiscs facilitate high-resolution structure determination of membrane proteins.* J Am Chem Soc, 2013. **135**(5): p. 1919-1925.
7. Lee, A.G., *Lipid-protein interactions in biological membranes: a structural perspective.* Biochim Biophys Acta, 2003. **1612**(1): p. 1-40.
8. Seddon, A.M., P. Curnow, and P.J. Booth, *Membrane proteins, lipids and detergents: not just a soap opera.* Biochim Biophys Acta, 2004. **1666**(1-2): p. 105-117.
9. Hust, M., et al., *Construction of human naive antibody gene libraries.* Methods Mol Biol, 2012. **907**: p. 85-107.
10. Griffin, L. and A. Lawson, *Antibody fragments as tools in crystallography.* Clin Exp Immunol, 2011. **165**(3): p. 285-291.
11. Hunte, C. and S. Richers, *Lipids and membrane protein structures.* Curr Opin Struct Biol, 2008. **18**(4): p. 406-411.
12. Tang, X.-l., et al., *Orphan G protein-coupled receptors (GPCRs): biological functions and potential drug targets.* Acta Pharmacol Sin, 2012. **33**(3): p. 363-371.
13. Indiveri, C., et al., *Strategies of bacterial over expression of membrane transporters relevant in human health: the successful case of the three members of OCTN subfamily.* Mol Biotechnol, 2013. **54**(2): p. 724-736.
14. Miroux, B. and J.E. Walker, *Over-production of proteins in Escherichia coli: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels.* J Mol Biol, 1996. **260**(3): p. 289-298.
15. Norholm, M.H., et al., *Improved production of membrane proteins in Escherichia coli by selective codon substitutions.* FEBS Lett, 2013. **587**(15): p. 2352-8.
16. Kenakin, T. and L.J. Miller, *Seven transmembrane receptors as shapeshifting proteins: the impact of allosteric modulation and functional selectivity on new drug discovery.* Pharmacol Rev, 2010. **62**(2): p. 265-304.
17. Bill, R.M., et al., *Overcoming barriers to membrane protein structure determination.* Nat Biotechnol, 2011. **29**(4): p. 335-340.
18. Mujic-Delic, A., et al., *GPCR-targeting nanobodies: attractive research tools, diagnostics, and therapeutics.* Trends Pharmacol Sci, 2014. **35**(5): p. 247-255.
19. Mullard, A., *Learning from the 2012-2013 class of breakthrough therapies.* Nat Rev Drug Discov, 2013. **12**(12): p. 891-893.
20. Kohler, G. and C. Milstein, *Continuous cultures of fused cells secreting antibody of predefined specificity.* Nature, 1975. **256**(5517): p. 495-497.

21. Green, L.L., et al., *Antigen-specific human monoclonal antibodies from mice engineered with human Ig heavy and light chain YACs.* Nat Genet, 1994. **7**(1): p. 13-21.
22. Schirrmann, T. and M. Hust, *Construction of human antibody gene libraries and selection of antibodies by phage display.* Methods Mol Biol, 2010. **651**: p. 177-209.
23. Doerner, A., et al., *Therapeutic antibody engineering by high efficiency cell screening.* FEBS Lett, 2014. **588**(2): p. 278-287.
24. Karlsson, A.J., et al., *Engineering antibody fitness and function using membrane-anchored display of correctly folded proteins.* J Mol Biol, 2012. **416**(1): p. 94-107.
25. Waraho, D. and M.P. DeLisa, *Versatile selection technology for intracellular protein-protein interactions mediated by a unique bacterial hitchhiker transport mechanism.* Proc Natl Acad Sci U S A, 2009. **106**(10): p. 3692-3697.
26. Berdougo, E., K. Mattia, and B. Doranz, *Isolating Membrane Protein Antibodies: Company Promotes Use of Phage Display with Lipoparticles to Isolate Human Antibodies.* Genetic Engineering and Biotechnology News, 2013. **33**(6).
27. Banik, S.S., E. Berdougo, and B. Doranz, *Antibodies Against Membrane Protein Targets: Lipoparticle Technology Facilitates Discovery in This Challenging Protein Class.* Genetic Engineering and Biotechnology News, 2011. **31**(2).
28. Lizak, C., et al., *N-Linked glycosylation of antibody fragments in Escherichia coli.* Bioconjug Chem, 2011. **22**(3): p. 488-96.
29. Ielmini, M.V. and M.F. Feldman, *Desulfovibrio desulfuricans PglB homolog possesses oligosaccharyltransferase activity with relaxed glycan specificity and distinct protein acceptor sequence requirements.* Glycobiology, 2011. **21**(6): p. 734-42.
30. Valderrama-Rincon, J.D., et al., *An engineered eukaryotic protein glycosylation pathway in Escherichia coli.* Nat Chem Biol, 2012. **8**(5): p. 434-436.
31. Lizak, C., et al., *X-ray structure of a bacterial oligosaccharyltransferase.* Nature, 2011. **474**(7351): p. 350-355.
32. Guarino, C. and M.P. DeLisa, *A prokaryote-based cell-free translation system that efficiently synthesizes glycoproteins.* Glycobiology, 2012. **22**(5): p. 596-601.
33. Rezwan, M. and D. Auerbach, *Yeast "N"-hybrid systems for protein-protein and drug-protein interaction discovery.* Methods, 2012. **57**(4): p. 423-429.
34. Wehrman, T., et al., *Protein-protein interactions monitored in mammalian cells via complementation of beta -lactamase enzyme fragments.* Proc Natl Acad Sci U S A, 2002. **99**(6): p. 3469-3474.
35. Chao, G., et al., *Isolating and engineering human antibodies using yeast surface display.* Nat Protoc, 2006. **1**(2): p. 755-768.
36. Fields, S. and O. Song, *A novel genetic system to detect protein-protein interactions.* Nature, 1989. **340**(6230): p. 245-246.
37. Milligan, G., *G protein-coupled receptor dimerisation: molecular basis and relevance to function.* Biochim Biophys Acta, 2007. **1768**(4): p. 825-35.
38. Milligan, G., *G protein-coupled receptor hetero-dimerization: contribution to pharmacology and function.* Br J Pharmacol, 2009. **158**(1): p. 5-14.
39. You, X., et al., *Intracellular protein interaction mapping with FRET hybrids.* Proc Natl Acad Sci U S A, 2006. **103**(49): p. 18458-18463.
40. Gavin, A.-C., et al., *Functional organization of the yeast proteome by systematic analysis of protein complexes.* Nature, 2002. **415**(6868): p. 141-147.
41. Michnick, S.W., et al., *Universal strategies in research and drug discovery based on protein-fragment complementation assays.* Nat Rev Drug Discov, 2007. **6**(7): p. 569-582.
42. Galarneau, A., et al., *Beta-lactamase protein fragment complementation assays as in vivo and in vitro sensors of protein protein interactions.* Nat Biotechnol, 2002. **20**(6): p. 619-622.

43. Lofdahl, P.A., et al., *Selection of TNF-alpha binding affibody molecules using a beta-lactamase protein fragment complementation assay.* N Biotechnol, 2009. **26**(5): p. 251-9.
44. Kostecki, J.S., et al., *Visualizing interactions along the Escherichia coli twin-arginine translocation pathway using protein fragment complementation.* PLoS One, 2010. **5**(2).
45. Fisher, A.C. and M.P. DeLisa, *Laboratory evolution of fast-folding green fluorescent protein using secretory pathway quality control.* PLoS One, 2008. **3**(6).
46. Koch, H., et al., *Direct selection of antibodies from complex libraries with the protein fragment complementation assay.* J Mol Biol, 2006. **357**(2): p. 427-441.
47. Skretas, G. and G. Georgiou, *Simple genetic selection protocol for isolation of overexpressed genes that enhance accumulation of membrane-integrated human G protein-coupled receptors in Escherichia coli.* Appl Environ Microbiol, 2010. **76**(17): p. 5852-5859.
48. Hu, C.-D., Y. Chinenov, and T.K. Kerppola, *Visualization of interactions among bZIP and Rel family proteins in living cells using bimolecular fluorescence complementation.* Mol Cell, 2002. **9**(4): p. 789-798.
49. Heuer, K.H., et al., *Development of a sensitive peptide-based immunoassay: application to detection of the Jun and Fos oncoproteins.* Biochemistry, 1996. **35**(28): p. 9069-9075.
50. Chen, M.M., et al., *Polyisoprenol specificity in the Campylobacter jejuni N-linked glycosylation pathway.* Biochemistry, 2007. **46**(50): p. 14342-8.
51. Chavan, M., A. Yan, and W.J. Lennarz, *Subunits of the translocon interact with components of the oligosaccharyl transferase complex.* J Biol Chem, 2005. **280**(24): p. 22917-22924.
52. Fisher, A.C., et al., *Production of secretory and extracellular N-linked glycoproteins in Escherichia coli.* Appl Environ Microbiol, 2011. **77**(3): p. 871-881.
53. Lizak, C., et al., *Unexpected reactivity and mechanism of carboxamide activation in bacterial N-linked protein glycosylation.* Nat Commun, 2013. **4**: p. 2627-2627.
54. Ruiz-Canada, C., D.J. Kelleher, and R. Gilmore, *Cotranslational and posttranslational N-glycosylation of polypeptides by distinct mammalian OST isoforms.* Cell, 2009. **136**(2): p. 272-83.
55. Liu, F., et al., *Rationally designed short polyisoprenol-linked PglB substrates for engineered polypeptide and protein N-glycosylation.* J Am Chem Soc, 2014. **136**(2): p. 566-569.
56. Gerber, S., et al., *Mechanism of bacterial oligosaccharyltransferase: in vitro quantification of sequon binding and catalysis.* J Biol Chem, 2013. **288**(13): p. 8849-8861.
57. Whitley, P., I.M. Nilsson, and G. von Heijne, *A nascent secretory protein may traverse the ribosome/endoplasmic reticulum translocase complex as an extended chain.* J Biol Chem, 1996. **271**(11): p. 6241-4.
58. Kowarik, M., et al., *N-linked glycosylation of folded proteins by the bacterial oligosaccharyltransferase.* Science, 2006. **314**(5802): p. 1148-1150.
59. Dell, A., et al., *Similarities and differences in the glycosylation mechanisms in prokaryotes and eukaryotes.* Int J Microbiol, 2010. **2010**: p. 148178-148178.
60. Larkin, A. and B. Imperiali, *The expanding horizons of asparagine-linked glycosylation.* Biochemistry, 2011. **50**(21): p. 4411-26.
61. Nasab, F.P., et al., *All in one: Leishmania major STT3 proteins substitute for the whole oligosaccharyltransferase complex in Saccharomyces cerevisiae.* Mol Biol Cell, 2008. **19**(9): p. 3758-3768.
62. Kowarik, M., et al., *Definition of the bacterial N-glycosylation site consensus sequence.* Embo j, 2006. **25**(9): p. 1957-66.

63. Schwarz, F., et al., *A combined method for producing homogeneous glycoproteins with eukaryotic N-glycosylation.* Nat Chem Biol, 2010. **6**(4): p. 264-266.
64. Li, L., et al., *Overexpression and topology of bacterial oligosaccharyltransferase PglB.* Biochem Biophys Res Commun, 2010. **394**(4): p. 1069-74.
65. Schagger, H., *Tricine-SDS-PAGE.* Nat Protoc, 2006. **1**(1): p. 16-22.
66. Hese, K., et al., *The yeast oligosaccharyltransferase complex can be replaced by STT3 from Leishmania major.* Glycobiology, 2009. **19**(2): p. 160-171.
67. Crichton, P.G., et al., *Lipid, detergent, and Coomassie Blue G-250 affect the migration of small membrane proteins in blue native gels: mitochondrial carriers migrate as monomers not dimers.* J Biol Chem, 2013. **288**(30): p. 22163-22173.
68. Loster, K., et al., *Chemical cross-linking leads to two high molecular mass aggregates of rat alpha 1 beta 1 integrin differing in their conformation but not in their composition.* FEBS Lett, 1995. **373**(3): p. 234-8.
69. Gilbreath, J.J., et al., *Change is good: variations in common biological mechanisms in the epsilonproteobacterial genera Campylobacter and Helicobacter.* Microbiol Mol Biol Rev, 2011. **75**(1): p. 84-8132.
70. Laughlin, S.T., et al., *Metabolic labeling of glycans with azido sugars for visualization and glycoproteomics.* Methods Enzymol, 2006. **415**: p. 230-250.
71. Kelleher, D.J. and R. Gilmore, *An evolving view of the eukaryotic oligosaccharyltransferase.* Glycobiology, 2006. **16**(4): p. 62.
72. O'Brien, S.P. and M.P. Delisa, *Split-Cre recombinase effectively monitors protein-protein interactions in living bacteria.* Biotechnol J, 2014. **9**(3): p. 355-61.
73. Aronson, D.E., L.M. Costantini, and E.L. Snapp, *Superfolder GFP is fluorescent in oxidizing environments when targeted via the Sec translocon.* Traffic, 2011. **12**(5): p. 543-548.
74. Zhou, J., et al., *An improved bimolecular fluorescence complementation tool based on superfolder green fluorescent protein.* Acta Biochim Biophys Sin (Shanghai), 2011. **43**(3): p. 239-244.
75. Link, A.J., et al., *Efficient production of membrane-integrated and detergent-soluble G protein-coupled receptors in Escherichia coli.* Protein Sci, 2008. **17**(10): p. 1857-63.
76. Engvall, E. and P. Perlmann, *Enzyme-linked immunosorbent assay (ELISA). Quantitative assay of immunoglobulin G.* Immunochemistry, 1971. **8**(9): p. 871-874.
77. Gai, S.A. and K.D. Wittrup, *Yeast surface display for protein engineering and characterization.* Curr Opin Struct Biol, 2007. **17**(4): p. 467-73.
78. Homola, J., *Surface plasmon resonance sensors for detection of chemical and biological species.* Chem Rev, 2008. **108**(2): p. 462-93.
79. Navratilova, I., J. Besnard, and A.L. Hopkins, *Screening for GPCR Ligands Using Surface Plasmon Resonance.* ACS Med Chem Lett, 2011. **2**(7): p. 549-554.
80. Stenlund, P., et al., *Capture and reconstitution of G protein-coupled receptors on a biosensor surface.* Anal Biochem, 2003. **316**(2): p. 243-250.
81. Rich, R.L., et al., *Biacore analysis with stabilized G-protein-coupled receptors.* Anal Biochem, 2011. **409**(2): p. 267-272.
82. Patching, S.G., *Surface plasmon resonance spectroscopy for characterisation of membrane protein-ligand interactions and its potential for drug discovery.* Biochim Biophys Acta, 2014. **1838**(1 Pt A): p. 43-55.
83. Thompson, A.A., et al., *GPCR stabilization using the bicelle-like architecture of mixed sterol-detergent micelles.* Methods, 2011. **55**(4): p. 310-317.
84. McLean, L.R. and M.C. Phillips, *Kinetics of phosphatidylcholine and lysophosphatidylcholine exchange between unilamellar vesicles.* Biochemistry, 1984. **23**(20): p. 4624-30.
85. Chung, K.Y., et al., *Role of detergents in conformational exchange of a G protein-coupled receptor.* J Biol Chem, 2012. **287**(43): p. 36305-36311.

86. Glover, K.J., et al., *Chemoenzymatic synthesis of glycopeptides with PglB, a bacterial oligosaccharyl transferase from Campylobacter jejuni.* Chem Biol, 2005. **12**(12): p. 1311-5.
87. Feldman, M.F., et al., *Engineering N-linked protein glycosylation with diverse O antigen lipopolysaccharide structures in Escherichia coli.* Proc Natl Acad Sci U S A, 2005. **102**(8): p. 3016-3021.
88. Kelleher, D.J., D. Karaoglu, and R. Gilmore, *Large-scale isolation of dolichol-linked oligosaccharides with homogeneous oligosaccharide structures: determination of steady-state dolichol-linked oligosaccharide compositions.* Glycobiology, 2001. **11**(4): p. 321-333.
89. Ryu, Y. and P.G. Schultz, *Efficient incorporation of unnatural amino acids into proteins in Escherichia coli.* Nat Methods, 2006. **3**(4): p. 263-5.
90. Tamura, K., et al., *MEGA6: Molecular Evolutionary Genetics Analysis version 6.0.* Mol Biol Evol, 2013. **30**(12): p. 2725-9.